Journal of Technology Law & Policy

Volume 28 Fall 2023 Number 1

ARTICLES

NOTES

Journal of Technology Law & Policy

Volume 28 Fall 2023 Number 1

THE PEOPLE'S WAR AND ITS APPLICATION TO CHINA'S LEGAL FRAMEWORK FOR CYBERSECURITY

Captain Christopher J. Lin[*]

Abstract

This Article addresses the growing threat of cyberattacks on critical infrastructure by examining China's response, particularly through its Cybersecurity Law (CSL), against the backdrop of global cybersecurity laws like the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA). The CSL, enacted in 2016, is analyzed within the context of Chinese military doctrine, specifically, the concept of the People's War introduced by Mao Zedong. Part I traces the historical evolution of the People's War, Part II explores its continued relevance in cyberspace, and Part III discusses how the People's War elements manifest in the CSL and related regulations. This Article argues that the CSL focuses on elevating China's defensive cyber capabilities across governmental and consumer sectors, diverging from the more consumer-privacy-centric approach of other global cybersecurity laws. Part IV delves into the challenges the United States faces in responding to the CSL and suggests potential paths forward to bridge strategic divides between the two countries in the realm of cyberspace. The introduction vividly portrays real-world scenarios of cyberattacks impacting critical infrastructure, setting the stage for the exploration of China's unique response in the subsequent sections.

* Judge Advocate, United States Army. Presently assigned as an LL.M. Candidate to Georgetown University Law Center. J.D., 2017, UCLA School of Law; B.A., 2013, University of California, San Diego. Member of the D.C. Bar. The author thanks Major D. Nicholas Allen, Brigade Judge Advocate, 1st Security Force Assistance Brigade, for his continuous support and encouragement, along with his insightful feedback and contributions. This Article is rooted in his infectious passion for scholarship and operational law. The views and opinions presented herein are those of the author and do not necessarily represent the views of the United States Government, the Department of Defense (DoD), or its components. Appearance of, or reference to, any commercial products or services does not constitute DoD endorsement of those products or services. The appearance of external hyperlinks does not constitute DoD endorsement of the linked websites, or the information, products, or services therein.

INTRODUCTION

A sudden power outage left a quarter-million residents without power or heat.[1] Telecommunication outages that rendered phone calls and data access impossible.[2] A control system failure that released raw sewage across public grounds.[3] Each scenario has occurred in the real world in the past two decades due to cyberattacks. In the United States, former Defense Secretary Leon Panetta voiced his concerns about a "cyber-Pearl Harbor" in which "aggressors can launch attacks with cyber-tools to gain control of our nation's critical infrastructure . . . causing physical destruction and loss of life on a scale that 'would paralyze and shock the nation.'"[4]

And China has observed and formulated a response to these concerns. On a cool Wednesday morning in the autumn of 2016, hundreds of attendees—including politicians and representatives from major technological companies—sat in plush white leather auditorium chairs, peering up at a video link projected behind a podium. Framed against a mahogany background and the striking red and gold colors of the Chinese

1. Sean Lyngaas, *Russian Military-Linked Hackers Target Ukrainian Power Company, Investigators Say*, CNN, Apr. 14, 2022, https://www.cnn.com/2022/04/12/politics/gru-russia-hackers-ukraine-power-grid/index.html [https://perma.cc/Y8KV-LRTV].

2. Kate Fazzini, *Power Outages, Bank Runs, Changed Financial Data: Here are the "Cyber 9/11" Scenarios that Really Worry the Experts*, CNBC (Nov. 18, 2018), https://www.cnbc.com/2018/11/18/cyber-911-scenarios-power-outages-bank-runs-changed-data.html [https://perma.cc/T4PA-V7BJ].

3. Tony Smith, *Hacker Jailed for Revenge Sewage Attacks*, REGISTER (Oct. 31, 2001), https://www.theregister.com/2001/10/31/hacker_jailed_for_revenge_sewage [https://perma.cc/85WS-HHYH].

4. Robert K. Palmer, *Critical Infrastructure: Legislative Factors for Preventing a "Cyber-Pearl Harbor*," 18 VA. J.L. & TECH. 289, 293–94 (2014).

flag, Xi Jinping, the general secretary of China, presented his opening remarks at the Wuzhen Summit, noting the importance of international cooperation in building a community in cyberspace while also ensuring inclusiveness and security.[5] His comments followed the National People's Congress's enactment of China's Cybersecurity Law (CSL) earlier that month, which would drive dialogue on the increasing awareness of cybersecurity and data rights, along with a rush by corporations to comply with the law.[6]

While the CSL was promulgated alongside a number of cyber-related laws across the globe in the past decade, including the General Data Protection Regulation (GDPR)[7] and the California Consumer Privacy Act (CCPA),[8] the CSL differs in that it places a distinct focus on elevating the country's defensive cyber capabilities across governmental and consumer sectors, as opposed to a more singular focus on consumer privacy. This Article argues that the CSL can be viewed within the framework of Chinese military doctrine—specifically, the CSL retains key elements of the People's War, a concept discussed by Mao Zedong, the founder of the People's Republic of China. Part I traces the evolution and legacy of the People's War from its origins in Mao's writings in the early 1900s to modern-day applications. Part II examines cyberspace as a new warfighting domain, with the People's War enjoying continued relevance. Part III discusses aspects of the People's War as they apply to the CSL and its surrounding regulations. Part IV explores the challenges that the United States faces in creating a balanced response to the CSL and a possible path forward in bridging the divide between the two countries' strategic approaches to cyberspace.

I. THE LEGACY OF THE PEOPLE'S WAR

Modern Chinese doctrine underwent numerous shifts within the past century, largely in response to external threats such as the Second Sino-Japanese War during World War II and observations of the Gulf War. These shifts can be broadly understood as three periods of differing focal points. Mobilization of the masses under the People's War was prominent

---

5. *Di San Jie Shijie Hulianwang Dahui* (*第三届世界互联网大会*) [*Third World Internet Conference*], YOUTUBE (Nov. 16, 2016), https://www.youtube.com/watch?v=cawjSOpXP-4 [https://perma.cc/577W-U62L].

6. *See, e.g.*, Huifeng He, *Cybersecurity Law Causing "Mass Concerns" Among Foreign Firms in China*, SCMP (Mar. 1, 2018), https://www.scmp.com/news/china/economy/article/2135338/cybersecurity-law-causing-mass-concerns-among-foreign-firms-china [https://perma.cc/N2WQ-CQPE].

7. *Data Protection in the EU*, EUROPEAN COMM'N, https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en [https://perma.cc/NU5F-P4WA].

8. *California Consumer Privacy Act (CCPA)*, STATE OF CALIFORNIA DEP'T OF JUST., https://oag.ca.gov/privacy/ccpa [https://perma.cc/3QTW-ZWL2].

from the early 1900s to the 1970s.[9] The People's War was then enveloped under the umbrella of active defense from the late 1970s until the early 1990s.[10] Finally, China focused on warfighting under informal conditions from approximately the early 1990s onwards.[11] However, active defense remains an important underpinning of Chinese military doctrine and strategic policy and is "a fundamentally defensive political and strategic stance, enabled—when required—by operational and tactical offense," characterized by multi-layered defenses that an adversary must overcome, alongside a no first-strike policy.[12] Because such multi-layered defenses call for leveraging the skills of the civilian populace, a core component of active defense is the concept of the People's War, which also remains an underlying principle, even with doctrinal shifts throughout the years. The People's War developed from its roots as a struggle against the gentry to its modern iteration of a civil-military fusion that preserves certain key traits of the original idea, including capitalizing on asymmetrical advantages and sustaining a protracted war.

A. *Mobilizing the Masses*

Pre-1949, China's military doctrine was largely rudimentary, with basic military schooling and doctrinal development that often drew from foreign sources, though Mao Zedong's philosophy also developed during this time period and is closely associated with the People's War.[13] The concept of the People's War in China can be found as early as 1927, during which Mao identified the potential power in leveraging the peasant population in Hunan in a revolutionary struggle against the gentry, deemed a powerful, oppressive social class that needed to be overthrown to ensure to the well-being of the masses.[14]

Initially, at an operational level, the emphasis was on organizing and consolidating the strength of the masses against an enemy so that "several hundred million peasants will rise like a mighty storm, like a hurricane, a force so swift and violent that no power, however great, will be able to

---

9. 1 MAO ZEDONG, SELECTED WORKS OF MAO TSE-TUNG 23 (1st ed. 1965); NGOK LEE, THE CHINESE PEOPLE'S LIBERATION ARMY 1980-82: MODERNISATION, STRATEGY AND POLITICS 50–51 (1983).

10. LEE, *supra* note 9, at 50–51; M. TAYLOR FRAVEL, CHINA'S MILITARY STRATEGY SINCE 1949 220 (2019).

11. FRAVEL, *supra* note 10, at 220.

12. *Zhongguo de Junshi Zhanlue* (*中国的军事战略*) [*China's Military Strategy*], GUOWUYUAN XINWEN BANGONGSHI (国务院新闻办公室) [STATE COUNCIL INFORMATION OFFICE], June 2015, http://www.scio.gov.cn/zfbps/ndhf/2015/Document/1435161/1435161.htm [https://perma.cc/G75R-BYZB]; Chinese Tactics, Army Techniques Publication, No. 7-100.3, 1-7 (Aug. 2021).

13. KA PO NG, INTERPRETING CHINA'S MILITARY POWER: DOCTRINE MAKES READINESS 49 (1st ed. 2004); FRAVEL, *supra* note 10, at 220.

14. ZEDONG, *supra* note 9.

hold it back."[15] Mao further noted in 1938, "Mobilization of the common people will create a vast sea in which to drown the enemy, create the conditions that will make up for our inferiority in arms and other things, and create the prerequisites for overcoming every difficulty in the war."[16]

While Mao reiterated the concept of the People's War throughout his written works, the term itself did not officially appear until Mao's political report to the Seventh National Congress of the Communist Party of China in 1945.[17] In his report, Mao stated that "all the anti-Japanese people in the Liberated Areas of China are called upon to join organizations of workers, peasants, youth and women, and cultural, professional and other organizations, which will wholeheartedly perform various tasks in support of the armed forces . . . [s]uch is a real people's war."[18] Additionally, Mao anticipated that a People's War would be a protracted war, one that—even where enemy forces struck deep into the mainland—there would be constant pockets of resistance, as the mobilized masses would gradually reinforce its main fighting effort to strain the enemy "under the trial of innumerable battles."[19] Though warfare in China shifted from a revolutionary movement against the gentry in Hunan to national liberation from the Japanese under the Second Sino-Japanese War, the core concept of the People's War remained the same—mobilization of the masses against a superior enemy to mitigate imbalances in military strength and to supplement the conventional army's warfighting functions in areas such as intelligence and logistical support.

B. *Active Defense*

The People's War shifted to a national strategic level by the late 1970s to the early 1980s, under the guideline of active defense, in response to the threat of a Soviet incursion into China and the 1973 Arab-Israeli War, wherein the United States and the Soviet Union employed advanced weaponry, marking a shift in the modernization of warfare.[20] The Central Military Commission (CMC) approved active defense in 1980 and focused on establishing a multi-layered defense—to include forward defensive positions—that the enemy must overcome so that China has time to mobilize its forces.[21]

However, the People's War remained necessary due to concerns regarding asymmetrical capabilities against adversaries. Specifically, the

15. *Id.*
16. 2 MAO ZEDONG, SELECTED WORKS OF MAO TSE-TUNG 154 (1st ed. 1965).
17. 3 MAO ZEDONG, SELECTED WORKS OF MAO TSE-TUNG 213 (1st ed. 1965).
18. *Id.* at 216–17.
19. MAO, *supra* note 19, at 188.
20. LEE, *supra* note 9, at 50–51; FRAVEL, *supra* note 10, at 456.
21. FRAVEL, *supra* note 10, at 454–66.

Soviet Union's defense capabilities surpassed China's during this time, and cuts to China's defense budget in favor of economic development further hindered the country.[22] In the post-Mao era and in recognition of the evolution of warfare, Deng Xiaoping preserved the link to Mao's interpretation of the People's War, emphasizing that "we can defeat a superior enemy with inferior equipment, for our wars are just, they are people's wars."[23] Under Deng's leadership, active defense focused on conventional forces that were directly supported by the mobilized masses in the form of militia.[24] For example, during this time, sixty percent of the People's Liberation Army relied on austere support systems, which materialized as regional militia providing logistical support by drawing from local resources such as truck transportation.[25] Active defense thus marks a strategic shift in warfighting philosophy that continues to this day; civil resources directly supplement the conventional military as an integral part of combat operations, resulting in a deterring effect given the whole-of-society approach and layered defenses.[26]

C. *Modern Defense Approaches*

With the advent of modern technology in the 1990s, the People's War transformed again into a concept that promoted close integration of the military and civilian sectors, with the rationale of fortifying military strength with commercial capabilities.[27] The CMC adopted a new strategic guideline in 1993 titled "winning local wars under modern,

22. LEE, *supra* note 9, at 50. Of note, in the late 1970s, a point of critique was whether the People's War was still relevant in light of future wars given technological advancements. Given such advancements, tactics that were previously successful, e.g., throwing grenades into sight openings on armored vehicles, may no longer be valid. Indeed, Su Yu, the commissar and party secretary of the Academy of Military Science beginning in 1972, felt that the People's War had largely been relegated to an abstract slogan. FRAVEL, *supra* note 10, at 472–75.

23. DENG XIAOPING, SELECTED WORKS OF DENG XIAOPING: VOLUME II (1975-1982) (1995).

24. FRAVEL, *supra* note 10, at 230. Beginning in 1978, the Central Committee of the Chinese Communist Party renewed focus on mobilization of people in warfare as militiamen, formalizing training and doctrine, e.g., having a separation of roles for urban and rural militia, where—for instance—the main effort for urban militia would be to construct city defenses. LEE, *supra* note 9, at 80–81.

25. LEE, *supra* note 9, at 73–75.

26. Chinese Tactics, *supra* note 12, at 1–7.

27. The term military-civil fusion and its various iterations can be found as far back as the Mao Zedong era as the basis of the People's War, i.e., making use of the civilian sector for warfighting, but in contrast to its initial inception that focused on mobilization of the peasantry, military-civil fusion in the modern day identifies the need for a symbiotic relationship between the military and civilian sectors, particularly within areas of technological development in which military and civilian technology should be mutually compatible. Jiang Ying (江英), *Jicheng Fazhan Junmin Shendu Ronghe Guangrong Chuantong* (*继承发展军民深度关荣传统*) [*Inherit and Develop the Glorious Tradition of Deep Military-Civil Fusion*], GUANGMING RIBAO (光明日报) [GUANGMING DAILY] July 18, 2017, https://epaper.gmw.cn/gmrb/html/2017-07/18/nw.D110000gmrb_20170718_2-02.htm [https://perma.cc/6PM8-6HR4].

high-technology conditions," largely in response to the Gulf War, which saw the use of precision-guided munitions, again signaling a shift in the advancement of warfare—in particular, technological augments to maneuver forces.[28] Indeed, against predictions by Chinese military analysts that the Gulf War would result in a protracted war, the United States and allied countries defeated the Iraqi military within one hundred hours from the start of the conflict, thus serving as a catalyst for change in Chinese military strategy.[29]

1. Warfighting Under "Informatization"

The 1993 guideline focused on developing a new approach to warfighting that combined an array of systems, e.g., precision-guided weapons, intelligence, and electronic, given that modern warfare was no longer strictly confined to targeting the forward line of troops or the support area; instead, attacks could also target information hubs and operational systems.[30] However, the 1993 guideline nevertheless remained rooted in the concept of active defense, honing in on regional, localized disputes along China's borders and regions, e.g., Taiwan, as opposed to a broader enemy invasion of mainland China.[31] Zhang Wangnian, the general chief of staff, acknowledged the challenges of warfighting given new technologies and the struggles of "being rooted in using inferior equipment to defeat an enemy."[32]

To remedy these obstacles while also adhering to active defense as a foundational strategy, Zhang proposed an emphasis on the mobility of naval, air, and missile forces to rapidly react to threats in addition to the development of advanced weaponry.[33] China further made minor adjustments to its military strategy in 2004 and 2014, with the 2004 strategy focusing on addressing informatization—the prevalence of information technology throughout all aspects of military operations—and the 2015 strategy focusing on integrated joint operations in addition to informatization, marking the continued recognition of the importance of technology and information.[34]

2. Civil-Military Fusion

While military strategies from 1993 onwards placed an emphasis on conventional military and multi-domain operations, the People's War remained a crucial principle, evolving from the organization of the

28. FRAVEL, *supra* note 10, at 590–94.
29. *Id.* at 608–09.
30. *Id.* at 618.
31. *Id.* at 599–600.
32. *Id.* at 651.
33. *Id.* at 651–52.
34. *Id.* at 699–702.

masses in the early 1900s into the concept of military-civil fusion in the modern day, echoing the whole-of-society approach of active defense.[35]

Military-civil fusion has numerous concepts, connotations, and nuances as it developed over a number of years but can be broadly understood as the integration and pairing of the civilian sector with the military with the goal of more effective warfighting.[36] One such term for military-civil fusion is *junmin jiehe*, or "combining the military and civilian sectors," which originated with Deng in 1978 as a strategy whereby—in a hands-off approach—the government encouraged the development of dual-use technologies in the 1980s.[37] Crucially, in the 1990s, alongside the 1993 strategic guideline, the government began to take an active role in the development of dual-use technologies, such as by providing defense firms with financial assistance and appropriate networking for creating such technologies.[38] Finally, the late 1990s saw a further increase in the importance of integrating the military and civilian sectors, as demonstrated by one of the key policy objectives of the Commission of Science, Technology, and Industry for National Defense, which highlighted "two-way civil-military technology cooperation, transfers, promotions, and joint development."[39]

The People's War ultimately persisted within Chinese military strategy across two centuries. It underwent a transformation from more political origins in leveraging the proletariat against the gentry, to a whole-of-society, layered defense approach that uses a close relationship between the military and civilian sectors, especially concerning technological integration. Modern Chinese doctrine retains aspects of the People's War, including asymmetrical warfare and "long-term combat [that] consumes the enemy in protracted contests."[40]

II. THE ADVENT OF CYBERSPACE

The development of cyberspace further changed the nature of warfare and is now largely considered a new warfighting domain or dimension.[41]

---

35. *Id.* at 231.

36. ALEX STONE, MILITARY-CIVIL FUSION TERMINOLOGY: A REFERENCE GUIDE 6–8 (2021).

37. *Junmin jiehe* contained four key principles: (i) developing dual-use technologies, (ii) ensuring that peacetime development took into account wartime mobilization, (iii), prioritizing military research and development in the civilian economy, and (iv) allowing the military to benefit from the effects of economic prosperity. TAI MING CHEUNG, FORTIFYING CHINA: THE STRUGGLE TO BUILD A MODERN DEFENSE ECONOMY 8 (2009).

38. CHEUNG, *supra* note 37, at 8.

39. *Id.*

40. *See generally* JUNSHI KEXUEYUAN (军事科学院) [ACADEMY OF MILITARY SCIENCE], ZHANLUE XUE (战略学) [SCIENCE OF MILITARY STRATEGY] (2020).

41. "Cyberspace is a global domain within the information environment consisting of interdependent networks of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and embedded processors and

China has recognized the importance of cyberspace, particularly in light of the proliferation of network connectivity and threats from geopolitical rivals and hackers. Indeed, the United States elevated its Cyber Command to the tenth Combatant Command in 2018.[42] Other countries, including the United Kingdom, France, and Japan, have followed suit, bringing cyber capabilities to a national strategic level and holding exercises that test offensive and defensive capabilities in cyberspace.[43] With over 730 million mobile internet users and 1.94 million network terminals infected with viruses per month within China, cyberspace has become a major concern for Chinese national security.[44] Despite the unique characteristics of cyberspace as a warfighting domain, the concept of the People's War under active defense holds lasting relevance in understanding China's approach to cyber operations.

In the past three decades, China has placed increased importance on informatization and cybersecurity in recent years. In China, cyber operations fall under the broader umbrella of information operations, which also includes other functions such as military information support operations and electronic warfare.[45]

Chinese scholars and military institutions have long forecasted cyberspace to be a new warfighting domain. A People's Liberation Army publication noted in 2006 that the twenty-first century is the century of information warfare; with over 170 countries and regions connected via computer networks, which can be attacked, cyberspace is the new combat space.[46] The Academy of Military Science further emphasized that

---

controllers." Operations, Field Manual No. 3-0, paragraph 1–31 (Dec. 6, 2017). "Cyberspace is highly vulnerable for several reasons, including ease of access, network and software complexity, lack of security considerations, in network design and software development, and inappropriate user activity." *Id.* at 1–33.

42. Li Minghai (李明海), *Wangluo Xinxi Tixi Junmin Ronghe Zhanlue de Sikao* (*网络信息体系军民融合战略的思考*) [*Reflections on the Strategy of Civil-Military Integration of Information Network Systems*], WANGLUO CHUANBO ZAZHI (网络传播杂志) [J. NETWORK COMMUNICATION], June 12, 2018, http://www.cac.gov.cn/2018-11/12/c_1123701001.htm [https://perma.cc/5ZFT-AXZC].

43. *Id.*

44. *Id.*

45. AMY CHANG, WARRING STATE: CHINA'S CYBERSECURITY STRATEGY 13 (2014). There are grounds for noting a possible semantic distinction in China's use of the word "cyber." *Id.* Because references to the cyber domain are noted in terms of *wangluo*, or network, in China, some scholars argue that network security or network space are more appropriate terms to avoid possible divergences in meaning. *Id.* In common parlance, however, Chinese media largely does not make the same distinction between the two terms. *See, e.g.*, *China's First Data Security Law and its Wider Impact*, CGTN, Sept. 7, 2021, https://news.cgtn.com/news/2021-09-07/China-s-first-data-security-law-and-its-wider-impact-13lgE8ufFsI/index.html [https://perma.cc/R6HV-RCVY].

46. *Lun Xin Shiji Xin Jieduan Wo Jun De Lishi Shiming* (*论新世纪新阶段我军的历史使命*) [*Regarding the Historical Mission of our Army in the New Century and Era*], Jiefangjun Bao

developing cyber capabilities is a priority, particularly because networks inevitably have vulnerabilities, and cyber defense can be difficult because of the numerous vulnerabilities that have yet to be identified.[47]

China's cyber concerns are elevated in light of numerous cybersecurity incidents, ranging from small data breaches to attacks on government networks. In 2011, unidentified foreign entities used the Indian government's National Informatics Centre servers to attack Chinese government servers.[48] In 2020, the coronavirus pandemic leaked the personal information of four to five hundred travelers from Wuhan, China, after submission to regulators and transportation entities.[49] More recently, in 2022, unknown hackers stole over 23 terabytes of personal information from the Shanghai police database, resulting in the largest cyberattack in Chinese history.[50] Over the past few years, China incurred over 2,700 advanced cyberattacks against a wide range of industries, spanning from scientific research institutions to major internet companies.[51]

The People's War persists even within the realm of cyberspace through the framework of active defense. The Science of Military Strategy (SMS), a doctrinal publication of the People's Liberation Army, addressed guidance for cyberspace for the first time in its 2013 edition and reiterated the concept of active defense.[52] The SMS contrasted China's military deterrence with those of Western countries, noting that rather than projecting military power to further global hegemony, China is defensively postured—adhering to the concept of active defense to

---

(解放军报) [PLA DAILY], Jan. 9, 2006, http://news.sohu.com/20060109/n241350798.shtml [https://perma.cc/LVH8-K9FR].

47. JUNSHI KEXUEYUAN (军事科学院) [ACADEMY OF MILITARY SCIENCE], *supra* note 40, ZHANLUE XUE (战略学) [SCIENCE OF MILITARY STRATEGY] 193 (2013).

48. Josy Joseph, *Govt Servers Used for Cyber Attacks on China, Other Countries' Networks*, TIMES OF INDIA (Nov. 17, 2011), https://timesofindia.indiatimes.com/tech-news/govt-servers-used-for-cyber-attacks-on-china-other-countries-networks/articleshow/10760699.cms [https://perma.cc/2EPB-55CF].

49. Yan Luo, *Cyberspace Administration of China Releases Notice on the Protection of Personal Information in the Fight Against Coronavirus*, INSIDE PRIVACY (Feb. 11, 2020), https://www.insideprivacy.com/international/china/cyberspace-administration-of-china-releases-notice-on-the-protection-of-personal-information-in-the-fight-against-coronavirus/ [https://perma.cc/98XV-MGND].

50. *China's Cabinet Stresses Cybersecurity After Data Leak*, BLOOMBERG (July 6, 2022), https://www.bloomberg.com/news/articles/2022-07-07/china-s-cabinet-urges-greater-cyber security-after-mass-data-leak [https://perma.cc/6JL4-7XJF].

51. *Over 2,700 Cyber Attacks Launched Against China, Chinese Security Company 360 Found*, GLOBAL TIMES (Mar. 4, 2021), https://www.globaltimes.cn/page/202103/1217364.shtml [https://perma.cc/P2EB-E2GK].

52. JUNSHI KEXUEYUAN (军事科学院) [ACADEMY OF MILITARY SCIENCE], *supra* note 40, at 145.

contain crisis and counteract invasion actions from other countries that may infringe on China's interests.[53]

In both the 2013 and 2020 editions of SMS, active defense entailed close cooperation between the political and civilian fields and the differing warfighting functions.[54] To this end, the General Secretary of the Chinese Communist Party, Xi Jinping, stated in a 2016 conference that the party, the country, the army, and individuals of all ethnic groups should move forward with one heart and one mind to overcome obstacles, setting forth another iteration of *junmin jiehe*, an echo of the People's War.[55] Specifically, within the cyber realm, Li Minghai, the deputy director of the War and Crisis Response Training Center, noted that close integration of military and civilian information systems is the foundation of victory in that it creates a joint force to respond to threats against networks.[56] To address cyberspace's challenges as a new warfighting domain, China thus continued its doctrinal legacy of a whole-of-country approach in unifying the military and civilian sectors to ensure a multi-layered defense.

III. CHINA'S CYBERSECURITY LAW AND IMPLEMENTING REGULATIONS

In conjunction with rising cybersecurity concerns and challenges, China has promulgated laws and guiding strategies to shape and secure its interests in cyberspace, with the view that there is no national security without cybersecurity.[57] As early as 2003, China published Document 27, also known as the Opinions of the Leading Group for Strengthening Information Security Assurance Work, which laid the groundwork for dynamic monitoring of the internet and protecting critical infrastructure.[58] By 2011, China's foray into data security at the national level was imminent, as the Ministry of Information and Industry Technology, China's internet regulator, issued guidelines for protecting

53. *Id.*

54. JUNSHI KEXUEYUAN (军事科学院) [ACADEMY OF MILITARY SCIENCE], *supra* note 28, at 33 (2020); JUNSHI KEXUEYUAN (军事科学院) [ACADEMY OF MILITARY SCIENCE], *supra* note 40, at 148.

55. *Xi Jinping: Jianchi Jun di Heli Junmin Tongxin Quanmian Tigao Shuang Yong Gongzuo Shuiping* (*习近平：坚持军地合力军民同心 全面提高双拥工作水平*) [*Xi Jinping: Sustain the United Efforts of the Military and the People, Raise the Quality of Dual-Use Efforts*], XINHUA SHE (新华社) [XINHUA NEWS], July 29, 2016, http://www.xinhuanet.com/politics/2016-07/29/c_1119306354.htm [https://perma.cc/3XH9-HWL6].

56. Li Minghai (李明海), *supra* note 42.

57. Hawke Johannes Gierow, *Cybersecurity in China: New Political Leadership Focuses on Boosting National Security*, MERCATOR INST. FOR CHINA STUDIES: CHINA MONITOR 2 (Dec. 9, 2014), https://merics.org/sites/default/files/2020-05/China_Monitor_20_Cyber_Security-National_Security_EN.pdf [https://perma.cc/UY8E-SHLS].

58. Adam Segal, *China Moves Forward on Cybersecurity Policy*, CFR (June 24, 2012), https://www.cfr.org/blog/china-moves-forward-cybersecurity-policy [https://perma.cc/RU7S-UBEV].

personal information; though the guidelines did not have the force of law, they nevertheless paved the way for a legal regime that responds to the evolving cyber environment via national standards.[59]

In 2016, the National People's Congress enacted the CSL, which came into effect in 2017 and was a landmark legislation that aimed to strengthen data protection to further national security. Importantly, it is the "first Chinese law that systematically lays out the regulatory requirements on cybersecurity, subjecting many previously under-regulated or unregulated activities in cyberspace to government scrutiny."[60]

In contrast to other data protection regulations, such as the GDPR or CCPA, which emphasize privacy and personal information protection, the CSL's foremost focus is on national security.[61] For example, the CSL seeks to impose security obligations on network operators, critical information infrastructure, and cross-border transfers of data; the broad applicability of concepts and terms within the CSL has the effect of exerting more control over data and information infrastructure, both foreign and domestic.[62] The CSL is accompanied by numerous other regulations that further clarify differing aspects and definitions within the field of data security. In particular, the Data Security Law regulates data processing activities with implications on national security, and the Personal Information Protection Law governs the protection of personal information, thereby "form[ing] an over-arching framework that will govern data protection and cybersecurity in China for years to come."[63]

The promulgation of the CSL and its implementing regulations drew a quick response from multinational corporations, particularly those with

---

59. *Release of China's First Personal Information Protection Standards Imminent*, INSIDE PRIVACY (Aug. 8, 2011), https://www.insideprivacy.com/international/release-of-chinas-first-personal-information-protection-standards-imminent [https://perma.cc/75MQ-B3C6].

60. *China Passes New Cybersecurity Law*, INSIDE PRIVACY (Nov. 8, 2016), https://www.cov.com/-/media/files/corporate/publications/2016/11/china_passes_new_cybersecurity_law.pdf [https://perma.cc/XYM4-NJ9Y].

61. The United States has responded to cyber challenges through executive and legislative means, such as the Biden Administration's Executive Order entitled "Improving the Nation's Cybersecurity," and noted the need for better communication between the public and private sectors, but the area of cybersecurity and privacy is largely covered by a patchwork of laws at the state and federal level, as opposed to having a broad, unified standard. Alan C. Raul and Snezhana S. Tapia, *In a Nutshell: Data Protection, Privacy, and Cybersecurity in USA*, LEXOLOGY (Nov. 5, 2021), https://www.lexology.com/library/detail.aspx?g=1df08bf2-622a-4674-ac31-51930f6a80f8 [https:// perma.cc/67FC-37AH].

62. *China Passes New Cybersecurity Law*, *supra* note 60.

63. *China Released Updated Draft Data Security Law and Personal Information Protection Law for Public Comments*, INSIDE PRIVACY (May 3, 2021), https://www.cov.com/-/media/files/corporate/publications/2021/05/covington-alert--china-released-updated-draft-data-security-law-and-personal-information-protection-law-for-public-comments-may-3-2021.pdf [https://perma.cc/S9SM-7RKW].

a significant online presence. Corporations from over forty countries issued a letter to Chinese premier Li Keqiang, with concerns including an assertion that regulator-led security reviews of information technology products and services under the CSL only create additional barriers to entry as opposed to heightened data security.[64] Despite an initial barrage of protests, corporations ultimately moved forward with regulatory compliance, given a heightened awareness of customer privacy rights during that time, in light of the passage of numerous privacy laws with global impacts, such as the GDPR. Major law firms pivoted towards establishing data privacy and cybersecurity practice groups to ease the transition towards compliance and redesigning privacy policies for corporations. Nevertheless, while Chinese regulators emphasized that the CSL's goal was to promote national security and safeguard the public's interests with a significant consumer privacy component, the CSL, at its core, reflects the government's focus on improving a defensive cyber posture, with key elements of the People's War in play—a whole-of-country defense, the ability to sustain a protracted war, and asymmetrical warfare.[65]

A. *Whole-of-Country Defense*

The CSL and accompanying regulations contain numerous provisions that set forth a broadly applicable security standard for all entities operating within the country. Article 21 provides that "[n]etwork operators shall perform . . . security protection duties according to the requirements of the cybersecurity multi-level protection system," with network operators broadly defined as "network owners, managers, and network service providers."[66] Additionally, Article 31 states that "[t]he State implements key protection on the basis of the cybersecurity multi-level protection system for public communication and information services, power, traffic, water resources, finance, public service, e-government, and other critical information infrastructure which—if destroyed, suffering a loss of function, or experiencing leakage of data—might seriously endanger national security, national welfare, the people's livelihood, or the public interest."[67] Finally, as both articles allude to, the

64. Tom Mitchell and Shawn Donnan, *Chinese Laws Prompt Global Business Backlash*, FIN. TIMES (Aug. 11, 2016), https://www.ft.com/content/8103baa0-5f9c-11e6-ae3f-77baadeb1c93 [https://perma. cc/H9R8-MQH6].

65. Charles Clover and Sherry Fei Ju, *China Cyber Security Law Sparks Foreign Fears*, FIN. TIMES (Nov. 7, 2016), https://www.ft.com/content/c330a482-a4cb-11e6-8b69-02899e8bd9d1 [https:// perma.cc/BB2P-KDYG].

66. Rogier Creemers, et al., *Translation: Cybersecurity Law*, DIGICHINA (June 29, 2018), https://digichina.stanford.edu/work/translation-cybersecurity-law-of-the-peoples-republic-of-china-effective-june-1-2017 [https://perma.cc/9YBL-RUW4].

67. *Id.*

multi-level protection system exists as a tiered system of classifying information systems and imposes security standards based on the risk and impact of a possible data breach.

In line with the People's War and the idea of civil-military unity, the CSL's expansive provisions in subjecting network operators and critical information infrastructure to common security standards further China's strategy of mobilizing the entirety of society in a defensive posture to minimize weaknesses across all networks within the country. The term network operator "covers virtually any business that operates an internal computer network, or even just a website, in China."[68] In other words, the CSL applies not only to government-operated networks but also to private-sector networks that belong to foreign and domestic companies.

All entities that operate a network in China are now required to adhere to security standards that assess the impact on national security, public interests, or social order, evaluated at a scale of one to five, with the most stringent standards applicable to network operators that pose the highest risk at level five.[69] Such standards range from requiring a qualified expert to conduct a security review of level two networks to requiring regulatory intervention in determining a schedule for reevaluating level five networks, which are often government-owned.[70] The broadly applicable wording with respect to network operators and the unified security standards of the CSL reflect the spirit of civil-military fusion because public and private network operators are equally obligated to implement cybersecurity measures, thus reducing reliance on purely governmental or military networks for national security and defense. Additionally, a whole-of-country defense that utilizes the civilian sector is important to reducing potential weaknesses in critical industries.[71]

This concern manifested in practice during large-scale combat operations in 2022, as Russia conducted a series of offensive cyber

---

68. Zachary S. Brez, et al., *Challenges and Advice for Multinational Companies in Complying with Chinese Cybersecurity Law*, KIRKLAND & ELLIS (Feb. 23, 2018), https://www.kirkland.com/publications/article/2018/02/challenges-and-advice-for-multinational-companies [https://perma.cc/P5ZM-WC7J].

69. U.S.-China Business Council, *The 5 Levels of Information Security in China*, CHINA BUS. R. (Dec. 5, 2016), https://www.chinabusinessreview.com/the-5-levels-of-information-security-in-china [https://perma.cc/TBZ4-ZQ2G].

70. Michael Pang and Jonathan Hsieh, *China's Cybersecurity Law: Multiple-level Protection Scheme*, PROTIVITI, https://www.protiviti.com/HK-en/insights/pov-multiple-level-protection-scheme [https://perma.cc/4QAD-Y6VY].

71. "Unlike military or intelligence networks, which are defended and overseen by the Department of Defense, or various civilian government networks, which are defended and overseen by the Department of Homeland Security, the National Institute of Standards and Technology, and the Office of Budget and Management, no one entity defends the private networks that most critical infrastructure relies upon." Robert K. Palmer, *Critical Infrastructure: Legislative Factors for Preventing a "Cyber-Pearl Harbor"*, 18 VA. J.L. & TECH. 289, 293 (2014).

operations against Ukrainian critical infrastructure, a mix of government and civilian systems, including targeting a power plant in an attempt to hinder electricity distribution and government websites from delaying distribution of relief supplies.[72] By imposing heightened cybersecurity obligations on critical information infrastructure, alongside bringing all network operators under the CSL's scope, China would be able to bolster its cybersecurity capabilities by mobilizing all entities operating within the country, thereby minimizing areas that may be vulnerable to exploitation.

B. *Protracted War – Big Areas and Little Areas*

In addition to the CSL, Chinese regulators have also promulgated a plethora of sector-specific cybersecurity requirements that further elevate its ability to withstand cyberattacks and address vulnerabilities. For example, in October 2019, the National People's Congress enacted the Encryption Law, which imposes, among other requirements, the obligation for critical information infrastructure operators to undergo a security assessment of commercial encryption product usage, where applicable, as well as an import-export framework that restricts encryption products that may impact national security.[73] In February 2020, the People's Bank of China issued the Personal Financial Information Protection Technical Specification, which governs how financial institutions collect and process personal information; e.g., where sensitive information is transmitted over public networks, financial institutions must ensure that such information is encrypted.[74] The aforementioned laws are a sampling of the sector-specific cybersecurity requirements that have been promulgated on top of the CSL and demonstrate China's commitment to additional security measures for sectors of concern, with some overlap with critical information infrastructure.

Sector-specific laws in the areas including encryption and finance allow for heightened protection of certain sectors that the government deems sensitive. Interconnectivity is a key nature of cybersecurity, which means that "[w]hile interdependencies among CI [critical infrastructure]

---

72. Jakub Przetacznik and Simona Tarpova, *Russia's War on Ukraine: Timeline of Cyber-attacks*, EPRS (June 2022), https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733549/ EPRS_BRI(2022)733549_EN.pdf [https://perma.cc/E67W-GFY4].

73. Eric Carlson and Yan Luo, *China Enacts Encryption Law*, INSIDE PRIVACY (Oct. 31, 2019), https://www.insideprivacy.com/data-security/china-enacts-encryption-law/ [https://perma.cc/2WCC-B5XM].

74. Yan Luo, *China Releases Personal Financial Information Protection Technical Specification*, INSIDE PRIVACY (Dec. 2, 2019), https://www.insideprivacy.com/international/china/china-releases-personal-financial-information-protection-technical-specification/ [https://perma.cc/5E8N-S5ER].

are often necessary to meet design specifications, they also lead to undesirable situations when a fault or attack occurs in one CI and escalates to other connected CI."[75] A case study was done regarding the impact of a cyberattack on interconnected systems of a water distribution center and a water treatment plant.[76] Here, after acquiring knowledge of points of weakness in the two systems, an attacker can manipulate multiple points simultaneously, with a larger number of interconnected nodes or links translating into a larger potential surface area for attack.[77]

The interconnectivity of systems can also mean that a single vulnerability can affect a multitude of systems and infrastructure. In January 2003, the SQL Slammer worm exploited unpatched SQL servers; an infected server would then prompt the host computer to search for and infect additional servers.[78] This cascading effect from a single point of weakness resulted in severe consequences, including ATM failures and canceled flights.[79] The sector-specific cybersecurity laws that exist on top of the CSL mitigate the dangers of such cascading effects of a cyberattack. For example, an attack on a specific node in one sector may be isolated, thus keeping the other sectors and the larger cyberinfrastructure intact.

The differing, heightened requirements across sectors lend to the concept of "big areas versus little areas" under the People's War, in which even if the enemy conquers and occupies a specific area of the country, the larger, remaining areas remain intact and in China's possession, with the latter continuously mobilizing to maintain sustained resistance against the enemy.[80] This would also allow China to fight a protracted war of attrition against a much stronger enemy by having constant pockets of defense.[81] By extension, China's sector-specific cyber regulations in conjunction with the CSL would, in theory, allow it to survive an initial cyberattack by limiting its impact and preserving the integrity of its remaining systems to fight a protracted war.

---

75. Venkata R. Palleti, et al., *Cascading Effects of Cyber-attacks on Interconnected Critical Infrastructure*, CYBERSECURITY 2 (Mar. 1, 2021), https://cybersecurity.springeropen.com/track/pdf/10.1186/s42400-021-00071-z.pdf [https://perma.cc/RRW8-UCZ8].

76. *Id.*

77. *Id*. at 16–17.

78. Roger A. Grimes, *SQL Slammer 16 Years Later: Four Modern-Day Scenarios that Could be Worse*, CSO (Jan. 31, 2019), https://www.csoonline.com/article/3337179/sql-slammer-16-years-later-four-modern-day-scenarios-that-could-be-worse.html [https://perma.cc/TZJ8-7DJJ].

79. *Protecting Interconnected Systems in the Cyber Era*, PWC 11, https://gita.org.in/Attachments/Reports/Protecting%20interconnected%20systems%20in%20the%20cyber%20era.pdf [https://perma.cc/L4JN-EPPM].

80. MAO, *supra* note 19, at 147–48.

81. *Id.* at 141–42.

C. *Asymmetrical Warfare*

Chinese regulators have additionally set forth laws that provide for actively monitoring potential vulnerabilities. In 2021, the Cyberspace Administration of China and the Ministry of Public Security promulgated the Provisions on the Management of Network Product Security ("Network Product Security Provisions"), which requires reporting of security vulnerabilities.[82] Article 7 states that network operators and network product providers shall report the vulnerability to the Ministry of Industry and Information Technology within two days of discovering a security vulnerability.[83] Article 9 also prohibits entities and individuals from publishing vulnerabilities to overseas entities and individuals.[84] Like the CSL, the aforementioned Articles broadly apply to network operators and network product providers of hardware and software operating within China.[85] Separately, the Data Security Law requires processors of important data to submit a regular risk assessment report that includes "the types and amounts of important data processed, information on data processing, data security risks and the response measures for them."[86]

The Network Product Security Provisions and Data Security Law show China's concerns with an interest in zero-day vulnerabilities. Zero-day vulnerabilities are vulnerabilities that entities have not yet patched. Importantly, according to a case study done by a cyber threat company based on tracking sixty vulnerabilities that occurred between 2018 and 2019, "[t]he average day between disclosure and patch availability was approximately 9 days," thereby providing attackers with a window of opportunity to manipulate the vulnerability.[87] Moreover, forty-two percent of vulnerabilities were exploited even after a patch was issued.[88] By being able to monitor such zero-day vulnerabilities under the Network Product Security Provisions, as well as having risk assessment reports that detail data processing and its corresponding risks as mandated by the Data Security Law, China would have a better understanding of new

82. Wangluo Chanpin Anquan Loudong Guanli Guiding (网络产品安全漏洞管理规定) [Provisions on the Management of Network Product Security] (2021).

83. *Id.*

84. *Id.*

85. *Id.*

86. Data Security Law of the People's Republic of China (2021).

87. Kathleen Metrick et al., *Think Fast: Time Between Disclosure, Patch Release and Vulnerability Exploitation—Intelligence for Vulnerability Management, Part Two*, MANDIANT (Apr. 13, 2020), https://www.mandiant.com/resources/time-between-disclosure-patch-release-and-vulnerability-exploitation [https://perma.cc/GBH9-NVQR].

88. *Id.*

vulnerabilities as they arise to protect its own networks and potentially use them against adversaries in offensive cyber operations.[89]

This further aligns with a core tenet of the People's War: overcoming a superior adversary requires flexible tactics and exploiting the enemy's weaknesses through asymmetrical warfare.[90] The PLA has long framed military strategy from a position of needing to prevail over a militarily superior adversary, and knowledge of newly discovered, obscure vulnerabilities would—in theory—present an opportunity for an advantageous attack on such adversary's systems or software where a patch has either not yet been released, or alternatively, has been released but has not seen widespread distribution.[91] Of note, monitoring vulnerabilities can also be viewed under the broader umbrella of active defense. Given the lack of geographical boundaries within cyberspace, networks, and nodes can be construed as vulnerable to attack on the fringes of China's area of operations or territory. Such monitoring can be viewed as a forward defensive posture in providing early warning of possible weaknesses in cyberspace.

IV. U.S. STRATEGIC CONCERNS AND A PATH TO BRIDGING THE DIVIDE

The CSL sits at the intersection between military, civilian, and legal cyber interests, thus posing unique challenges to the United States in crafting an effective response. At the outset, the United States and China have differing views on their respective strategic approaches to cyber governance and cyber sovereignty, resulting in a higher possibility for misunderstandings or mistrust.[92] Moreover, while the United States should prioritize establishing a better system for sharing cyber-threat information in response to the CSL, the legislative process can be lengthy and needs to account for the competing interests of the public and private sectors.[93] A possible, more immediate path forward would be restarting high-level bilateral dialogues on cyber interests between the two countries to eliminate pockets of misunderstanding, establish red lines, and create a code of conduct to facilitate predictability in cyber operations further.

---

89. Brad D. Williams, *China's New Data Security Law Will Provide it Early Notice of Exploitable Zero Days*, BREAKING DEFENSE (Sept. 1, 2021), https://breakingdefense.com/2021/09/chinas-new-data-security-law-will-provide-it-early-notice-of-exploitable-zero-days [https://perma.cc/T7NB-ZA4L].

90. MAO, *supra* note 19, at 167.

91. Caitlin Campbell, *China's Military: The People's Liberation* Army, CRS 22 (June 4, 2021), https://sgp.fas.org/crs/row/R46808.pdf [https://perma.cc/48MQ-6YQ4].

92. Kristen E. Eichensehr, *The Cyber-Law of Nations*, 103 GEO. L.J. 317, 329 (2015); Anqi Wang, *Cyber Sovereignty at its Boldest: A Chinese Perspective*, 16 ISJLP 395, 397 (2020).

93. Palmer, *supra* note 4, at 197.

A. *A Fundamental Divide*

While in recent years, numerous countries have reached a consensus that cyberspace constitutes a new warfighting domain, the laws passed by each country regulating cybersecurity as it relates to national security ties into a broader issue of cyber governance.[94] From a Chinese perspective, cyberspace has intangible territorial borders that each country can exert control over for a number of goals, including the preservation of social stability, copyright protection, and national security; in other words, China promotes the concept of cyber sovereignty, which divides cyberspace into country-based jurisdictions.[95] Conversely, the United States prioritizes a free and open internet that embraces a multi-stakeholder approach to governance.[96] The advancement of cyber capabilities in both countries and the divergence in their strategic approach to cyberspace creates the potential for misunderstandings and, consequently, escalation of force.[97] For example, China may view the CSL as a legal framework that is necessary to safeguard its critical information infrastructure against malicious actors and possible foreign threats, but the United States may view the same law as destabilizing to the international community with respect to the free flow of information and also dangerous with respect to increasing its offensive cyber capabilities.[98] Indeed, even if a common interest in preventing escalation exists, the divergent views of cyberspace governance and strategy may result in what one party views as addressing legitimate domestic concerns as prepping the battlefield by another party.[99]

B. *Seeking Mutual Understanding of Strategic Interests in Cyberspace*

The United States has the challenge of formulating a balanced response to China's CSL, with the need to navigate the nuance between having an effective counter to the potentially offensive elements within the CSL and avoiding a spiral of mistrust and military escalation, as both

94. Eichensehr, *supra* note 92, at 329.

95. Wang, *supra* note 92, at 397.

96. Eichensehr, *supra* note 92, at 330.

97. Michael Kolton, *Interpreting China's Pursuit of Cyber Sovereignty and its Views on Cyber Deterrence*, 2 CYBER DEF. REV. 119, 137 (2017).

98. Laura Dobberstein, *China is Likely Stockpiling and Deploying Vulnerabilities, Says Microsoft*, REGISTER (Nov. 7, 2022), https://www.theregister.com/2022/11/07/china_stockpiles_vulnerabilities_microsoft_asserts [https://perma.cc/56B7-YQUC]; Tom Miles, *U.S. Asks China Not to Enforce Cyber Security Law*, REUTERS (Sept. 26, 2017), https://www.reuters.com/article/us-usa-china-cyber-trade/u-s-asks-china-not-to-enforce-cyber-security-law-idUSKCN1C11D1 [https://perma.cc/KSN3-6RQP].

99. Kolton, *supra* note 97, at 140.

countries would begin to enter into a feedback loop in responding to the other's actions.[100]

To mitigate the CSL's vulnerability reporting requirements, which can potentially be used offensively, the United States should continue its efforts in building a tailored, robust cyber-threat sharing framework between the public and private sectors to anticipate zero-day vulnerabilities similarly.[101] Real-time sharing and analysis of data trends and unusual behaviors would assist in identifying and stopping malicious activity.[102] To this end, the United States already has Information Sharing and Analysis Centers (ISACs), established by Presidential Decision Directive-63 in 1998, wherein "each critical infrastructure sector . . . establish[ed] sector-specific organizations to share information about threats and vulnerabilities," with most ISACs having "24/7 threat warning and incident reporting capabilities."[103] However, the effectiveness of ISACs remains questionable due to artificial self-imposed limits in cyber-threat sharing, where, for example, some ISACs share information only with trusted members, as opposed to allowing for broad, simultaneous dissemination of information.[104] A centralized entity or organization that aggregates and shares the cyber-threat information may be more effective in minimizing the shortcomings of the preexisting ISAC framework, particularly if the types of information to be shared is clearly delineated to filter for critical information and is screened to deconflict with the patchwork of applicable privacy laws.[105] However, a number of competing interests remain in play and have long hindered legislative progress in this area; whether the government should mandate information sharing or minimum security standards continues to be a point of contention.[106] Proponents of government-required standards believe that market forces and voluntary behavior are inadequate to address the cyber threats against the United States.[107] On the other hand,

100. *See, e.g.*, *id.* at 137.

101. Palmer, *supra* note 4, at 314.

102. *Id.* at 314–15.

103. *Id.* at 316; *About ISACs*, National Council of ISACs, https://www.nationalisacs.org/about-isacs [https://perma.cc/7UT5-NP49].

104. Palmer, *supra* note 4, at 317–18.

105. *Id.* at 355–56.

106. *Id.* at 297.

107. *Id.* Of note, under the Biden Administration, the United States government has taken steps to address collaboration between the public and private sectors; indeed, the United States government "announced and operated under a new model for cyber incident response by including private companies in the Cyber Unified Coordination Group." *The United States, Joined by Allies and Partners, Attributes Malicious Cyber Activity and Irresponsible State Behavior to the People's Republic of China*, WHITE HOUSE (July 19, 2021), https://www.whitehouse.gov/briefing-room/statements-releases/2021/07/19/the-united-states-joined-by-allies-and-partners-attributes-malicious-cyber-activity-and-irresponsible-state-behavior-to-the-peoples-republic-of-china [https://perma.cc/9WZ6-4QKB].

opponents of such standards believe that the government cannot effectively address the needs of varying industries and sectors and may stifle innovation instead.[108]

In light of the legislative and legal barriers that lengthen the timeline to establish an effective scheme of sharing cyber-threat information, a more immediate step the United States can take to address concerns surrounding the CSL would be to reestablish and participate in regular bilateral dialogue on cyber concerns, as well as create a code of conduct for cyber operations. The formal dialogue on cybersecurity that began under the Obama[109] and Trump[110] Administrations should continue to build a robust understanding of differing strategic interests and also enumerate the red lines that each country may have to prevent or deescalate potential crises in cyberspace. This is especially crucial in cyberspace, where the rapidness of a potential attack or response can come without the early warning signals of ground maneuver, such as troop buildup, and attribution can be unclear.[111] Accordingly, there must be a reversal of the current status, in which, after multiple years of the coronavirus pandemic, "many government channels [have been] canceled, suspended or lapsed, [and] unofficial dialogues have been among the few tools left to keep the two sides from continuing to talk past each other."[112] Importantly, the United States and China should agree on a code of conduct concerning cyber operations to further minimize areas of uncertainty. While the Tallinn Manual exists as what experts consider the "current black letter law on jus ad bellum and jus in bello rules relevant to cyber operations,"[113] some Chinese scholars[114] have been critical that the Tallinn Manual 2.0 does not adequately address

---

108. Palmer, *supra* note 4, at 297.

109. *First U.S.-China High-Level Joint Dialogue on Cybercrime and Related Issues Summary of Outcomes*, Dept. Justice (Dec. 2, 2015), https://www.justice.gov/opa/pr/first-us-china-high-level-joint-dialogue-cybercrime-and-related-issues-summary-outcomes-0 [https://perma.cc/LT94-ETCU].

110. *First U.S.-China Law Enforcement and Cybersecurity Dialogue*, Dept. Justice (Oct. 6, 2017), https://www.justice.gov/opa/pr/first-us-china-law-enforcement-and-cybersecurity-dialogue [https://perma.cc/A4H6-NSCS].

111. Ernest J. Monitz, et al., *U.S. Nuclear Policies for a Safer World*, NTI (June 10, 2021), https://www.nti.org/analysis/articles/us-nuclear-policies-safer-world [https://perma.cc/MVL4-N8Q2].

112. Christian Shepherd and Lyric Li, *China Wants to Mend Ties with the U.S. but it Won't Make the First Move*, WASH. POST (Nov. 13, 2022), https://www.washingtonpost.com/world/2022/11/13/china-united-states-relations-xi-jinping [https://perma.cc/V5VM-6824].

113. Ashley Deeks, *Tallinn 2.0 and a Chinese View on the Tallinn Process*, LAWFARE (May 31, 2015), https://www.lawfareblog.com/tallinn-20-and-chinese-view-tallinn-process [https://perma.cc/XE8T-J5B9.

114. While such scholars may not necessarily represent the official views of the Chinese government or the People's Liberation Army, their views are nevertheless edifying in exploring how Chinese views may differ from Western views with respect cyberspace.

certain concerns, including the consequence-based view of cyberattacks[115] and data as a military objective.[116] The latter, in particular, has been controversial even between Chinese scholars who take differing stances on whether data should be considered a "non-object" military objective, in light of how the Tallinn Manual 2.0 considers military objectives to be objects, and even if it were to be considered a valid military objective, whether data should be segregated into military and civilian data.[117] Rather than simply following the Tallinn Manual 2.0, a code of conduct could codify the intent and stances of both governments and further explore red lines to reduce concerns of unintended or misinterpreted signals. Additionally, the less formal nature of a code of conduct, as compared with a treaty-based option, would be a good step forward in developing a better understanding of areas of concern with respect to cyber operations between the United States and China without locking either country into a potentially difficult political position.

CONCLUSION

Though the People's War has its origins in Mao Zedong's philosophy of class warfare in the early 1900s[118] and pre-dates the Second Sino-Japanese War, it has remained relevant in modern Chinese military doctrine as more than just an antiquated slogan. The People's War has evolved alongside doctrinal shifts throughout the decades, from active defense in the early 1970s[119] to fighting under informatized conditions in the early 1990s.[120] In this time, the People's War transformed from a more literal mobilization of the masses to overthrow the gentry into military-civil fusion under the umbrella of active defense.[121]

Even in the new cyber domain, the concept of the People's War is applicable and features heavily in the CSL. Indeed, the CSL's broadly mandated security standards across the public and private sectors[122] tie

115. The Tallinn Manual 2.0 takes a consequence-based view of cyberattacks, in which a cyber operation is considered a cyberattack where the operation is reasonably expected to cause death, damage, or injury to persons or objects. However, under this view, "assessment of the damage turns out to be extremely tricky, especially when the consequences are mostly indirect," and "the consequence-based approach limits the notion of the attack so as to exclude those operations that result in severe and disruptive non-physical harm." Zhixiong Huang and Yaohui Ying, *The Application of the Principle of Distinction in the Cyber Context: A Chinese Perspective*, 913 IRRC 335, 343 n.32 (2020).

116. *Id.* at 360.

117. *Id.* at 362–63.

118. Mao, *supra* note 12 at 23.

119. LEE, *supra* note 9, at 50–51.

120. FRAVEL, *supra* note 10, at 220.

121. *Id.* at 231.

122. *See, e.g.*, Creemers et al., *supra* note 66.

into the concept of a layered defense, using the strength of the entirety of the country under the People's War. Sector-specific regulations[123] on top of the CSL increase survivability through isolating threats, thereby setting conditions to fight a protracted war of attrition against the adversary. Finally, the CSL's vulnerability reporting mechanisms are suspected to have a secondary function of gathering zero-day vulnerabilities in an offensive capacity,[124] again tying into a familiar concept under the People's War—asymmetrical warfare, in which an adversary's weaknesses can be leveraged and exploited through non-conventional means.

In turn, the United States faces challenges in crafting a measured response to the CSL. A forceful response in shoring up offensive capabilities may not be ideal. Mike McConnell, a former director of the National Security Agency, noted:

> Let's say you take an action. We depend on this stuff more than anyone else. We're more vulnerable than anybody else in the world. If we could put a map of the world up here with the US on the center and we put bandwidth on top of it, it's a bell curve. Most of the communications in the world flow through the United States; we are the biggest users and beneficiaries. So, there's a great hesitancy to use anything in a cyber context because it's relatively easy to punch back aggressively.[125]

Additionally, in light of the legislative barriers to creating an effective cyber-threat information-sharing system,[126] the United States may find more immediate success in resuming high-level dialogue in identifying the red lines of each country and areas of potential misunderstanding, particularly as the United States and China have fundamental differences in their respective approaches to cyberspace and strategy. The United States should also formulate a bilateral code of conduct to eliminate further ambiguities in signaling and intent with respect to cyber operations, thereby reducing the risk of escalation. Notwithstanding the above, cyberspace will likely be a continued area of tension for the United States in the coming years, particularly with the increasing intersection between civilian and military purposes within cyberspace and the diverging views between countries with respect to cyber sovereignty.

---

123. *See, e.g.*, Carlson and Luo, *supra* note 84.

124. Dobberstein, *supra* note 98.

125. Kevin J. Deleney, Why the US Doesn't Use Cyber-weapons to Attack its Enemies More Often, QUARTZ (June 30, 2013), https://qz.com/99162/why-the-us-doesnt-use-cyber-weapons-to-attack-its-enemies-more-often-mike-mcconnell [https://perma.cc/UE67-GKSJ].

126. Palmer, *supra* note 4, at 297.

STATISTICAL SECURITIES COMPLIANCE

Brian Haney[*]

Abstract

This Article makes three main contributions. First, this Article introduces the Solana blockchain as a public good and provides policy analysis for open innovation. Second, this Article introduces a new dataset for SEC blockchain enforcement, supporting empirical compliance analysis. Third, this Article draws on the legal informatics literature to provide a mechanism for applied analysis of digital assets on the Solana blockchain in the context of securities law. The main purpose of this Article is to introduce new methods for using natural language processing to automate compliance services on the Solana blockchain.

INTRODUCTION

The problem this work sets out to solve is how to differentiate between security and non-security tokens under U.S. law legally. While many digital tokens are not classified as "securities" and are thus not subject to SEC jurisdiction, many others are classified as securities by the SEC. It is often unclear whether a given token is a security, and making this determination through traditional legal analysis can prove to be quite

* Thank you to the Veronica Root Martinez, Mark Lemley, Justice Amy Barrett, Justice Clarence Thomas, Fernando Martinez, Athena Aherrera, Andrea Baglioni, Gemma Guerrero-Wiest, Jimmy Bingham, Prosper Adeoti, Eugene Nnamdi, Sam Tosin, and David Kazeem.

challenging. This ambiguity regarding the legal classification of digital tokens cuts to the core of digital asset regulation and will be a defining feature of twenty-first-century finance. This problem is so central to the future of finance because it underlies the essence of blockchain technology as a mechanism for decentralization and as a stimulus for economic opportunity, transparency, and legitimacy.

The solution to this classification problem is a statistical method for analyzing digital assets in the context of U.S. securities law. Drawing on the law and informatics scholarship, this Article introduces a new process and software for statistically analyzing digital assets on the Solana blockchain. The statistical strategies introduced in this Article provide a fair and concise method for differentiating between digital assets that are securities and digital assets that are not securities.

Part I provides an overview of blockchain technology, emphasizing Solana, a cutting-edge and global information technology. Part II discusses and describes data regarding securities law enforcement in the blockchain space and introduces a novel dataset. Part III draws on legal informatics to introduce new mechanisms for measuring blockchain compliance and applies those mechanisms to produce a computational analysis of assets on the Solana blockchain.

I. BLOCKCHAIN

Blockchains[1] are decentralized databases that are maintained by global computer networks.[2] According to scholar Primavera De Filippi, "Blockchain technology constitutes a new infrastructure for the storage of data and the management of software applications, decreasing the need for centralized middlemen."[3] Consisting of computers called nodes,[4] blockchains connect computers via the Internet.[5] This type of relationship among the various nodes is called a peer-to-peer network, a dynamic

1. *See Generally* Emily Wells, et al., *Blockchain Benefits and Risks* (May 2018), https://www.researchgate.net/profile/Igor_Linkov/publication/325385235_Blockchain_Benefits_and_Risks/links/5df6b251a6fdcc2837245f1e/Blockchain-Benefits-and-Risks.pdf. *See also* Elona Marku et al., *General Purpose Technology: The Blockchain Domain, Int. J. of Bus. and Mgmt.* (Oct. 2020), https://www.researchgate.net/publication/346557624_General_Purpose_Technology_The_Blockchain_Domain.

2. *See* Marku et al., *supra* note 1.

3. PRIMAVERA DE FILIPPI, AARON WRIGHT, BLOCKCHAIN AND THE LAW 33 (2018).

4. Each node maintains a transaction record called a ledger, which acts as a new data structure on the Internet.

5. David Mills et al., Distributed Ledger Technology in Payments, Clearing, and Settlement, 10 (Fed. Rsrv. Bd. Fin. & Econ. Discussion Series, Working Paper No. 95, 2016).

information technology that facilitates global, programmatic, and online protocols.[6]

A. *Solana*

At the turn of the 20th Century, Thomas Edison and Nikola Tesla were competitively inventing new machines and processes to generate electric power.[7] Edison is famous for inventing the light bulb, but long forgotten is his plan to bring electricity to the world using a direct current.[8] Tesla invented a better model for electricidal transmission, an alternating current.[9] Where direct currents only flow in one direction, alternating currents flow in multiple directions, which increases both magnitude and transmission range.[10] Today the entire world runs on the alternating electrical current invented by Tesla, which now connects computers all around the world.

Just as early electrical transmission designs were divided into direct current and alternating current, blockchain technology is currently segmented into two essential models—the proof-of-work (PoW) blockchain and the proof-of-history (PoH) blockchain. The most prominent PoW blockchain is Bitcoin.[11] Much like Edison's light bulb, Bitcoin was a breakthrough technology.[12] However, much like the direct current technology underlying the light bulb, Bitcoin's PoW model does not scale. The major problem is that PoW requires expensive mining

6. SATOSHI NAKAMOTO, BITCOIN: A PEER-TO-PEER ELECTRONIC CASH SYSTEM 8 (2008) ("The peer-to-peer network developed to solve the double spending problem, where the same digital token is spent more than once."); *see also* David Mills et al., Distributed Ledger Technology in Payments, Clearing, and Settlement, 10 (Fed. Rsrv. Bd. Fin. & Econ. Discussion Series, Working Paper No. 95, 2016).

7. Apparatus for The Electrical Transmission of Power, U.S. Patent No. 265,786 (filed Oct. 10, 1882); *see also* Electric Light, U.S. 219,628 (filed Sept. 16, 1879).

8. He wanted to power homes with the same direct current he used in the light bulb, but the problem was the direct current model couldn't scale. As such, Edison's design was limited to providing electricity within a few blocks of a power station.

9. Method of Converting and Distributing Electric Currents, U.S. Patent No. 382,282 (filed May 1, 1888); *see also* Pyromagnetic Electric Generator, U.S. Patent No. 428,057 (filed May 13, 1890).

10. Tesla's model crushed Edison's because the alternating current could scale.

11. *See* Satoshi Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System*, BITCOIN (2008), https://bitcoin.org/bitcoin.pdf [https://perma.cc/JJ6N-4YJQ]. *See also* SAIFEDEAN AMMOUS, THE BITCOIN STANDARD: THE DECENTRALIZED ALTERNATIVE TO CENTRAL BANKING (2018).

12. The PoW model relies on a complex cryptographic hashing algorithm for a process called mining, which is used to distribute new assets and incentivize network maintenance.

operations, which demand massive computing power and electricity consumption.[13]

The PoH model is necessary for custom compliance program creation to meet the specific needs of global and instantaneous information transfer at no cost. The best PoH blockchain is Solana. The Solana blockchain architecture is based on PoH, which is a computational proof for verifying order and temporal relationships.[14] In short, PoH is an innovation that allows for encoding trustless time-lapse on a distributed ledger. When used alongside a consensus algorithm such as proof-of-work (PoW) or proof-of-stake (PoS), PoH can reduce messaging overhead and enable faster transactions than previous blockchain mechanisms. As such, Solana offers an order of magnitude improvement in global payments and transactions cost-efficiency with its novel smart contract technology.[15]

B. *Open-Source Software*

Open-source software (OSS) is the best software. Open source innovation drives the edge across the industry, from information technology[16] to defense.[17] The general effect of creating an open-source license is to grant a free license while limiting liability for the holder without warranty.[18] For decentralized projects and startups, one idea behind open innovation is the creators of new ideas do not have to be within an organization to be helpful.[19] Solana has applied this principle to allow for the global development of the world's best high-performance blockchain.

---

13. As a result, network maintenance for PoW blockchains is extremely expensive and economically inefficient. So, the cost for transactions is unnecessarily high and, in some instances, can cost hundreds of dollars for a single transaction.

14. ANATOLY YAKOVENKO, SOLANA: A NEW ARCHITECTURE FOR A HIGH, PERFORMANCE BLOCKCHAIN V0.8.13 (2022).

15. On Solana, smart contracts are a generalized term for transactions, or programs that run on nodes and modify the blockchain with transactional data.

16. Major open source in information technologies include Solana, Bitcoin, TensorFlow, Ethereum, Selenium, React, Python, and React.

17. *See* youshixun, *Versatile model of cognitive electronic warfare with countermeasures*, GITHUB (2019), https://github.com/youshixun/vCEW [https://perma.cc/A87T-VES7]; *see also* Shixun You et al., *Completing Explorer Games with a Deep Reinforcement Learning Framework Based on Behavior Angle Navigation*, ELECTRONICS 17 (2019), https://www.mdpi.com/2079-9292/8/5/576 [https://perma.cc/8AWN-SDDQ].

18. Heli Koski, *OSS Production and Licensing Strategies of Software Firms*, 2 REV. ECON. RSCH. ON COPYRIGHT ISSUES 111, 117 (2005) (explaining OSS is often attractive because it introduces software for free which makes it easier to establish a large user base and increase revenue from complementary service provisions).

19. JOHN PALFREY, INTELLECTUAL PROPERTY STRATEGY 107 (2011); *see also* PETER THIEL, ZERO TO ONE 129 (2014).

More generally, OSS has many advantages compared to proprietary software development.[20] For example, OSS can also be extremely helpful for developing secure code because the public nature of the product allows anyone to report bugs or issues. OSS also serves as a public good by advancing human knowledge in science, technology, and innovation. Moreover, OSS develops contributions from a global talent pool of diverse inventors, creators, and engineers. According to Stanford Law Fellow Fernando Morera, organizations need to collaborate within decentralized ecosystems to be effective and maximize knowledge and value.[21]

OSS is a keystone to Solana's ability to be able to innovate and invent new technologies. Solana is an open-source and decentralized blockchain built for optimal performance.[22] In fact, Solana's development of OSS is a public good that creates transparency for global finance. The Solana Foundation, which maintains the Solana OSS code base, solves this problem by incentivizing developers with grants to support the public good and open innovation. This helps to correct the common economic and opportunistic inequality in institutional technology development.

The two most prominent open-source licenses the Solana Network uses are the Apache and the MIT License. In fact, these two licenses are the most used licenses on the Solana Foundation GitHub.[23] The Apache License expressly offers the software—as is and without warranty.[24] Interestingly, the MIT License is relatively similar in structure because, like the Apache License, the MIT License grants a license to use the technology while limiting liability for the copyright holder.[25] However, one main difference between the two licenses is that the Apache License is expressly irrevocable, meaning that it is permanent once the invention is disclosed.[26] Solana's open-source strategy, software, and licenses are enabling a robust ecosystem of layer-2 applications to flourish on Solana.

---

20. *See* Jeanne C. Fromer, *Machines as the New Oompa-Loompas: Trade Secrecy, the Cloud, Machine Learning, and Automation*, 94 N.Y.U. L. REV. 706, 708 (2019) ("This Article argues that, in light of the technological shifts in computing, the incentives that trade secret law currently provides to develop these contemporary Oompa-Loompas are excessive in relation to their worrisome effects on follow-on innovation and competition by others.").

21. Fernando Morera, *Governing Open Innovation – A Transatlantic Perspective*, STANFORD LAW SCHOOL, https://law.stanford.edu/projects/governing-open-innovation-a-transatlantic-perspective/ [https://perma.cc/53WL-HFUA] (last visited Nov. 5, 2022).

22. *Introduction*, SOLANA DOCUMENTATION, https://docs.solana.com/introduction [https://perma.cc/29D2-B2YC] (last visited Nov.21, 2022).

23. Solana Labs, GitHub (2022), https://github.com/orgs/solana-labs/repositories [https://perma.cc/RQ54-QBL8].

24. *Apache License, Version 2.0* (2004), THE APACHE SOFTWARE FOUNDATION, http://www.apache.org/licenses/LICENSE-2.0 [https://perma.cc/4YMP-M4L9].

25. *The MIT License*, OPEN SOURCE INITIATIVE, https://opensource.org/licenses/MIT [https://perma.cc/3RAP-J7PT] (last visited Nov. 21, 2022).

26. *Apache License, Version 2.0*, *supra* note 24, at §§ 2–3.

C. *SPL Tokens*

SPL tokens are Solana layer-2 tokens typically associated with applications on Solana. For example, mSOL is a staked asset associated with value increases proportional to Solana staking rewards.[27] Another example is ORCA,[28] an SPL token that governs the Orca decentralized exchange.[29] On the decentralized protocol, users supply tokens in liquidity pools, allowing algorithms to set prices based on supply and demand. The ORCA asset is used for various purposes in DeFi and for governance and voting on future protocol development.[30] Below, Figure 1 is a list of select SPL tokens, with data regarding asset supply and market capitalization.

Figure 1[31]

Asset	Name	Total Supply	Total Market Capitalization
Solana	SOL	511,616,946.00	$21,721,850,558.00
Orca	ORCA	99,999,998.70	$83,365,598.92
Green Satoshi Token	GST	68,730,903.87	$110,658,502.00
GenesysGo Shadow	SHDW	199,999,997.36	$146,709,722.00
Samoyedcoin	SAMO	7,236,693,918.49	$47,227,829.00
Nova Finance	NOVA	9,999,999.82	$8,789,762.00
Serum	SRM	9,992,475,560.59	$10,092,400,316.20
StepN	GMT	5,872,455,632.91	$5,713,984,262.00
Star Atlas DAO	POLIS	359,999,998.74	$217,475,968.00
Raydium	RAY	554,999,996.00	$561,158,653.00
Step Finance	STEP	4,000,000.00	$94,193,278.00
Dust Protocol	DUST	15,999,947.71	$40,319,868.23
MonkeyBucks	MBS	999,999,985.09	$78,341,998.83
Learning Star	LSTAR	400,369,233.00	$8,007,384.66
Solend	SLND	99,999,999.96	$114,999,999.95
Larix	LARIX	9,999,999,420.72	$16,783,399.02
Solice	SLC	399,999,999.86	$44,013,999.98

27. *Marinade*, GITHUB, https://github.com/marinade-finance [https://perma.cc/C6VV-L376] (last visited Nov. 21, 2022) (defining mSOL as a type of collateralized asset developed by Marinade and a DAO that makes Solana more decentralized and capital efficient through liquid staking).

28. *Trader FAQs*, ORCA (2022), https://docs.orca.so/orca-for-traders/master [https://perma.cc/K5BJ-DCVR] (last modified Oct. 31, 2022).

29. *Id.* (discussing how Orca enables near-instant token swaps using an automated market maker model).

30. OrcaORCA, COINBASE, https://www.coinbase.com/price/orca [https://perma.cc/ACW6-GM25] (last visited Nov. 21, 2022).

31. COINBASE, https://www.coinbase.com/ (last visited Nov. 21, 2022) (data gathered between 05/28/22 and 07/22/22).

Oxygen	OXY	10,000,000.00	$576,182,932.00
Only1	LIKE	500,000,000.00	$15,548,737.00

One great advantage of SPL tokens is the fact that SPL tokens have consistently partnered with Coinbase, a leading centralized exchange with multiple SEC approvals, including a public offering of equities.[32] These partnerships are largely possible due to Solana's trusted technology and reputation for blockchain business and ethics excellence. In fact, Solana's accelerating growth in the market is largely due to both brilliant tokenomics and robust business development on the network. As such, SPL tokens make Solana a great place for entrepreneurs and a public good for economic opportunity.

II. SECURITIES

Broadly, there are two main types of securities: debt and equity. The SEC website includes six types of securities: stocks, membership interests, stock options, restricted stock units, convertible instruments, and debt.[33] However, digital assets, cryptocurrencies, or blockchains are not listed among these six securities. This Part explores securities in the context of blockchain regulation, enforcement, and policy.

A. *Regulation*

Securities are financial instruments that represent an interest in equity or debt. The Securities Act of 1933 defines *Security* in the following way,

> (1) The term "security" means any note, stock, treasury stock, security future, security-based swap, bond, debenture, evidence of indebtedness, certificate of interest or participation in any profit-sharing agreement, collateral-trust certificate, preorganization certificate or subscription, transferable share, investment contract, voting-trust certificate, certificate of deposit for a security, fractional undivided interest in oil, gas, or other mineral rights, any put, call, straddle, option, or privilege on any security, certificate of deposit, or group or index of securities (including any interest therein or based on the value thereof), or any put, call, straddle, option, or privilege entered into on a national securities exchange relating to foreign currency, or, in general, any interest or instrument commonly known as a "security", or any certificate of interest or participation in,

32. *Id.*

33. What different types of securities are issued to startup investors?, SEC (2022), https://www.sec.gov/education/capitalraising/building-blocks/startup-securities [https://perma.cc/AVU8-VLRN] ("Many startups and investors refer simply to equity or an ownership interest in a company.").

> temporary or interim certificate for, receipt for, guarantee of, or warrant or right to subscribe to or purchase, any of the foregoing.[34]

The definition includes thirty total financial instruments as securities. Of course, digital assets are not one of them. Thus, only when a digital asset is used primarily as a security does it fall under the scope of SEC regulation.[35] Most analyses of digital assets as securities focus on a seminal case in securities law, *Securities and Exchange Commission v. W.J. Howey Co.* (*Howey*).[36]

In American Jurisprudence, *Howey* established the principle that:

> The test of an investment contract within [the] Securities Act is whether [the] scheme involves an investment of money in a common enterprise with profits to come solely from the efforts of others, and, if the test is satisfied, it is immaterial whether the enterprise is speculative or nonspeculative or whether there is a sale of property with or without intrinsic value.[37]

In other words, the holding in *Howey* establishes over seven decades of precedent that guarantee digital assets, such as Solana, Bitcoin, and Ethereum, are not inherently securities under U.S. Law—and that there are limits on the SEC's power. The SEC does take securities fraud and unregistered offerings seriously in the context of digital assets. However, its enforcement agenda is far too focused on aggrandizing the scope of its regulatory authority rather than prosecuting legitimate legal violations, which results in exceptional financial waste within the agency.

B. *Enforcement*

SEC allegations against unregistered offerings have increased in the past two years, both in litigation and administrative actions.[38] Much of

---

34. 15 U.S.C. § 77(b) (2022).

35. *See* Fernando Morera, Central Bank Digital Currencies – Recent Transatlantic Developments, STANFORD-VIENNA TRANSATLANTIC TECHNOLOGY LAW FORUM NEWSLETTER (Apr. 16, 2021), https://ttlfnews.wordpress.com/2021/04/16/central-bank-digital-currencies-recent-transatlantic-developments/ [https://perma.cc/N6VT-5G3Q] (explaining that one way regulation might work is if there was a regulated CBDC, while digital assets remained unregulated by law. CBDC is a "…form of digital money, intended to have both currency and legal tender status, which is issued, backed, and governed by central banks…").

36. SEC v. W.J. Howey Co., 328 U.S. 293 (1946).

37. *Id.* (citing 15 U.S.C. § 77b (2022) ("An 'investment contract', as used in the Securities Act, means a contract, transaction, or scheme whereby a person invests his money in a common enterprise and is led to except profits solely from the efforts of [a] promoter or a third party. . . .").

38. Cornerstone Research, SEC Cryptocurrency Enforcement: 2021 Update 8 (2021), https://www.cornerstone.com/wp-content/uploads/2022/01/SEC-Cryptocurrency-Enforcement-2021-Update.pdf [https://perma.cc/B8BD-D4HW].

the focus for the SEC's enforcement is on initial coin offerings (ICOs),[39] a fundraising technique involving exchanging cryptocurrency for digital assets.[40] What differentiates ICOs from other token offerings is that tokens sold through an ICO represent an equity interest in a company and are thus more likely to be both an investment and a security.[41] More broadly, the SEC continues making allegations and engaging in civil litigation with blockchain software projects across the decentralized Internet.

For example, in December 2020, the SEC sued Ripple Labs bringing allegations the Ripple cryptocurrency (XRP) was sold as an unregistered security.[42] The civil action involved the SEC making a claim for $1.3 billion in damages from Ripple. While XRP is a layer-1 digital asset and is not inherently a security, given that certain XRP tokens were allegedly sold as securities, the SEC chose to prosecute the case as a plaintiff. The key reason the SEC chose to bring a lawsuit against Ripple was that the SEC alleged Ripple raised $1,388,227,062.70 from sales of XRP to institutional investors.[43] If the SEC's allegations are proven true, then the SEC must also prove that XRP is, in fact, a security, which is highly unlikely given the asset's clear purpose and use of the Foundation for distributed ledger technology. Even if Ripple did sell XRP as a security to institutional investors, XRP is still not a security as a matter of law.

Similarly, in March 2021, the SEC brought allegations against LBRY, Inc., for an unregistered securities offering pursuant to the Securities Act.[44] The complaint alleged that LBRY, Inc. sold LBRY Credits (LBC)[45] to fund LBRY, an Ethereum project offering a free, open, and community-run digital marketplace.[46] The project uses LBC to power its

---

39. Cornerstone Research, SEC Cryptocurrency Enforcement: 2021 Update 9 (2021), https://www.cornerstone.com/wp-content/uploads/2022/01/SEC-Cryptocurrency-Enforcement-2021-Update.pdf [https://perma.cc/L7PL-UTMV].

40. *Id.*

41. In such a case, the company is offering coins as a security. See Edward O. Thorp, A Man for All Markets 301 (2017) ("Derivative securities, which include warrants, options, convertible bonds, and many later complex inventions, derive their value—as we have seen—from that of an "underlying" security such as a common the common stock of a company.").

42. Complaint at 1, SEC v. Ripple Labs, Inc., No. 1:20-cv-10832 (S.D.N.Y. Dec. 22, 2020) ("From at least 2013 through the present, Defendants sold over 14.6 billion units of a digital asset security called 'XRP,' in return for cash or other consideration worth over $1.38 billion U.S. Dollars ('USD'), to fund Ripple's operations and enrich Larsen and Garlinghouse. Defendants undertook this distribution without registering their offers and sales of XRP with the SEC as required by the federal securities laws, and no exemption from this requirement applied.").

43. *Id.* at 20.

44. SEC v. LBRY, Inc., Civ No. 1:21-cv-00260 (D.N.H. Mar. 29, 2021).

45. *See* LBRY Credits, CoinMarketCap, https://coinmarketcap.com/currencies/library-credits/ [https://perma.cc/ZPG7-23UK] (last visited May 5, 2022).

46. *See* LBRY, https://lbry.tech/ [https://perma.cc/PG9W-EAZV] (Oct. 31, 2022, 8:03 PM).

decentralized platform and support open-source software development predominantly provided to the public under the MIT License.[47]

The SEC alleged in its complaint that key facts include: (1) LBRY, Inc. offered LBC to institutional investors at a discount to the secondary market trading price, and (2) LBRY, Inc. made multiple direct sales of LBC to several investment funds.[48] LBRY allegedly received more than $11 million in U.S. dollars,[49] but this is unlikely because the total of LBC tokens, which are speculatively worth approximately $18 million total,[50] would not yield nearly $11 million in capital income as investments.[51] Still, according to the SEC, LBCs were offered and sold as investment contracts and, therefore, as securities without first registering with the federal government.

In March 2022, the SEC alleged the creators of Ormeus Coin "acted as modern-day snake oil salesmen, using social media, promotional websites, and in-person roadshows to mislead retail investors for their own personal benefit."[52] Ormeus Coin is an Ethereum asset using the ERC20 standard and is marketed as a "new digital money system backed by a fully audited industrial crypto-mining operation."[53] In this case, the allegations included deceptive fraud in addition to an unregistered offering. According to the SEC, "…the defendants falsely stated that Ormeus Coin had a $250 million crypto mining operation and was producing $5.4 million to $8 million per month in mining revenues."[54]

The complaint alleged John and JonAtina Barksdale, the Ormeus creators, defrauded retail investors out of approximately $124 million through two unregistered and fraudulent offerings of Ormeus Coin.[55] In

47. GitHub LBRY (2022), https://github.com/lbryio [https://perma.cc/CCF9-7PLQ]; *see also* Mass. Inst. of Tech. The MIT License, Open Source Initiative (2021), https://opensource.org/licenses/MIT [https://perma.cc/PDM8-7X3B].

48. Complaint at 2, SEC v. LBRY, Inc., 26 F.4th 96 (1st Cir. 2022) (No. 1:21-cv-00260).

49. *Id.*

50. LBRY Credits, *supra* note 45.

51. For example, even if some LBC tokens were sold as securities, it is unlikely all the LBC were sold as securities or in the same way. For example, given that LBC is on several exchanges and other decentralized protocols, at least some LBC must have been used as a market efficiency mechanism by arbitrage bots and, therefore, would not function as a security. Similar developer payments or rewards allow for active participation in the network and fail the *Howey Test* because there is no promise of profits from the efforts of a promoter when the user must actively engage to earn value.

52. SEC v. Barkdale and Barkdale, No. 1:22-cv-01933 (S.D.N.Y Mar. 8, 2022).

53. *See* ORMEUS COIN (May 5, 2020), https://ormeuscoin.com/ [https://perma.cc/A6EN-XPT6].

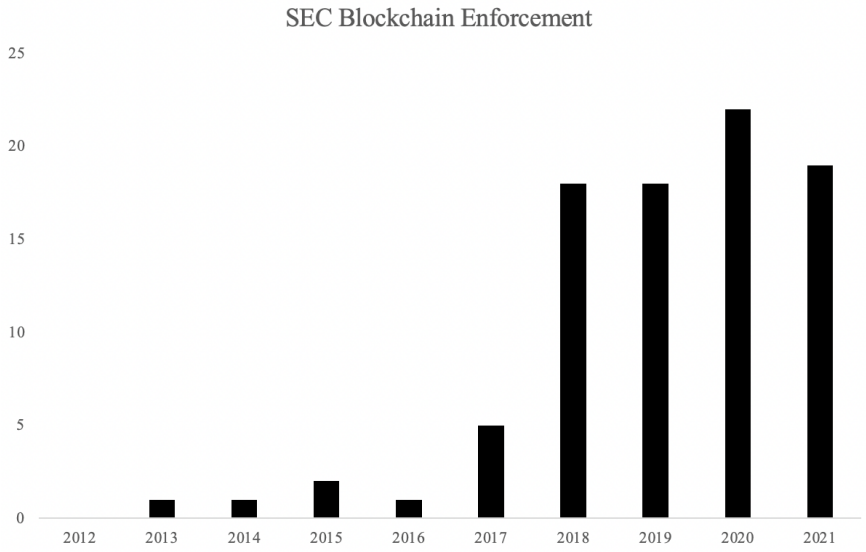
54. SEC Press Release, *SEC Charges Siblings in $124 Million Crypto Fraud Operation that included Misleading Roadshows, YouTube Videos*, SEC (Mar. 8, 2022), https://www.sec.gov/news/press-release/2022-37 [https://perma.cc/RKE3-VFZL].

55. Complaint at 3-4, Securities and Exchange Commission v. Barkdale and Barkdale, No. 01933 (S.D.N.Y Mar. 8, 2022), https://www.sec.gov/litigation/complaints/2022/comp-pr2022-37.pdf [https://perma.cc/G92X-7TCR].

addition to the civil penalties sought by the SEC, criminal charges were also brought in a parallel action by the United States Department of Justice (DOJ) against John Barksdale in this case.[56] Of 95 total enforcement actions sampled until May of 2022, 82 were only civil, 2 were criminal and 11 involved both civil and criminal charges.[57]

Importantly, not all digital assets are securities. In a June 2018 statement, the SEC declared that Bitcoin and Ethereum were decentralized enough, such that neither BTC nor ETH were considered securities.[58] Figure 2 graphs SEC blockchain-related enforcement actions by year.

Figure 2[59]



One goal for an effective compliance policy for cryptocurrencies is to use the public information available regarding enforcement actions to proactively structure an asset to distinguish it from assets the SEC alleges

56. Complaint, United States v. Barksdale, No. 00684 (S.D.N.Y Mar. 8, 2022), https://www.justice.gov/usao-sdny/press-release/file/1480836/download [https://perma.cc/RX4G-EF59]; *see also* SEC Press Release *supra* note 54, https://www.sec.gov/news/press-release/2022-37 [https://perma.cc/T57Q-ZPA7].

57. ChoiceCoin, Solana-Compliance, GITHUB (2022), https://github.com/ChoiceCoin/Solana-Compliance/blob/main/Database/Enforcement/SecuritiesEnforcement.xlsx [https://perma.cc/ULQ3-9TFP] (follow "View raw" hyperlink).

58. William Hinman, Director, Div. of Corp. Fin., Speech, *Remarks at the Yahoo Finance All Markets Summit: Crypto*, SEC (June 14, 2018), https://www.sec.gov/news/speech/speech-hinman-061418 [https://perma.cc/DL5X-66UD].

59. ChoiceCoin, *supra* note 57.

are securities.[60] For example, consider the different factors between Bitcoin and Ormeus Coin to better understand the law and inform policy.

C. *Policy*

Probably the biggest policy challenge for regulating digital asset securities[61] is defining the word security.[62] Most digital assets are not securities because most digital assets do not produce any profits solely from the efforts of others and often lack a common enterprise. In the case of decentralized assets not on a centralized exchange, any profits coming from the asset are only derivable from active participation in a decentralized protocol.[63] As another example, the types of characteristics measured in decentralization, such as supply, token distributions, and liquidity, are extremely volatile and vary greatly in relation to the existence of a common enterprise.

Defined digital asset securities are only those assets that represent an equity interest in a company or common enterprise and are sold as investments. By contrast, non-security digital assets are any intentionally decentralized assets intended for use within a product or service or operating as a cryptographic key. Clearly defining non-security tokens as assets not regulated by the SEC will ensure that opportunities can remain for open-source software projects developing blockchain technologies and digital assets.

Most digital assets should not be considered securities because they are not what we typically ascribe to the word security, such as a stock or a mortgage. Instead, digital assets are new technologies with a plethora of properties separate and apart from financial investments.[64] One of the key things that differentiates digital assets from securities is that digital assets can maintain their value irrespective of price.[65] This is because all

---

60. Jurisdiction may provide an important part of a compliance analysis, whether a project is domestic, foreign, or even what state or Federal Circuit a project principally operates because federal law is not static across the country when some Courts are legislating from the bench on behalf of the SEC. For Enforcement actions by SEC office, see Appendix B.

61. There are two types of securities: equities and debt. More generally, most cryptocurrencies and digital assets are not securities, money, or debt. Instead, this new asset class is something completely new that cannot be forced e*x post* into an existing framework of legal analysis.

62. Lummis-Gillibrand Responsible Financial Innovation Act, S.4356, 117th Cong. (2022) (defining ancillary assets, a new class of assets, as a specific type of security token having additional properties that yield additional regulation).

63. This is a lot of work and not consistent with traditional conceptions of securities, such as buying stock—which can be inherently passive.

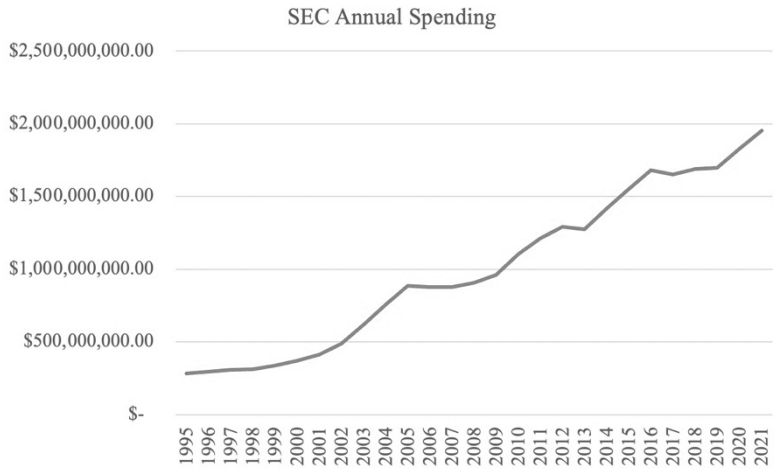
64. It's imperative this reality is respected by the Bill and any new law on the subject.

65. For example, if you buy a stock in $MSTR at 281.92 and sell it at 282.92, then you made one dollar and if you sell at $280.92, then you lost one dollar. However, with digital assets, if you buy $BTC at $22,010.10, your gains or losses are just as much dependent on what you do with your $BTC as they are the price of the asset.

digital assets are potentially revenue-producing assets, independent of price.[66]

One problem is that the SEC has a financial interest in arguing that most digital assets are securities.[67] The more the SEC has authority to regulate, the more money Congress will appropriate to the agency, but SEC spending is already out of control. Figure 3 shows the annual spending by the SEC. Between 2011 and 2021, the SEC overspent on its congressionally appropriated funds by more than 275 million dollars.[68]

Figure 3[69]



Still, despite holding positions of public service, SEC employees reap a fortune through annual compensation. For example, in the year 2020, the SEC had 4,495 employees with an average salary of $200,613.09 among all employees, not including paid time off and full benefits.[70]

66. For example, staking is a method for decentralization rather than a promise of profits because it helps distribute the asset fairly across a network. Instead, the network and the asset are decentralized in control, use, and development. Thus, there is no common enterprise.

67. U.S. Sec. & Exch. Comm'n, Fiscal Year 2023 Congressional Budget Justification and Annual Performance Plan; Fiscal Year 2021 Annual Performance Report (Mar. 28, 2022),https://www.sec.gov/cj [https://perma.cc/35F3-GMYL] (last modified July 21, 2022).

68. *Id*.

69. U.S. Sec. & Exch. Comm'n, Budget History—BA vs. Actual Obligations ($ in 000s), https://www.sec.gov/foia/docs/budgetact [https://perma.cc/XT6Y-S3TP] (last modified Nov. 3, 2019).

70. U.S. Sec. & Exch. Comm'n, Securities and Exchange Commission Salaries of 2020, https://www.federalpay.org/employees/securities-and-exchange-commission.

By arguing that more things are securities, such as digital assets, the SEC hopes to increase its authority and budget.[71] Unsurprisingly, the SEC cites regulating digital assets to support requesting a budget increase for 2023 to a total of over $2.17 billion.[72] Yet, security tokens are actually few and far between.[73] The SEC should be incentivized to reduce spending rather than increase spending. Moreover, the SEC should be incentivized to respect and promote the public good rather than its own bottom line.

Open-source software[74] is a public good.[75] Open innovation is a fundamental phenomenon that drives blockchain technology. Moreover, the transition from proprietary financial technology to open-source financial technology on blockchains serves the public by promoting financial transparency. Moreover, using a public ledger for asset dissemination provides a major combatant against rampant government agency fraud, abuse, and wasteful spending.[76] Thus, the open nature of blockchain technology is a critical public good.

A major part of the Solana Foundation's mission is the development of open-source software projects, which benefit the public by providing open, transparent, and affordable access to financial technologies and innovation. In fact, Solana is the largest open-source blockchain not based on proof-of-work technology. This adds additional public benefit

---

71. But the collective is a network of individuals with their own respective interests and motivations. *See* MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION 7 (1971) (arguing the State's members often have interests separate and apart from the people).

72. U.S. Sec. & Exch. Comm'n, Fiscal Year 2023 Congressional Budget Justification and Annual Performance Plan; Fiscal Year U.S. Sec. & Exch. Comm'n, Fiscal Year 2021 Annual Performance Report (Mar. 28, 2022), https://www.sec.gov/files/FY%202023%20Congressional%20Budget%20Justification%20Annual%20Performance%20Plan_FINAL.pdf [https://perma.cc/UP33-M4JC].

73. Only digital assets offered through an express ICO should be considered security tokens. An ICO is a specific type of action where a project backs a new asset with equity and then sells the asset to the public. Very few projects use an ICO, and they are generally vulnerable to much higher regulatory scrutiny for good reason. In fact, most projects decentralize assets through other mechanisms, removing any common enterprise or any expectation of profit.

74. Often, open-source software projects evolve on the decentralized Internet to build applications.

75. The two most abundant licenses for open-source software on the Solana Network are the Apache License and the MIT License. *See* The MIT License, *supra* note 47; *see also* APACHE LICENSE, VERSION 2.0 (2004), http://www.apache.org/licenses/LICENSE-2.0 [https://perma.cc/Q3MF-8K9M].

76. Craig Whitlock, Bob Woodward, Pentagon Buries Evidence of $125 Billion in Bureaucratic Waste, WASH. POST (Dec. 5, 2016), https://www.washingtonpost.com/investigations/pentagon-buries-evidence-of-125-billion-in-bureaucratic-waste/2016/12/05/e0668c76-9af6-11e6-a0ed-ab0774c1eaa5_story.html [https://perma.cc/EZ7W-T32F]; *see also* NATO, Defense Expenditure of NATO Countries (2012-2019) (June 25, 2019), https://www.nato.int/nato_static_fl2014/assets/pdf/pdf_2019_06/20190625_PR2019-069-EN.pdf [https://perma.cc/FP4E-VLKN].

because proof-of-work blockchains are notoriously less energy-efficient and environmentally friendly.

There should be a presumption of non-security for open-source software projects. In other words, it should be the SEC's burden to prove an asset is a security in civil court. Moreover, that burden should be beyond a reasonable doubt given that, in many cases, the federal government is suing a private citizen or small business.[77] This would allow for open innovation to persist and protect decentralized projects, developers, and entrepreneurs from unnecessary regulatory risk and illegitimate enforcement. At the same time, it would allow the SEC to focus its efforts on only those digital assets that are securities and where actual fraud occurs.

With respect to code, open-source software programs using digital assets for various purposes or in a decentralized way are also a public good. Open-source software projects forgo the ability to drive high-profit margins from proprietary software development and instead focus on product creation for the public good. Most open-source projects are also decentralized because anyone around the world can contribute. Additionally, assets associated with open-source projects are more likely to be used as tools rather than passive investments. Moving forward, legislation for blockchain technologies should respect the confluence of open-source software and the public good.[78]

III. APPLIED COMPLIANCE

Compliance is a process by which companies follow the law. *The Compliance Process* distills the corporate compliance function to foundational formalism.[79] In doing so, Duke Law Professor Veronica Root Martinez "demonstrates how focusing on process reforms will allow complex organizations to adopt more integrated and complex compliance programs that are better equipped to address corporate misconduct."[80] Compliance is important for blockchain projects and startups. This Part applies compliance to analyze various assets on the Solana blockchain from a securities perspective. Additionally, this Part discusses digital assets with respect to ethics, scalable use, and the role of regulation.

---

77. In fact, the SEC does not have the authority to prosecute individuals criminally and instead works with the DOJ to prosecute certain cases of Fraud criminally.

78. The most recent legislation relating to blockchain technology is the Lummis-Gillibrand Responsible Financial Innovation Act, which was introduced to the Senate Finance Committee by Senators Lummis and Gillibrand. *See* S. 4356, 117th Cong. (2022).

79. Veronica Root, *The Compliance Process*, 94 IND. L.J. 203 (2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3151893 [https://perma.cc/5EBC-NJF3] (reasoning that fundamental principles about the role compliance plays within the firm to create a formative model for organizational excellence).

80. *Id.*

A. *Legal Informatics*

Legal informatics is an approach to law based on information theory. Both a method of practice and theory, legal informatics focuses on the confluence of computer science and the law. The idea is to build information systems that improve law practice in terms of time, accuracy, and efficiency. In theory, legal informatics provides means for statistical analysis of the law through natural language processing to help address fundamental issues in jurisprudence, such as defining law and understanding the basic linguistic mechanisms underpinning law practice. Harvard Law Fellow Ron Dolin argues one method of formalizing human intuition in decision-making is a weighted geometric mean.[81]

Whether an asset is a security or non-security asset depends on various factors and proper analysis of certain attributes associated with the asset. Using legal informatics for securities compliance, the analysis may be conducted using defined variables. The variables may be weighted and processed to produce probabilistic measurements. Measurements may correspond with a number between 0 and 1, where an asset with a score of 1 is statistically unlikely to be a security, and an asset with a score of 0 is likely to be a security.

Equation 1

$$compliance\_score = \sqrt[\Sigma_{j=1}^{n} w_j]{\prod_{i=1}^{n} F_i^{W_i}}$$

In Equation 1, the compliance factors F_i may be assigned based on various features for a specific asset or regulatory corpus. For example, one factor to consider may be utility because if a cryptocurrency is used for governance or voting, it is almost certainly not a security token.[82] For purposes of applied analysis on Solana, the following nine factors were used to calculate the probability certain SPL tokens would be considered securities.

1. Equity: An asset is less likely to be a security if it does not represent an equity interest in a company.

2. Decentralization: An asset is less likely to be a security if it is decentralized.

---

81. Ron A. Dolin, *Measuring Legal Quality*, HARV. L. SCH., CTR. ON THE LEGAL PRO. (June 18, 2017), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2988647 [https://perma.cc/U2JZ-5AL3].

82. Similarly, if a token is backed by or tied to the value of another asset and pays dividends to investors, then the token is likely a security. Factors may also take account of existing legal frameworks for securities analysis—for example, a scorecard approach. *See* Cryptocurrency Rating Council, About Our Asset Rating Framework, Importance of the Howey Test for Classifying Digital Assets (2021), https://www.cryptoratingcouncil.com/framework.

3. Participation: An asset is less likely to be a security if users earn the asset through participation.

4. Investment: An asset is less likely to be a security if it is not marketed or sold as an investment.

5. Utility: An asset is less likely to be a security if it has a specific utility.

6. Purpose: An asset is less likely to be a security if the asset has an intended purpose for use aside from financial return.

7. Control: An asset is less likely to be a security if the asset gives the user control over an organization's decision-making.

8. Derivatives: An asset is less likely to be a security if it does not offer users derivatives or cash returns.

9. Commonality: An asset is less likely to be a security if it is not dedicated to the furtherance of a common enterprise.

The algorithm may be applied using a cognitive computing framework,[83] a collaborative process allowing humans and computers to perform the kinds of intelligent activities that they perform best.[84] The basic idea provides a means for cognitive information manipulation, which is required for commonsense reasoning.[85]

83. Jeanne C. Fromer, *Machines as the New Oompa-Loompas: Trade Secrecy, the Cloud, Machine Learning, and Automation*, N.Y.U. L. REV., 706, 720 (2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3359746 [https://perma.cc/N9YE-P723] ("In recent years, these techniques have been among the most successful and prominent ways of imbuing computers with artificial intelligence, or human-like cognitive abilities."). *See also* Emily Berman, *A Government of Laws and Not of Machines*, 98 B.U. L. REV. 1277, 1278 (2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3098995 [https://perma.cc/5FEV-85TV] (Machine learning is a strand of artificial intelligence that sits at the intersection of computer science, statistics, and mathematics, and it is changing the world.).

84. Dean Alderucci, *The Automation of Legal Reasoning: Customized AI Techniques for the Patent Field*, 58 DUQ. L. REV. 73, 74 (2020); *see also* Olga Russakovsky et al., *Best of both worlds: human-machine collaboration for object annotation*, "IEEE Xplore" (2015), https://ieeexplore.ieee.org/document/7298824 [https://perma.cc/X673-WV4L]; KEVIN D. ASHLEY, ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS: NEW TOOLS FOR LAW PRACTICE IN THE DIGITAL AGE 22 (2017).

85. David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653, 717 (2017) (explaining collaboration between lawyers and technologists will be key for tackling some of the most intractable problems at the juncture of law and Machine Learning).

Figure 4[86]

Project	Asset Name	Compliance Score[87]
Solana	SOL	0.84
Marinade	mSOL	0.84
Orca	ORCA	0.78
Serum	SRM	0.78
Raydium	RAY	0.75
Solice	SLC	0.64
Oxygen	OXY	0.63

Here, the algorithm may be used to assess the compliance of SOL, the Solana layer-1 asset, and SPL Tokens. The human expert must analyze and process each factor and make an expert scoring.

Figure 5

Assessment	Range	Explanation
Minimal risk	0.77 +	Strongest level of compliance.
Mild risk	0.61 – 0.76	Satisfactory level of compliance.
Moderate Risk	0.44 – 0.60	Mild risk of illegal activity or lacking in compliance mechanisms.
High Risk	0.00 – 0.43	Moderate to severe risk of illegal activity, no compliance mechanism, or other critical failure.

Figure 5 provides interpretive guidance for compliance score calculations and analysis. Moreover, the guidance allows for further ordinal categorization of projects from a securities compliance perspective. Still, all organizations should strive for organizational excellence in both compliance and ethics.

86. Choice Coin, Solana-Compliance, Software, v1, GitHub (2022), https://github.com/ChoiceCoin/Solana-Compliance/blob/bce4bcb05f24381c158bed7dbd6bea6e4145365b/Software/v1/SolanaStatisticalCompliance.py [https://perma.cc/UH24-9NHP] (the software algorithm is available on an open-source basis under the Apache License on GitHub).

87. All scores are estimates for academic purposes only.

B. *Ethics*

All else being peripheral, projects need to intend to follow the law and run their operations the right way. An honest commitment to ethics is critical to developing effective compliance mechanisms. Ethics are principles governing human behavior.[88] The study of ethics[89] is inherently limited by the subjective nature of personal ethics.[90] Indeed, what one person finds to be unethical may be considered entirely appropriate by another.[91] The evolution of ethical norms across the decentralized Internet is progressing at slower rates, which are in part dependent on ideological shifts supporting stronger ethical codes.[92]

One important ethical consideration is that of public waste due to the high costs associated with enforcement.[93] Regulation stifles competition by picking winners and losers based on capital allocations but does not yield a net public good because blockchain regulation otherwise falls under the Law of the Horse.[94] Moreover, additional regulation would likely stifle opportunity, not only for development but also for decentralizing economic freedom.[95]

Another important ethical consideration is the development of compliance and ethics programs on decentralized projects across the World Wide Web. The fact is that projects with ethics and compliance programs will put both the organizational and software infrastructure in place to succeed over the long hall. Perhaps one of the most amazing things about blockchain technology is that the decentralized network is *de facto* immutable. Given the importance of blockchain technology, a bedrock to building every project should be professionalism, compliance, and ethics.

Professor Veronica Root Martinez is a zealous advocate for moral and ethical courage in compliance. In *More Meaningful Ethics*, she argues for reconceiving the role ethics plays in modern business.[96] As explained,

---

88. Thomas M. Madden, *Law and Strategy and Ethics?*, 32 GEO. J. LEGAL ETHICS 181, 200 (2019) (discussing law firm competition).

89. Alternatively, ethics are more often used as a justification for maintaining socio-economic order.

90. Veronica Root Martinez, *More Meaningful Ethics*, U. CHI. L. REV. ONLINE 1, 3 (2020) [hereinafter Martinez, *More Meaningful Ethics*].

91. *Id.* at 6.

92. MARYAM JAMSHIDI, THE FUTURE OF THE ARAB SPRING: CIVIC ENTREPRENEURSHIP IN POLITICS, ART, AND TECHNOLOGY STARTUPS, 27 (2014).

93. Veronica Root Martinez, *Coordinating Compliance Incentives*, 102 CORNELL L. REV. 1003, 1029 (2017) (discussing regulatory agencies deficiencies in information and coordination).

94. Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501 (1999).

95. *See generally id.* ("The argument so far is that law can change the constraints of code, so that code might regulate behavior differently.").

96. Martinez, *More Meaningful Ethics*, *supra* note 90, at 54.

some legal scholars speculate ethics and compliance are separate and distinct concepts with a gray boundary.[97] However, Professor Martinez goes further, exploring perspectives on ethical relativism.[98]

Professor Martinez argues for developing ethical infrastructures within firms, promoting a more moral corporate culture.[99] Indeed, she contends firms should not retreat from difficult ethical dilemmas but rather should engage directly by implementing specific and explicit ethical infrastructures. Moreover, she recognizes the necessity for custom compliance program creation to meet the specific needs of each unique firm.[100] This is particularly true when developing ethical and compliant practices on blockchains.

The process of creating an ethical infrastructure may not be easy, but given persistent scandals across the blockchain space, excellent projects and startups have a grand opportunity to excel. Therefore, a commitment to ethics is critical to every crypto compliance program. In fact, ethics are one of the main missing mechanisms in traditional finance, which serves primarily to aggrandize income inequality[101] in the United States. Thus, every blockchain and decentralized project should commit to excellence in compliance and ethics. The two go hand in hand. In fact, adopting compliance mechanisms is a key factor for developing scalable blockchain technology.

C. *Scalable Use*

Non-security tokens will continue to represent most digital assets and will be a keystone to the scalable adoption of blockchain technology. Two critical elements for creating a non-security token are decentralization and participation. With decentralization, assets lack a common enterprise and should fail the Howey Test. Similarly, when users actively participate in a decentralized protocol with an asset, their efforts generate profits of their own accord, and thus, the asset should fail the Howey Test. Singularizing decentralization and participation, the most important thing

---

97. *Id.* at 7 ("Legal academic scholarship discussing the interplay of ethics and compliance often leans more heavily on compliance than ethics.").

98. *Id.* at 9 (explaining that "[w]hat one person finds to be unethical may be considered entirely appropriate by another individual.").

99. *Id.* at 21 ("This Essay argues that it is time for firms to adopt explicit and specific ethical infrastructures within compliance programs.").

100. Martinez, *More Meaningful Ethics*, *supra* note 90, at 67 ("Each organization has its own unique structure, industries, risks, and concerns, and compliance programs regularly reflect that fact. Firms hoping to include more meaningful ethics norms within their ethics and compliance programs will need similar flexibility to implement ethical infrastructures that will work well for their particular firms.").

101. *See generally* R. von Gleichen et. al., Affordable childcare when you need it? Childcare opening hours in the context of the Childcare Act 2016 (2016), www.oxpolicy.co.uk (discussing market rates rising and the effect for costs in the context of childcare).

for new projects creating non-security assets is to focus on use. In other words, projects must create digital assets that are used as cryptographic keys on the blockchain rather than passive instruments for financial investment.

Moving forward, Solana is the leading blockchain for processing payments because its PoH model provides a cost-efficient method for transactions. Unlike Bitcoin and Ethereum transactions, Solana transactions are fast with de minimis network costs. These two elements allow Solana to have a competitive advantage compared to other networks, which will need additional scaling solutions to process payments affordably. Ultimately, the goal of scaling blockchain is to allow the blockchain to serve as a foundational financial infrastructure for generations to come. In the future, anyone will have access to the permissionless network and be able to pay for coffee, shop on Amazon, or even buy a car with digital assets on the blockchain.

What the blockchain industry needs is not more investors but rather more customers. A harsh criticism of blockchain technology is that blockchains are just another form of stagnation caused by speculative solutionism. But there isn't anything of substance or, more importantly, demand for technology. This may be true with proof-of-work blockchains because the underlying information technology is slow and expensive. In fact, for this reason, Ethereum is moving to a completely new token distribution model with Ethereum 2.0.[102]

Solana changes things. One of the great things about Solana is that its novel proof-of-history technology allows for faster and cheaper transactions than its predecessors, Bitcoin and Ethereum. For the way Edison is remembered for harnessing electricity in the light bulb, but Tesla is remembered for bringing electricity to the world—Satoshi is remembered for blockchain, but Solana is bringing blockchain to the masses.

One of the amazing things about the Solana blockchain is that, as an information technology, it has already reached a global scale less than a decade into its development. Globalization creates significant regulatory challenges for such a young technology,[103] and it is important that the United States continues to respect and foster both entrepreneurship and innovation in creating policy and legislative developments. It is important to recognize blockchains, like Solana, as public goods and to respect the reality that adding even more regulation to an already overregulated

---

102. Upgrading Ethereum to radical new heights (2022), https://ethereum.org/en/upgrades/ [https://perma.cc/43XN-X4WS].

103. Olia Kanevskaia, *The Law and Practice of Global ICT Standardization*, 116 (Mar. 31, 2020) ("But despite the challenges "open standards," such as inclusion of proprietary solutions into Internet specifications and lack of sufficient governmental recognition, this concept, as well as the OpenStand principles, cannot be ignored in the context of modern standardization.").

industry will only reduce the public benefit otherwise provided by blockchains as open access recourses.

CONCLUSION

In a recent Tweet, Chairperson Gary Gensler of the SEC said, "Let's not risk undermining 90 years of securities law."[104] Gensler's statement is correct insofar as policymakers and courts must respect that for over 90 years, the SEC has been limited to only regulating securities, which are debt and equities. Moreover, we cannot risk misclassifying the millennium's greatest financial technology as securities simply because we do not have an existing legal infrastructure to control it. Instead, we should respect the innovation occurring and only classify digital assets intended to be securities as securities.[105]

Blockchains have value as an information technology. In other words, blockchains have value because the databases of both globalized and decentralized information have value. This is fundamentally different than traditional securities, debt—which has value supported by a legally and physically attached instrument, such as a social security number or house; and equities which have value attached to revenue streams, interest, and dividends. So, consider a bright line rule, respecting individual property rights and the 5th Amendment of the United States Constitution,[106] digital assets are not securities unless their creators expressly intend the asset to be a security and the asset represents an expressly secured interest in equity or debt.

In conclusion, Part I provided an overview of Solana, the largest PoH blockchain, as measured by market capitalization. Part II discussed and described securities compliance and enforcement data in the blockchain space. Part III introduced new mechanisms for measuring blockchain compliance statistically and applied those new mechanisms to various assets on Solana. Moving forward, Solana is a public good, and information technology strives to be a staple for excellence in securities compliance, ethics, and economic engineering.[107]

---

104. Gary Gensler (@GaryGensler), Twitter (10:00 AM) (July 28, 2022), https://mobile.twitter.com/GaryGensler/status/1552700562533236739?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Etweet [https://perma.cc/KHH3-WUBR].

105. While we may not be able to say with certainty whether the SEC will argue a certain asset is a security, we can make predictions about the matter using legal informatics.

106. U.S. CONST. AMEND. V.

107. Veronica Root, *The Compliance Process*, 94 IND. L.J. 203, 216 (2019) ("Despite the focus by regulators and prosecutors on the importance of developing an effective compliance program, it is commonly understood that it would be inefficient for firms to strive to obtain "perfect" compliance.").

MOVING THE UNITED STATES INTO THE 21ST CENTURY FOR CHILDREN'S ONLINE PRIVACY RIGHTS

Zackary A. Blanton[*]

Abstract

It has been more than twenty-five years since the Children's Online Privacy Protection Act (COPPA) was first implemented in the United States. Since its enactment—well over a decade ago—there has been only one instance in which Congress successfully passed noteworthy modifications to the Act. While there has been a recent increase in proposed amendments to the Act better to protect children in our current reality of everchanging technology, little has been done to initiate the much-needed change. The increased focus on children's online rights has been sparked primarily by changes made in the United Kingdom. At the forefront of the drive for greater protection of the privacy rights of children, the United Kingdom's transformation has left the world considering what alterations need to be made to their current systems to stay up to date with this growing demand.

Despite the mounting need for change, online service providers have stalled the process, leaving children in a world of new technologies without adequate protections in place. As market giants, online service providers influence ongoing debates to limit legislative changes and the potential economic burden of those changes. Several scholarships have identified issues with the current system in the United States, but few have taken on the task of proposing a practical solution. To effectuate change, it is imperative to zero in on the most essential needs of children to adequately protect them online while balancing the concerns of those opposing large-scale modifications. This Note will begin by looking at the current law of child online privacy protections in the United States, COPPA, exploring how the act works, how violations are handled, and how the original version of COPPA has changed. Next, it will explore the approach recently taken by the United Kingdom and then evaluate how COPPA compares, as well as the discussions currently taking place regarding this topic. Lastly, this Note will set out a five-point plan to implement the necessary changes to bring children's online privacy protections into the 21st century.

---

* 2023 J.D. Candidate, University of Florida Levin College of Law. I would like to thank Professor Stacey Steinberg for being my faculty advisor and assisting me throughout this process. I would also like to thank my family, specifically my mother, Stacey Blanton, for encouraging me and assisting me throughout this journey. Lastly, I would like to thank Nadia Rossbach for always being a great friend and someone I could bounce ideas off. I could not have completed this work without the assistance of my friends, family, and mentors.

INTRODUCTION

Google was founded in 1998. Since then, there has been an array of innovative technologies, search engines, and social media developments, including Wikipedia in 2001, Facebook in 2004, YouTube in 2005, Twitter in 2006, the iPhone in 2008, and one of the most recent advancements to social media, TikTok in 2016.[1] These technological advancements over the last twenty-five years have been some of the most

1. Joshua Kim, *Technology Since 1998*, INSIDE HIGHER ED (Jan. 6, 2014), https://www.insidehighered.com/blogs/technology-and-learning/technology-1998 [https://perma.cc/T623-W8SE].

life-altering developments since the inception of the computer for both adults and children alike. However, since Google was founded, the United States has not implemented any new regulations for handling children's online privacy protections.[2] In fact, the most current updates since the creation of Google over two decades ago came about in 2012, marking another decade with minimal change.[3] This means that the decade-long gap since the last update to the regulations protecting children's online privacy goes back to before most of the children still considered protected under the regulation were born.[4]

The apparent negligence of the legislature and other involved parties is exactly what will be addressed in this Note, in addition to determining what advancements have been made in Great Britain and the changes that can be implemented now to ensure the online safety of our future generations. This Note will also explore why it has taken so long to change an obviously broken system and the efforts currently underway to help effectuate change in this area. Using social media platforms and other online service providers as a guide, the focus will be on exploring how to expand protections to include teenagers that are aged thirteen through seventeen. I will also examine other alternatives and resources for expanding the protection of children's online privacy rights by comparing the two policies at the forefront of these issues. The first is the current United States policy, the Children's Online Protection Policy Act (COPPA), and the second is the current policy in Great Britain, The Age-Appropriate Design Code (or the Children's Code), which was included in the 2018 Data Protection Act.[5] In this way, my analysis will shed light on the essential elements of the United States' policy in need of revision to bring the country into the 21st century with respect to online privacy protection for children.

---

2. *Id.*

3. *Id.*

4. The FTC issued a notice of proposed rulemaking to COPPA in 2011 and a supplemental notice of proposed rulemaking to COPPA in 2012. The FTC announced the publication of the amended rules to COPPA on December 19, 2012. Because of this the amendment is commonly referred to as the 2012 amendment to COPPA and will be referred to as such for the purpose of this Note. However, the amended rules to COPPA took effect starting on July 1, 2013. *See* Federal Trade Commission, *16 C.F.R. Part 312: Children's Online Privacy Protection Rule: Final Rule Amendments and Statement of Basis and Purpose* (Dec. 19, 2012), *available at* http://ftc.gov/os/2012/12/121219copp arulefrn.pdf [https://perma.cc/5D3K-8JJ8] (Final Rule and SBP); *see also* 16 C.F.R. § 312.

5. Byrin Romney, *Screens, Teens, and Porn Scenes: Legislative Approaches to Protecting Youth from Exposure to Pornography*, 45 VT. L. REV. 43, 45 (2020).

I. THE CHILDREN'S ONLINE PRIVACY PROTECTION ACT

COPPA was created in 1998 to help protect the personal information of children on the internet who are under the age of thirteen.[6] "COPPA applies to 'operators'[7] of commercial Web sites and certain other online services that are 'directed'[8] to children under thirteen."[9] The finding that an operator reaches children under the age of thirteen is not based on the actual express intent of the online service provider but rather on characteristics such as images or graphs used, the language used to reach individuals, and the presentation of the website as a whole.[10] Even if the service providers are not directing their attention specifically toward reaching children under the age of thirteen, as long as there is actual knowledge that providers are collecting personal information[11] from these children, then the online provider will still fall under COPPA.

---

6. 16 C.F.R. § 312.2.

7. After the 2012 amendment to COPPA, operator

> means any person who operates a Web site located on the Internet or an online service and who collects or maintains personal information from or about the users of or visitors to such Web site or online service, or on whose behalf such information is collected or maintained, or offers products or services for sale through that Web site or online service, where such Web site or online service is operated for commercial purposes involving commerce among the several States or with 1 or more foreign nations; in any territory of the United States or in the District of Columbia, or between any such territory and another such territory or any State or foreign nation; or between the District of Columbia and any State, territory, or foreign nation.

Id.

8. As long as an operator knowingly collects information from children in the United States then they are bound by COPPA. Even if a web-based operator is a foreign entity and they intend to reach children under the age of thirteen in the United States, they still fall under the parameters of COPPA. Anita L. Allen, *Minor Distractions: Children, Privacy and E-Commerce*, 38 HOUS. L. REV. 751, 760 (2001).

9. *Id.* Many questions arise as to why the cutoff age is thirteen when there are so many other programs, like FERPA, where parents can still access the school record of a child under the age of eighteen even if the teen objects. *Id.* at 759. The reasoning given by the FTC is limited, "that the age of thirteen is the standard for distinguishing adolescents from young children who may need special protections." *Id.* Nevertheless, the FTC fails to explain why it would assume that children between the ages of thirteen and seventeen do not need such protection and also that those children would fully understand the negative ramifications of revealing private personal information to operators of online services. *Id.*

10. *Id.* at 760–61.

11. Personal information in the eyes of COPPA "is defined broadly to include a person's name, address, e-mail address, phone number, social security number, and any other identifier deemed to enable physical or online contact." Allen, *supra* note 8, at 761.

A. *The Requirements for Online Service Providers*

COPPA has five requirements that must be met in order to comply with the regulation: notice, verifiable parental consent, parental review, security, and limits on the use of games and prizes.[12] To fully understand what COPPA truly entails, it is important to break down each component individually. To start, online service providers must provide notice to the parents of children who want to access the websites that collect information about users before any of the children's information is collected.[13] The notice requirement must provide parents with the following information:

> (1) "a description of the specific types of personal information collected from the child by [the] operator"; (2) "the opportunity at any time to refuse to permit the operator's further use or maintenance . . . of personal information from that child"; and (3) "a means that is reasonable . . . for the parent to obtain any personal information collected from that child."[14]

Further, this information "must be within the four corners of the notice . . . [c]ompanies must also send this notice directly to the parent and must post a prominent and clearly labeled link to an online notice of its information practices . . . ."[15] This is important because it allows parents to continuously regulate what type of information a site obtains so that even if certain personal information is revealed by the child without the parent's knowledge, the parent can attempt to have the information removed.[16]

The parental consent and review requirement involves gaining the consent of the parent in a verifiable way and giving the parent a reasonable avenue for reviewing the personal information collected on the child.[17] COPPA does not specifically outline a defined mechanism for obtaining this consent.[18] Therefore, "[t]he operator of a Web site may obtain parental consent online and verify that consent via e-mail or

---

12. Tianna Gadbaw, *Legislative Update: Children's Online Privacy Protection Act of 1998*, 36 CHILDREN'S LEG. RIGHTS J. 228, 228 (2016).

13. Gianna Korpita, *It's a Small World After All: How Disney's Targeted Advertisements Implicate COPPA*, 19 J. HIGH TECH. L. 407, 414–15 (2019).

14. Allen, *supra* note 8, at 763.

15. Korpita, *supra* note 13, at 417.

16. *See Complying with COPPA: Frequently Asked Questions*, FTC (July 2020), https://www.ftc.gov/business-guidance/resources/complying-coppa-frequently-asked-questions [https://perma.cc/Q2KX-BRRS].

17. Shannon Finnegan, *How Facebook Beat the Children's Online Privacy Protection Act: A Look into the Continued Ineffectiveness of COPPA and How to Hold Social Media Sites Accountable in the Future*, 50 SETON HALL L. REV. 827, 831 (2020).

18. Allen, *supra* note 8, at 761.

telephone if the personal information is used only internally."[19] There are certain exceptions to the requirement of parental consent.[20] For one, an online service provider can gather personal information if it is used "to protect the safety of children, the security of the site, and to satisfy the demands of law enforcement."[21] Operators may also, on a one-time basis, collect only email addresses from a child in order to process the request as long as such information is properly deleted afterward.[22] Another important point of distinction is that COPPA only regulates commercial sites. If such sites are not considered commercial for the purpose of COPPA, they are not restricted.[23]

For purposes of the security requirement, the language of the regulation states that "[a]n operator shall not be considered to have collected personal information under this paragraph if it takes reasonable measures to delete[24] all or virtually all personal information from a child's postings before they are made public and also to delete such information from its records. . . ."[25]

COPPA also states that operators must use "reasonable security" measures to protect personal information.[26] However, neither the Federal Trade Commission (FTC) nor the statute specifically define what this entails.[27] Instead, operators are left with the suggestion "to minimize the amount of data collected from children, retain this data for as short a period as possible, and make certain that any third parties who access this data maintain strong security."[28] Consequently, the guidelines leave loopholes for operators and allow them to make their own rules when it comes to the reasonable security requirement under COPPA.[29]

The last requirement is the limit on the use of games and prizes.[30] This limitation consists of prohibiting operators from using incentives that lead to a large influx of private and personal information from the children who play such games due to the appeal these incentives have on influencing the child's decision to take part in the activity.[31] In other

19. *Id.*
20. *Id.*
21. *Id.*
22. *Id.*
23. *Id.* at 762.
24. "Delete means to remove personal information such that it is not maintained in retrievable form and cannot be retrieved in the normal course of business." 16 C.F.R. § 312.2.
25. *Id.*
26. Jeremy Greenberg, *Dangerous Games: Connected Toys, COPPA, and Bad Security*, 2 GEO. L. TECH. REV. 170, 176 (2017).
27. *Id.*
28. *Id.*
29. *Id.*
30. Emily DiRoma, *Kids Say the Darndest Things: Minors and the Internet*, 2019 CARDOZO L. REV. DE NOVO 43, 53 (2019).
31. *Id.* at 46.

words, the limitations on the use of games and prizes provide that operators can only acquire personal information that "is reasonably necessary to participate in the activity."[32] Again, this allows operators to determine what is reasonable in terms of the information they acquire relative to the use of games and prizes.

B. *The Interworking of COPPA and How Violations Are Addressed*

COPPA allows the FTC to act against violators of COPPA,[33] specifically the "operators" of websites and other online services.[34] However, a preemption provision in COPPA restricts private parties from filing a claim under statutes pertaining to state consumer protection.[35] Further, COPPA explicitly states that "[n]o State or local government may impose any liability for commercial activities or actions by operators in interstate or foreign commerce in connection with an activity or action described in [this regulation] that is inconsistent with the treatment of those activities or actions under this section."[36] In other words, state and local governments cannot bring action against online service providers under state consumer protection laws.[37]

In addition, COPPA limits the Attorney General from producing claims that fall under state consumer protection laws, or the equivalent, that interfere with COPPA.[38] Different courts have interpreted this provision in different ways.[39] For example, the Courts of Appeal for the Third Circuit determined that a claim could be brought as long as the operator was deceptive in how the children's information was acquired "as to create a false expectation of privacy."[40] However, COPPA allows the Attorney General "to bring suit to enjoin practices in violation of the statute, enforce compliance, obtain damage, restitution or other compensation on behalf of residents of the applicable state, or obtain other such relief as a court may deem appropriate."[41] Therefore, it is possible for the Attorney General to bring legal action against the operators under certain limited circumstances.[42]

Ultimately, legal action regarding children's online privacy is taken at the federal level, primarily through the Federal Trade Commission.[43]

---

32. Allen, *supra* note 8, at 764.
33. 16 C.F.R. § 312.2.
34. 1 Robert Brownstone & Tyler Newby, *Data Sec. & Privacy Law* § 9:89 (2022–2023).
35. *Id.*
36. *Id.*
37. *Id.*
38. *Id.*
39. *Id.*
40. *Id.*
41. *Id.*
42. *Id.*
43. *Id.*

Unfortunately, the reality is that it took the FTC three years to bring any action after the COPPA amendments were codified in 2013, and the Attorneys General rarely, if ever, use their power to bring claims against operators.[44]

C. *The 2012 Amendments to the Children's Online Privacy Protection Act*

Until December 2012, online privacy protections for children were handled through the 1998 version of COPPA. However, in 2012, the FTC amended COPPA due to the expansive development of technology in the new century.[45] With the intent of creating impactful changes to the way we handle children's online privacy, the modifications served only to resolve ambiguities from the 1998 version, along with a few other minor updates.[46] For example, the 2012 version redefined operators,[47] websites, and/or online services directed toward children and personal information.[48] The change in the definition of personal information has provided "parents additional control over the collection of their children's data."[49]

The 2012 changes also kept children's online information more secure.[50] The 2012 amendment to COPPA limits operators from keeping personal information of children "only as long as reasonably necessary."[51] When an operator decides the information is no longer needed, operators have the duty to use reasonable measures to protect the information from unauthorized access.[52] This is different from the 1998 version of COPPA, where operators were not instructed to discard information when no longer needed, but rather, they were left to decide what to do with the information. Furthermore, the new law clarified that operators must take "reasonable steps to release personal information

44. *Id.*

45. Gadbaw, *supra* note 12, at 229.

46. *Id.*

47. In the 2012 amendment, Operator includes any "operator of a child-directed site or service where it allows outside services to collect personal information from its visitors." This allowed an ongoing issue to be resolved where third parties were collecting the personal information of children on behalf of the online service providers. *Id.*

48. In the 2012 amendment, personal information "was re-defined to include 'geological information as well as photos, videos, and audio files of a child's image or voice.'" This allowed physical information to be protected as well. *Id.*

49. Diana S. Skowronski, *COPPA and Educational Technologies: The Need for Additional Online Privacy Protections for Students*, 38 GA. ST. U. L. REV. 1219, 1230 (2022).

50. Gadbaw, *supra* note 12, at 229.

51. *Id.*

52. *Id.* at 229–30.

only to service providers and third parties who are capable of maintaining the confidentiality, security, and integrity of such information."[53]

Another noteworthy improvement relates to the use of "safe harbor programs."[54] Online service providers who wish to take advantage of safe harbor provisions are now required to "conduct annual comprehensive reviews of their member's information practices and submit to the FTC annual reports of the results of these annual reviews."[55] Further, the 2012 revisions gave service providers other ways of acquiring parental consent.[56] For instance, it is now permissible for companies to acquire parental approval "through electronic scans of signed parental consent forms, videoconferencing, use of government-issued ID, and alternative payment systems."[57] Operators can also attain approval by adhering to "a 120-day notice and comment process conducted by the FTC."[58] With these new methods, companies can match faces to different forms of personal identification of the parents to acquire the consent needed.[59] The amendments to the 1998 version of COPPA have helped to make impactful changes to children's online privacy. Since then, the United States has fallen behind in comparison to other countries, like the United Kingdom, that have made substantial changes to keep up to date with the growing number of technological advances.

II. THE UNITED KINGDOM POLICIES ON CHILDREN'S ONLINE PRIVACY: THE AGE-APPROPRIATE DESIGN CODE (THE CHILDREN'S CODE)

Due to the growing number of children being exposed to the internet, the need for increased protection for children's online privacy rights has sparked action in the United Kingdom. The United Kingdom has become aware of the use of data collection by online service providers and the fact that the collection process begins once an individual downloads an application and commences to play or use the app.[60] They also recognized

53. *Id.* at 230. *See also* David R. Hostetler & Seiko F. Okada, *Children's Privacy in Virtual K-12 Education: Virtual Solutions of the Amended Children's Online Privacy Protection Act (COPPA) Rule*, 14 N.C.J.L. & TECH. ONLINE 167, 168 (2013) (stating that the 2012 amendment "strengthens regulation over website operators and by expanding COPPA's reach to mobile application developers and third-party vendors . . . .").

54. Under the 1998 version of COPPA, "the safe harbor provision encouraged industry self-regulation by allowing approved industry members to create their own COPPA oversight programs with their own compliance guidelines." Gadbaw, *supra* note 12, at 230. Furthermore, "[w]ebsite operators who participated in these approved safe harbor programs were subject only to the provisions of their own self-created and self-regulated safe harbor program in lieu of FTC enforcement." *Id.*

55. *Id.*

56. *Id.*

57. *Id.*

58. *Id.*

59. *Id.*

60. INFORMATION COMMISSIONER'S OFFICE, AGE APPROPRIATE DESIGN: A CODE OF

that one out of five users is a child.[61] Further, the amount of time all humans currently spend using services from online providers has grown exponentially.[62] The upsurge in time spent using these services has also augmented how this type of content is shaping the lives of everyone, especially children.[63] Without certain safeguards in place to protect them, the risk of harmful consequences is higher than ever.[64] Thus, the United Kingdom enacted the Age-Appropriate Design Code, or the Children's Code, becoming a force of law on September 2, 2020.[65]

The Children's Code outlines fifteen standards that companies must follow, keeping the child's best interest at the forefront.[66] The Code applies to all online service providers likely to be accessed by children in the country, which the United Kingdom calls information society services (ISS).[67] If a company is found not adhering to the guidelines of the Children's Code, the company would be in violation of the Privacy and Electronic Communication Regulation (PECR) and the General Data Protection Regulation (GDPR).[68] As a consequence of the violation, action may be taken against the company or organization, including "assessment notices, warnings, reprimands, enforcement notices, and penalty notices . . . . For serious breaches of the data protection principles, [the agencies] have the power to issue fines of up to €20 million . . . or 4% of [a company's] annual worldwide turnover, whichever is higher."[69] This can result in a hefty penalty for those who do not obey the standards; however, violators are often given a chance to rectify the issues associated with the violation.[70]

The Children's Code's specific standards include the children's best interest, data protection impact assessments, age-appropriate application, detrimental use of data, default settings, geolocation, parental controls, and online tools.[71] Children's best interest comes from the United Nations Convention on the Rights of Children (UNCRC).[72] Article Three of the Convention states that "[i]n all actions concerning children, whether

---

PRACTICE FOR ONLINE SERVICES, 9 (Sept. 2, 2020).

61. *Id.* at 3.

62. *Id.*

63. *Id.*

64. *Id.* at 30.

65. *Id.* at 32.

66. *Id.* at 7–8

67. "The definition of an ISS is 'any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services.'" This encompasses most for-profit online services and even includes electronic services for controlling connected toys and other devices. *Id.* at 16.

68. *Id.* at 5.

69. *Id.* at 12.

70. *Id.*

71. *Id.* at 7–8.

72. *Id.* at 24.

undertaken by public or private social welfare institutions, courts of law, administrative authorities or legislative bodies, the best interests of the child shall be a primary consideration."[73] Indeed, one of the main goals of this new regulation is to allow children more access to the Internet, which includes more access to information, more opportunities to interact with others, and more ways to further the promotion of their development through various forms of technology and games.[74] Further, relying on the best interest standard, the United Kingdom asserts that children should have the right to privacy and freedom from companies' economic exploitation.[75] The Code also incorporates another important standard: the detrimental use of data.[76] This standard is in place to ensure a child's personal data is not used in such a way that has been shown to be detrimental to the well-being of the child.[77] It also ensures that providers' policies do not contradict industry or government-set standards.[78]

Next, the data protection impact assessments (DPIA) standard is a seven-step assessment[79] with goals to "help you identify and minimize the data protection risks of your service—and in particular, the specific risks to children who are likely to access your service which arises from your processing of their personal data."[80] Under the GDPR, this type of assessment is required before starting any "type of processing that is likely to result in a high risk to the rights and freedoms of individuals."[81] One of the more important aspects of this seven-step process is the consultation with parents and children, which requires a hands-on approach to reviewing the risks to privacy associated with certain company protocols and conducting research from consumers of the online service to ensure they are aware of how personal information is being used.[82] This hands-on approach is instrumental in allowing operators to see exactly what kinds of activities are occurring within the companies regarding the collection of children's data.

The age-appropriate application is one of the most important standards in the Children's Code and one that other countries, including

---

73. *Id.*
74. *Id.*
75. *Id.*
76. *Id.* at 43.
77. *Id.* at 43–44.
78. *Id.* at 43.
79. The seven-step program includes: "identify[ing] the need for a DPIA; describe[ing] the processing; consider[ing] consultation; assess[ing] necessity and proportionality; identify[ing] and assess[ing] risks arising from your processing; identify[ing] measures to mitigate the risks; sign[ing] off, record[ing] and integrat[ing] outcomes. Importantly, the process was created to be a more flexible and scalable system." *Id.* at 27.
80. *Id.*
81. *Id.*
82. *Id.* at 28.

the United States, should implement to move into the 21st century where children's online privacy protection is concerned.[83] The application assesses the different needs of children based on each child's age level and stage of development.[84] Using this type of information, children are afforded the appropriate level of protection by allowing for flexibility in determining the proper standards based on the online services children are actually using.[85] Additionally, the Code gives online service providers a standard for all users so they do not have to assess what age bucket a child fits into that could potentially require a different form of protection.[86]

The delineated age periods include zero to five or pre-literate and early literacy, six to nine or core primary school years, ten to twelve or transition years, thirteen to fifteen or early teens, and sixteen to seventeen or approaching adulthood.[87] However, it is important to note these are not the required age ranges or classifications but rather a guide as to what age groups may need a different protection category.[88] This concept also allows the online service provider to use any method necessary to determine a user's age as long as the information is obtained accurately.[89] A few methods to determine the user's age include self-declaration, artificial intelligence, third-party verification services, account holder confirmation, technical measures, and hard identifiers such as formal documents, like a passport.[90] However, as innovative and creative as these methods may be, online service providers have been reluctant to implement these methods due to the additional cost and time.

The next standards are the default settings and geolocation, which highlight the idea that the use of certain settings ensures online privacy protection.[91] This means setting a "high privacy" standard as a default unless a company can provide a compelling reason why the standard should be different.[92] Likewise, for geolocation, the standard of the Children's Code is for those settings to be turned off in order to protect the child's location.[93] The default settings are simple aspects of the standards that can profoundly affect ensuring children are protected from inadvertently oversharing personal information by simply using the tools within the provider's application or program.

---

83. *Id.* at 32.
84. *Id.*
85. *Id.*
86. *Id.* at 32–33.
87. *Id.*
88. *Id.* at 32.
89. *Id.* at 33.
90. *Id.* at 34.
91. *Id.* at 5.
92. *Id.* at 7.
93. *Id.*

Parental controls provide another layer of protection within the arsenal of standards while also ensuring children can freely express themselves on the internet.[94] These types of controls provide age-appropriate information to children regarding how the parents monitor their use of certain applications.[95] The idea is that, depending on the child's age, if a parent is given access to monitor the child's activity or track their location, he/she should be made aware that a parent is monitoring them.[96]

The final standard involves the use of online tools to ensure that children have the proper resources needed so they are able to exercise their data protection rights and report any concerns regarding their personal information.[97] These standards can be used by other countries, especially the United States, as a guideline to understanding the methods and ideas implemented in other areas that could be helpful in making impactful changes to how children's online privacy is treated.[98]

III. TIME FOR CHANGE: THE CURRENT DEBATE ON REVISIONS TO COPPA AND WHAT IS CURRENTLY BEING DONE

A lot of the changes that have been implemented around the world, like the Children's Code in the United Kingdom, have not been passed free from debate. After all, there is a reason the last change to COPPA was over ten years ago, despite the growing number of technological advancements. Both sides of the debate have valid reasons and viewpoints as to why certain changes to COPPA should or should not be implemented. To truly advocate for substantive change to COPPA, both sides must be discussed, and the arguments for and against should be fleshed out.

A. *Potential Downsides of Revision to COPPA and What Can Be Done to Counter the Issues*

One issue that commonly emerges is the idea that, given the current restrictions of COPPA, children have been removed from certain online platforms, impairing their ability to freely express themselves on the internet, especially children under the age of thirteen who are currently affected by COPPA.[99] Often, online service providers take the easy and sometimes cheaper way out when adhering to the regulations of COPPA

94. *Id.*
95. *Id.* at 10.
96. *Id.* at 40.
97. *Id.* at 8.
98. *Id.* at 3.
99. Sasha Grandison, *The Child Online Privacy Protection Act: The Relationship Between Constitutional Rights and the Protection of Children*, 14 U. D.C. L. REV. 209, 219 (2011).

by simply banning use by children under the age of thirteen.[100] These online service providers recognize that they will be able to withstand the "missed opportunity" of not allowing children under the age of thirteen to join because children are not easily thwarted by a simple age verification screen.[101] In other words, children simply lie about their age to circumvent this barrier.[102] As a result, children are now truly unprotected when it comes to these sites acquiring their private information, similar to the circumstances going back to the mid-nineties before there was any protection.[103]

Individuals and groups opposed to revisions to COPPA raise concerns that extra regulations will further hinder the ability of youth to access the internet and freely express themselves without government intervention.[104] This concept may seem reasonable to outsiders who are not familiar with COPPA and other protections that have been executed globally, but to those who truly understand what increased privacy protection for children will do, this is not the case at all. In fact, it will do exactly the opposite.[105] The practical effect of regulation around online privacy is not to stop children from participating online or using the applications of online service providers. Instead, it allows children to play and interact freely on the internet without fear, or even worse, the lack of fear due to ignorance.[106]

Another critique of COPPA and any further revision is the idea that an increase in restrictions that cause companies to implement safeguards creates an economic burden.[107] Thus, small businesses, specifically those in the midst of growth, are now affected at the front end and unable to afford the cost of putting proper protections in place as required by law.[108] This can inadvertently lead to online service powerhouses that control a majority of the market, stifling startup companies and essentially creating a monopoly of large companies that control everything.[109] This includes the power to force and push through legislation that will allow these powerhouses to gain even more strength in their respective markets and

100. DiRoma, *supra* note 30, at 61.

101. *Id.*

102. *Id.*

103. *Id.*

104. EY UK Privacy and Data Governance Channel, *Protecting Children Online: The Age Appropriate Design Code*, at 18:23 (Nov. 16, 2021), (downloaded using Spotify) https://open.spotify.com/episode/5t9AVgXhaQ2FCdrwe34NAl?si=Zfxqu5JJRQKqj1MPdE3PKQ [https://perma.cc/9S8X-GR9D].

105. *Id.* at 06:30.

106. *Id.*

107. *Id.*

108. *Id.*

109. *Id.*

to further absolve them of their responsibility to protect the online privacy of our youth.[110]

To illustrate, the Children's Code in the United Kingdom has implemented more restrictions and heightened regulatory requirements that online service providers must follow to comply with the new laws.[111] The commonly mentioned concern with the Children's Code is related to the same issue of stifling development and impeding small businesses from flourishing.[112] This is a valid concern because the law involves increased regulatory requirements and a push for an age-appropriate standard, increasing operator costs.[113] Still, this issue can be absolved by using the different resources the United Kingdom has made available to assist with these problems.[114] One such resource is a technical standard published by the British Standards Institute, created for the purpose of training companies on how to perform an identity attribute check to verify a user's age.[115]

The standard verifies an assertion of parental responsibility in a way that does not violate children's privacy and still adheres to the requirements of the Children's Code by only collecting data on a temporary basis.[116] Companies are wary of the technical implications of the standard.[117] However, the technical aspect is mostly API integrations,[118] which are common in credit reference agencies, so this is not a new notion.[119] It should be noted that "[n]o matter the business and the size of the enterprise, APIs enable seamless operation and performance of applications and web systems."[120] Additionally, the models and procedures in the published guides by the British Standards Institute have suggestions regarding methods that can be used to implement the new regulations.[121] It can be used over and over in a

110. *Id.*

111. INFORMATION COMMISSIONER'S OFFICE, *supra* note 60, at 11.

112. DiRoma, *supra* note 30, at 61.

113. EY UK Privacy and Data Governance Channel, *supra* note 104, at 08:26.

114. *Id.* at 08:35.

115. *Id.*

116. *Id.* at 08:38.

117. *Id.*

118. API integrations can be described as "the connection between two or more applications via their APIs (application programming interfaces) that allow systems to exchange data sources. API integrations power processes throughout many sectors and layers of an organization to keep data in sync, enhance productivity and drive revenue." Thomas Jones, *What is an API Integration? (A guide for non-technical people)*, GENERATION DIGIT. (May 11, 2021), https://www.gend.co/blog/what-is-api-integration-a-guide-for-non-technical-people[https://perma.cc/TRT6-XEVY].

119. EY UK Privacy and Data Governance Channel, *supra* note 104, at 09:10.

120. Jones, *supra* note 118.

121. *Id.*

formalistic process that is a zero data and knowledge model[122] to protect children and parental information.[123]

Lastly, the chief question typically posed relates to the cost of acquiring the information and the burden of implementing the process for smaller companies that are just starting out, but under the United Kingdom's system, these resources are provided at no cost.[124] Therefore, if a common model or procedure could be employed at the same time as a revision to COPPA, it could provide resources for smaller companies, and the issue would be greatly diminished.[125] Also, it should be noted that the model published by the British Standards Institute is a globally used model.[126]

B. *What Is Currently Being Done in the Area of Children's Online Privacy Protection in the United States*

There are several groups around the country that are working to revise and update the much-outdated system that is COPPA.[127] These groups are comprised of individuals with various degrees of interest, including concerned parents advocating for change, state governments and legislatures that are working to make a difference within their own borders, as well as federal legislatures that are vying for support on bills that can generate change directly to the current law.[128] An examination of these projects is the best way to understand what local and state governments have been doing and what matters are being pushed to their legislatures to create impactful changes in child privacy rights.[129]

There are currently three bills being considered that would expand online protection for children; however, none have gained enough support to revise COPPA, and only two are worth mentioning for this Note.[130] The third bill, the Eliminating Abusive Rampant Neglect of Interactive Technologies or EARN IT Act, is directed toward the

122. Zero data and knowledge model means that this procedure can be completed without taking any data or information of the potential user or the parent of the potential user, so no information or data is obtained by running this age verification check. EY UK Privacy and Data Governance Channel, *supra* note 104, at 09:24.

123. *Id.*

124. *Id.* at 09:48.

125. *Id.*

126. *Id.* at 08:53.

127. Electronic Privacy Information Center, *Children's Privacy*, EPIC.ORG (last visited Feb. 21, 2023), https://epic.org/issues/data-protection/childrens-privacy/ [https://perma.cc/S6AL-9NUC].

128. *Id.*

129. Michael P. Canty, Carol C. Villegas, & Danielle Izzo, *Assessing 3 Bills To Expand Kids' Online Protections In 2022*, LABATON SUCHAROW (Feb. 4, 2022), https://www.labaton.com/blog/assessing-3-bills-to-expand-kids-online-protections-in-2022 [https://perma.cc/S78F-GLKK].

130. *Id.*

insulation of online service providers and directly relates to specific instances of child exploitation that are beyond the scope of this Note.[131]

The first bill worth mentioning is a proposed amendment to COPPA, which includes changing the cutoff age from thirteen to fifteen; lowering the standard for knowledge from actual knowledge to constructive knowledge; forbidding advertising that targets minors; providing a feature that will allow minors the opportunity to delete any personal information obtained by online providers; forcing obligations for online providers to label detailed disclosures in regard to the information obtained; creating a program within the FTC to regulate online marketing directed at minors.[132] The second is a new act called the Kids Internet Design and Safety, or KIDS.[133] The major components of this act include changing the age threshold for protection to sixteen, similar to the previous idea; lowering the standard to constructive knowledge; prohibiting particular interfaces or functional components that target children; limiting the scope of algorithms; increasing guidelines and prohibiting certain explicit content from reaching children.[134]

Both proposals are very forward-thinking and would help resolve several issues relative to the current system. Yet what they seem to lack are more details and resources that can be implemented to create substantive change. For instance, limiting the scope of algorithms is a great tool for keeping service providers from acquiring personal information from young users to develop marketing and advertising focused directly on the specific wants of children. However, the lack of specific guidelines provided to service providers regarding what they can and cannot do and the lack of resources to help the providers adhere to these guidelines is problematic. Essentially, the providers are largely left unregulated because the FTC and others like them fail to keep up to date on current business systems and technologies and knowledge of current issues children face online. If the FTC neglects to provide the proper resources to manage the additional requirements, the cycle of having regulations in place without proper enforcement will continue to render all the changes considered ineffective. However, with the proper resources and enforcement in place, one can ensure that service providers can continue to do business efficiently while simultaneously protecting the interest of children's privacy rights.

131. *Id.*
132. *Id.*
133. *Id.*
134. *Id.*

IV. MEETING THE NEEDS OF THE 21ST CENTURY: A FIVE-POINT PLAN FOR REVISING COPPA

As emphasized throughout this Note, the amount of time that has passed since changes to COPPA were last made is astounding. Consequently, both advocates and critics of COPPA tend to agree that some type of change is in order; however, what exactly should be done is the more challenging question. To garner the necessary support for legislative reform, it is critical to balance protecting children's interests online and enacting changes that service providers can easily implement. Therefore, one must determine the most critical issues currently endangering children's online privacy and provide only the most essential safeguards for their protection. As such, I have laid out a five-point plan of the most significant revisions that need to be implemented to create a lasting impact on children's online privacy rights.

A. *Expansion of Protections to the Ages of Thirteen to Seventeen*

First, one of the most imperative changes for COPPA to be as effective as possible is the expansion of protection to cover children from ages thirteen to seventeen. The increase in age protection was discussed in the two bills currently being debated by Congress and was also mentioned in the United Kingdom's Age-Appropriate Design Code.[135] Importantly, "[t]eenagers ages thirteen to seventeen are going online increasingly more frequently than ever before. A recent study by the Pew Research Center found that ninety-two percent of teenagers report going online daily--including twenty-four percent who say they go online almost constantly."[136] As such, those who need online privacy protection the most are, in fact, children between the ages of thirteen and seventeen.

A common argument against expanding protection to children in this upper age bracket is that they have enough life experience or knowledge to be properly protected without outside intervention. However, this is often not the case.[137] These groups include individuals beginning to transition into high school, beginning to drive, and actively and independently participating as consumers in the market for the first time. As a result, a number of these individuals step into a vulnerable position of acquiring a new form of freedom while lacking a complete

135. Canty, Villegas, & Izzo, *supra* note 129; *See also* INFORMATION COMMISSIONER'S OFFICE, *supra* note 60.

136. DiRoma, *supra* note 30, at 47.

137. Indeed, other areas of the law seem to agree with the notion that children of this age range require extra protection. First, "[t]eens are still legally defined as minors and cannot legally enter into binding contracts--including privacy policies frequently found on the Internet." *See id.* at 61. Moreover, under FERPA, a child under the age of eighteen cannot prevent her parents from accessing her school records nor can a child override a parent's veto of her school record disclosure if the record is sought by the child. Allen, *supra* note 8, at 759.

understanding of the possible ramifications of their actions. Indeed, "California's legislature concluded that children and teenagers, compared to their adult counterparts, were at greater risk online because they lack fully developed self-regulating abilities and easily succumb to online-driven peer pressure."[138]

For instance, "[h]igh social media use can lead minors to [be] inundated with numerous advertisements and products. Simply by logging into a social media site, internet users of all ages are exposed to advertisements on a wide range of services from clothing stores to restaurants to the newest indoor tanning locations."[139] Most troubling, online-directed advertisements and marketing promotions often expose children in this age bracket to products that can be sexually explicit or related to the tobacco or vaping industries.[140] These industries are mindful that starting children off at an early age can enhance the possibility of addiction and continued use of their product.[141] Their unregulated advertisements directed toward older children boost the peer pressure already prevalent in a teenager's daily life.[142] Ultimately, the vulnerabilities of older children also require online privacy protection to prevent service providers from exploiting personal information to market certain products to these children coercively.

B. *Age-Appropriate Design*

Next, some form of the United Kingdom's age-appropriate design should be implemented.[143] Specifically, initiating a different set of protections based on the ages of the children involved, as done in the United Kingdom, would be invaluable to children's online privacy.[144] To effectuate this change, online service providers should complete DPIAs, and the information collected should be turned over to the appropriate governmental agencies.[145] These agencies will then use the information to create guidelines based on a child's age. This will ensure that safeguards and standards are properly constructed based on the age of users.

The individualized protection would resolve much of the debate surrounding the issue of expanding COPPA protections to those under eighteen. The age-specific structuring of the system would expand protection while recognizing a seventeen-year-old's protection needs are

138. DiRoma, *supra* note 30, at 57.
139. *Id.*
140. *Id.* at 45
141. *Id.* at 48.
142. *Id.*
143. INFORMATION COMMISSIONER'S OFFICE, *supra* note 60, at 23.
144. *Id.* at 32.
145. *Id.* at 27–31.

uniquely different from the protections needed for a seven-year-old. Children have different capacities of understanding and behaviors at different ages. Therefore, an arrangement in place that does not allow latitude in conjunction with a child's developmental stage may impose far too much protection on some and far too little on others. As noted previously, child privacy standards should not be addressed with an all-or-nothing approach, but rather, the standards should be structured for a child's particular online use in a way that will meet their needs as they develop.[146]

An individualized approach to child privacy also helps alleviate some of the concerns commonly debated regarding the restriction of a child's free access to the use of the internet. Protections tailored toward a specific age range will not restrict a child's ability to access the internet freely because children tend to use it according to their developmental stage. Any protection, in this case, would be implemented precisely to make up for a specific lack of capacity a child may have based on his or her age.

C. *Right to Have Personal Information Deleted*

The third part of the plan is the ability for children, or the parents of children, to request that certain private information be deleted.[147] This goes hand in hand with the concept of the "right to be forgotten."[148] In essence, depending on the child's age, a child, or his parents should be able to remove personal information on the internet, which is deemed detrimental.[149]

Contrary to the notion that "[g]rowing up is synonymous with learning from one's mistakes and teenagers deserve the chance to erase their foolish mistakes in private, *without the threat of future repercussions from future onlookers*,"[150] there is no current right to remove personal

146. *Id.* at 32–33.

147. This is a highly debated topic due to the effect it could have on the First Amendment's freedom of the press. *See* Amy Gajda, *Privacy, Press, and the Right to Be Forgotten in the United States*, 93 WASH. L. REV. 201, 203 (2018). However, the narrow classification of private personal information disclosed only to online service providers for age verification, or a similar purpose, should not offend the First Amendment, as it is not information that an individual voluntarily and under no pressure from an additional source decided to post or reveal. Instead, it is personal information required to be provided for an individual to use the online provider's service that is then used for advertising and other purposes. Additionally, the personal information acquired by providers is often not of the nature that can be viewed as furthering any new product or storyline that the public has the right to access, but rather it is used by providers to increase revenue and trick young users into marketing ploys.

148. Jeffrey Rosen, *The Right to Be Forgotten*, 64 STAN. L. REV. ONLINE 88, 89 (2012); *see also* Stacey B. Steinberg, *Sharenting: Children's Privacy in the Age of Social Media*, 66 EMORY L.J. 839, 864 (2017).

149. *See* Ashley Stenning, *Gone but Not Forgotten: Recognizing the Right to Be Forgotten in the U.S. to Lessen the Impacts of Data Breaches*, 18 S.D. INTL. L.J. 129, 132 (2016).

150. DiRoma, *supra* note 30, at 65 (emphasis added).

information once it exists in the online world. A minor's online image can drastically affect her life, as that minor will eventually enter the working world or attend college, and how these individuals are portrayed on social media is a common way for employers or admissions personnel to assess an individual. In fact, "[s]chools and employers are rejecting young people for school programs, internships, college admissions, and jobs after researching applicants' online activities and posts."[151] Therefore, the ability of children to request their information be deleted is a fundamental concept and one of grave importance in children's online protection.

D. *Default Settings*

Another part of the plan is the implementation of certain default settings. Children, and even parents, are not always aware of exactly how and what information is being obtained, which can leave children vulnerable by default.[152] Thus, the required default settings should be that of high privacy protection rather than defaulting to little or no protection. The whole idea of increasing protections is because children do not have the capacity to understand what safeguards they need to protect their information.[153] Therefore, having the privacy protections default on the higher end makes sense.

As mentioned previously, a common argument against using default settings as a means of data protection is the presumption that parents help decide on and implement certain settings.[154] However, the unfortunate reality is that not all parents are involved in the process the way one might think.[155] Children often have parents who do not understand the complexities of the internet, parents who are too busy working and handling other tasks to implement the proper protections or even parents who have no knowledge that their child is using an application that is acquiring personal information.[156] Furthermore, "[p]arents find it difficult to restrict access because children are often savvier than their parents at finding and accessing Internet materials."[157] As a result,

151. *Id.* at 49.

152. Susan G. Archambault, *Student Privacy in the Digital Age*, 2021 B.Y.U. EDUC. & L.J. 1, 8 (2021).

153. *Id.*

154. Melanie L. Hersh, *Is COPPA A Cop Out? The Child Online Privacy Protection Act As Proof That Parents, Not Government, Should Be Protecting Children's Interests on the Internet*, 28 FORDHAM URB. L.J. 1831, 1835–36 (2001).

155. Brooke Auxier, Monica Anderson, Andrew Perrin & Erica Turner, *Parents' Attitudes – and Experiences – Related to Digital Technology*, PEW RSCH. CTR. (July 28, 2020), https://www.pewresearch.org/internet/2020/07/28/parents-attitudes-and-experiences-related-to-digital-technology/ [https://perma.cc/M5G3-ZYAG].

156. *Id.*

157. Hersh, *supra* note 154, at 1832.

children often use the internet with no or very little parental restriction or supervision.

Additionally, the government is the one best situated to understand what is affecting our children online and the nuances that cover this growing topic.[158] In fact, "it has been held that both parents and the government have a legal basis for protecting children."[159] Governmental agencies, specifically the FTC, are the intermediaries between online service providers, children, and parents. As such, they are the ones getting up-to-date information on violations to COPPA and what is happening in the real world in relation to this issue. Therefore, the government should be the one to apply and regulate these default settings to ensure children are protected while having the parents as an additional safeguard. In the end, ensuring adequate default settings are in place as frontline protection will result in a step in the right direction for protecting children's private information online, and together with the last part of the plan, will serve to maximize that protection.

E. *Best Interest of the Child*

The last part of the five-point plan is to require that online service providers and all players involved in the process always account for the child's best interest. At first, it may seem to be an ambiguous provision to include, but it is vital to the success of the entire plan. The essence of this provision serves its purpose whenever any ambiguity arises or when an online service provider is unclear about what action should be taken. At that point, the provider should follow the guideline that works in the child's best interest. This should always prevail, no matter the situation. All the plan components work together to increase protection and ensure that every child under eighteen has the proper safeguards; however, without constant reminders, children can be forgotten or overlooked. As mentioned previously, this is not a new idea created by the United Kingdom when crafting the Age-Appropriate Design Code; instead, it is an idea grounded in basic human rights and coined by the United Nations.[160] Online providers must recognize the importance of this protection and make it a part of their daily tasks to keep the interest of the children at the forefront of their operations.

158. *Id.* at 1859.

159. *Id.* at 1835.

160. INFORMATION COMMISSIONER'S OFFICE, *supra* note 60, at 3.

CONCLUSION

As technology and the manner in which children interact on the internet change, the law, too, must change. Technology is advancing at too great a rate for children's online privacy protection to be stuck in the late 20th century. The plan proposed in this Note incorporates only a fraction of amendments that may be implemented to protect our youth better. However, it is an essential first step to creating substantive change. Implementing a practical solution will help alleviate stress and overcome the greatest hurdle preventing the law from developing alongside technology—the economic and administrability burden imposed on online service providers. Ultimately, children today are being exposed to risky circumstances, and it is our responsibility as parents, online service providers, and even young adults who were recently in the same predicament to step up and push for lasting change that will bring children's online privacy rights into the 21st century while simultaneously protecting the innocence of our youth.

SYNTHETIC DATA AND GDPR COMPLIANCE: HOW ARTIFICIAL INTELLIGENCE MIGHT RESOLVE THE PRIVACY-UTILITY TRADEOFF

Michael Cairo[*]

Abstract

Data is in many ways the lifeblood of the digital economy. High-quality data oftentimes requires significant detail which may be at odds with the privacy concerns of the human subjects from whom data is extracted. The tension between the usefulness of a dataset and the data subject's privacy has been referred to as the "privacy-utility tradeoff." A novel application of artificial intelligence has potentially made it possible to resolve this tradeoff through the creation of "synthetic data," anonymized data generated through general adversarial neural networks from authentic raw data. Unlike pseudonymized data, synthetic data retain properties that are statistically equivalent to the underlying data gathered from data subjects. As the cost of compliance with privacy laws across the world increases, synthetic data may prove to be a viable solution to the tension between protecting individual privacy rights and the demand in the big data market.

This Note argues that large BigTech companies should incorporate synthetic data into their business models to protect users' private, personal data while retaining large profits derived their ad-driven business models. Part I provides an overview of GDPR, the patchwork of U.S. privacy laws, and recent caselaw that illustrates EU regulators' strict approach to enforcement compared to their U.S. counterparts. Part II discusses how the Privacy-Utility Tradeoff and BigTech's current business model renders compliance with data privacy regulations difficult. Part III explains how synthetic data can be used to resolve the Privacy-Utility Tradeoff.

* University of Florida Law '22. Associate attorney at Horizons Law & Consulting Group specializing in corporate and securities work for startups in the blockchain and digital asset industry. The author wishes to thank the University of Florida Levin College of Law for receipt of the Governor's Scholarship and Dean Amy Stein of the University of Florida Levin College of Law for her work on this Article and for an outstanding education in the intersection of artificial intelligence and the law.

INTRODUCTION

> "The classic saying is: 'if you're not paying for the product, then you are the product,' . . . [but t]hat's a little too simplistic. It's the gradual, slight, imperceptible change in your own behavior and perception that is the product."[1]

As the role of technology continues to expand in the daily lives of most people around the globe, data companies are having difficulty complying with a shifting regulatory landscape as new laws governing data privacy emerge.[2]

Europe's behemoth, the General Data Protection Regulation (GDPR or the Regulation),[3] is the most robust, comprehensive, and jurisdictionally far-reaching data privacy regulatory framework to date. U.S. privacy law lags behind, with no equivalent omnibus federal privacy law and only a few state laws working to fill the federal gap caused by a sector-specific patchwork of privacy laws.[4] As such, within the global regulatory landscape, GDPR stands as a significant concern for BigTech giants like Google and Facebook, whose entire business models depend upon freely collecting and processing their users' personal data to sell advertisements,[5] a common industry practice that persisted largely unencumbered for nearly two decades—until GDPR's enactment.

Data privacy laws, and GDPR in particular by way of European regulators' unrelenting enforcement regime, are forcing BigTech companies to adapt.[6] Four years into its implementation, GDPR now applies to every company that interacts with the personal data of any

---

1. THE SOCIAL DILEMMA (Exposure Labs 2020).

2. *See* Matthew Humerick, *Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence*, 34 SANTA CLARA HIGH TECH. L.J. 393, 418 (2018); Thomson Reuters, *Top Five Concerns with GDPR Compliance*, https://legal.thomsonreuters.com/en/insights/articles/top-five-concerns-gdpr-compliance (last visited Sept. 25, 2020); Elizabeth L. Feld, *United States Data Privacy Law: The Domino Effect After the GDPR*, 24 N.C. BANKING INST. 481, 486 (Mar. 2020).

3. Council Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Council Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) 1 (EU) [hereinafter GDPR].

4. *See, e.g.*, California Consumer Privacy Act, CAL. CIV. CODE § 1798.100–199.95 (West 2018); *see also* Biometric Information Privacy Act, 740 ILL. COMP. STAT. 14/15 (2008).

5. *See* Facebook, Inc., Annual Report (Form 10-K at 7) (Dec. 31, 2019) ("We generate substantially all of our revenue from selling advertising placements to marketers. Our ads enable marketers to reach people based on a variety of factors including age, gender, location, interests, and behaviors."); *see also* Alphabet, Inc., Annual Report (Form 10-K at 9) (Dec. 31, 2019) ("We generated over 83% of total revenues from the display of ads online in 2019.").

6. Bob Violino, *Data privacy rules are sweeping across the globe, and getting stricter*, CNBC (Dec. 22, 2022, 11:21 AM), https://www.cnbc.com/2022/12/22/data-privacy-rules-are-sweeping-across-the-globe-and-getting-stricter.html [https://perma.cc/FP8P-QMBL].

resident of the European Union (EU).[7] The Regulation grants every EU resident a private right of action against any entity which processes that resident's personal information in violation of the rights included in the Regulation, such as compulsory data breach and recourse notices, or the penalties for data being obtained, collected, or used without data subjects' informed consent.[8] Violations have cost data companies a whopping $2,779,699,894 from GDPR's implementation in May of 2018 through February of 2023.[9]

Compliance with GDPR and other new data privacy laws that protect personally identifiable information has been extremely costly and labor-intensive for data companies, as the current approach to achieving compliance focuses on deidentification (manually removing users' personally identifiable characteristics from large datasets) of billions of users' personal data that have been collected and maintained by these companies for nearly twenty years.[10] Aside from the direct cost of deidentifying user data, this approach comes with an additional utility cost due to the "Privacy-Utility Tradeoff": the inverse relationship between the lack of personally identifiable characteristics within collected data and the utility of the data.[11] Data that is rich with personal information enables its custodian to use it for a wide range of purposes, like personalized advertisements tailored toward user preferences. Thus, while deidentification allows for compliance with data privacy laws, it diminishes the value of collected data to BigTech companies.

The advent of artificial intelligence (AI) and its role in efficiently processing massive datasets has further complicated the ongoing privacy discussion with respect to healthcare, policing and surveillance.[12] However, AI and its role in efficiently processing massive datasets may offer an alternative to deidentification and a solution to the Privacy-

7. GDPR, *supra* note 3, at art. 3.

8. *Id.* at art. 77.

9. *GDPR Enforcement Tracker*, CMS LEGAL, https://www.enforcementtracker.com/ [https://perma.cc/9PNB-FP6A] (last visited Feb. 17, 2023). Fines are reflected in Euros and were converted to U.S. Dollars by a EUR/USD exchange rate of 1.0696 as of February 17, 2023.

10. *See* David M. Parker et al., *Privacy and Informed Consent for Research in the Age of Big Data*, 123 PENN ST. L. REV. 703, 711 (2019). *See also* Jeffrey Dobin, *The CCPA, Facebook's Potential $60 Billion Fine & How AI Improves Compliance*, MOSTLY AI (Feb. 18, 2020), https://mostly.ai/2020/02/18/the-ccpa-facebooks-potential-60-billion-fine-how-ai-improves-compliance/ [https://perma.cc/29N6-3RMW].

11. Steven M. Bellovin et al., *Privacy and Synthetic Datasets*, 22 STAN. TECH. L. REV. 1, 4 (2019).

12. *See generally* Kashmir Hill, *The Secretive Company That Might End Privacy as We Know It*, N.Y. TIMES (Jan. 18, 2020), https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html [https://perma.cc/8J68-86H4]; Parker et al., *supra* note 10.

Utility Tradeoff in the form of synthetic data. As the name suggests,[13] synthetic data is essentially "fake," AI-generated data that mimics authentic user-generated data without using personal data which could be used to identify a user.[14] Crucially, GDPR and similar U.S. privacy laws do not apply to such data that cannot be used to identify an individual user because synthetic data generally falls within the definition of "anonymous" data under GDPR Recital 26,[15] rending the GDPR inapplicable.[16]

This Note focuses on the differences between EU and U.S. privacy law as applied to synthetic data. Additionally, this Note focuses on Google and Facebook in particular because they are widely recognized BigTech companies which have drawn much of the recent ire from regulators and the public over their handling of users' personal data. However, the implications discussed herein pertain to many data companies of all sizes. Part I provides an overview of GDPR and the patchwork of U.S. privacy laws and recent caselaw that illustrates EU regulators' heavy-handed approach to enforcement compared to their U.S. counterparts. Part II discusses how the Privacy-Utility Tradeoff and BigTech's current business model renders compliance with data privacy regulations difficult. Part III explains how synthetic data can be used to resolve the privacy-utility tradeoff and proposes a new business model designed for compliance with GDPR.

I. EU & U.S. PRIVACY LAW PRIMER

The EU and U.S. have taken considerably different approaches to data privacy regulation. In the information age, those differences are starting to reignite the debate surrounding data privacy in the U.S.[17] One such

13. Webster's Dictionary defines "synthetic" as "devised, arranged, or fabricated for special situations to imitate or replace usual realities." *Synthetic*, MERRIAM-WEBSTER DICTIONARY, https://www.merriam-webster.com/dictionary/synthetic [https://perma.cc/C755-QH8L] (last visited Nov. 21, 2020).

14. Javier Tordable, *Synthetic Data Creates Real Results*, FORBES (Aug. 26, 2020, 1:10 PM), https://www.forbes.com/sites/googlecloud/2020/08/26/synthetic-data-creates-real-results/ [https://perma.cc/N67R-7QA2].

15. GDPR, *supra* note 3, at Recital 26.

16. *See infra* Part III.A.

17. *See, e.g.*, Washington Post Editorial Board, *Congress Has Another Chance at Privacy Legislation. It Can't Afford to Fail Again*, WASH. POST (May 9, 2021), https://www.washingtonpost.com/opinions/congress-has-another-chance-at-privacy-legislation-it-cant-afford-to-fail-again/2021/05/08/9409fa28-af5c-11eb-ab4c-986555a1c511_story.html [https://perma.cc/H5K9-5YK6]; Lauren Feiner, *Congress Has Failed to Pass Big Tech Legislation in 4 Years Leading Up to the Next Election*, CNBC (Oct. 31, 2020), https://www.cnbc.com/2020/10/31/congress-fails-to-pass-big-tech-legislation-ahead-of-election.html [https://perma.cc/M44Z-8QU9]. *See also* Lauren Feiner, *FTC Commissioners Agree They*

difference is how data privacy is conceptualized in the EU versus the U.S. In Europe, for example, privacy is a fundamental right,[18] while in the U.S., it is a bit more complicated.[19] Another key difference is that the EU takes a centralized, uniform approach to data privacy regulation through GDPR, while the United States has instead opted for a sector-specific patchwork of federal legislation, whereby narrow regulations are promulgated by specific federal agencies and individual states are permitted to impose supplemental privacy laws.[20] The U.S. Congress is nowhere near passing comprehensive privacy legislation at the federal level.[21]

To understand the data privacy compliance issue discussed in this Note, this Part will serve as an overview of the relevant provisions of GDPR, federal privacy laws in the United States, California's CCPA, and other emerging state privacy laws that might provide insight about what a federal privacy framework may look like.[22] This Part will begin with their jurisdictional reach, key definitions, the rights of data subjects, basic requirements for compliance, the manner in which violations are adjudicated, and the magnitude of the penalties for violations. Next, this Part will highlight the major differences in EU and U.S. privacy law and discuss how the fall of the EU-U.S. Privacy Shield[23] has given rise to substantial uncertainty for global compliance.

Should Act to Protect Consumer Privacy if Congress Doesn't, CNBC (Apr. 20, 2021), https://www.cnbc.com/2021/04/20/ftc-commissioners-agree-they-should-protect-consumer-privacy.html [https://perma.cc/L6RL-8KT5].

18. *See* Charter of Fundamental Rights of the European Union, 2012 O.J. (C 326) 2 at arts. 7–8, 10–11, https://www.refworld.org/docid/3ae6b3b70.html [https://perma.cc/5LRL-KNZA].

19. A hypothetical conversation between an EU and a U.S. citizen highlights how much more complex privacy law is in the U.S. compared to the EU even well before GDPR's passage. *See* Daniel Solove, *The Chaos of US Privacy Law*, LINKEDIN (Oct. 24, 2012), http://www.linkedin.com/today/post/article/20121024165918-2259773-the-chaos-of-us-privacy-law [https://perma.cc/N48C-Q38D].

20. *See infra* Part I.B.

21. *See, e.g.*, Maria Curi, *Outlook for Big Tech Dims as Omnibus Excludes Key* Measures, BL (Dec. 20, 2022, 9:51 AM), https://news.bloomberglaw.com/privacy-and-data-security/outlook-for-big-tech-bills-dims-as-omnibus-excludes-key-measures [https://perma.cc/4MYC-D9MN]; Alex LaCasse et al., *A look back at privacy and data protection in 2022*, IAPP (Dec. 20, 2022), https://iapp.org/news/a/a-look-back-at-privacy-and-data-protection-in-2022/# [https://perma.cc/PF2C-BNMM]; Gopal Ratnam, *Lawmakers will face familiar technology issues next Congress*, Roll Call (Dec. 13, 2022, 7:00AM), https://rollcall.com/2022/12/13/lawmakers-will-face-familiar-technology-issues-next-congress/ [https://perma.cc/FAZ9-7SWU].

22. *See* California Privacy Rights Act 2020 Cal. Legis. Serv. Prop. 24 (West) (amending CCPA, effective Jan. 1, 2023).

23. The EU-U.S. Privacy Shield previously shielded U.S. technology companies from liability for violations of EU law when collecting European residents' data until it was invalidated by European regulators. Ruth Boardman & Ariane Mole, *Schrems II: Privacy Shield Invalid, SCCS Survive. What Happens Now?*, BIRD & BIRD (July 15, 2020),

A. *EU: GDPR*

Until 2018, BigTech giants operated largely unrestricted in their data collection methods and in their use of collected data.[24] But this all changed on May 25, 2018, when GDPR went into effect.[25] Today, GDPR is largely regarded as the groundbreaking data privacy gold standard by data privacy experts.[26] In essence, EU Member States regard privacy as a fundamental right[27] with GDPR's primary aim being the protection of EU residents' personal data.[28] There are five key provisions of GDPR which are relevant to this Note, each of which are discussed, in turn, in the following "Basic Overview" section.

1. Basic Overview

The first relevant GDPR provision relates to its jurisdictional scope. GDPR grants EU regulators expansive extraterritorial jurisdiction over "controllers" and "processers" both within the EU and beyond, so long as the actions taken by these entities involve the personal data of an EU resident.[29] "Controllers" and "processors" are among those subject to liability under the Regulation. "Controller" means "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data," and "processor" means "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller."[30]

---

https://www.twobirds.com/en/insights/2020/global/schrems-ii-judgment-privacy-shield-invalid-sccs-survive-but-what-happens-now [https://perma.cc/FD5F-X8HF].

24. Stephen Zafarino, *The GDPR and the Effect on US Ad Tech*, CIO (June 28, 2018, 9:40AM), https://www.cio.com/article/3285667/the-gdpr-and-the-effect-on-us-ad-tech.html [https://perma.cc/3QP3-745K].

25. GDPR, *supra* note 3; *see also* Kimberly A. Houser & W. Gregory Voss, *GDPR: The End of Google and Facebook or A New Paradigm in Data Privacy?*, 25 RICH. J.L. & TECH. 3, 58 (2018).

26. Giovanni Burttarelli, *The EU GDPR as a Clarion Call For a New Global Digital Gold Standard*, EUROPEAN DATA PROTECTION SUPERVISOR (Apr. 1, 2016), https://edps.europa.eu/press-publications/press-news/blog/eu-gdpr-clarion-call-new-global-digital-gold-standard_en [https://perma.cc/VMZ5-WNPX].

27. GDPR, *supra* note 3, at Recital 1 ("The protection of natural persons in relation to the processing of personal data is a fundamental right.").

28. *Id.* at Recital 4 ("The processing of personal data should be designed to serve mankind. The right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality.").

29. *Id*. at art. 3. *See also* Ben Wolford, *Does the GDPR apply to companies outside of the EU?*, GDPR.EU, https://gdpr.eu/companies-outside-of-europe/ [https://perma.cc/S7PT-5VDM] (last visited Feb. 18, 2023).

30. *Id.* at arts. 4(7)–(8).

The second relevant set of provisions, Articles 3 and 4(1)–(2), also address the jurisdictional scope of GDPR.[31] Specifically, GDPR applies to any enterprise or individual who is engaged in the "processing of personal data . . . regardless of whether the processing takes place in the [European] Union or not," where "processing" means "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation [sic], structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission."[32] Further, the Regulation accounts for extraterritorial processing of personal data. Such "cross-border processing" as defined in Article 4(23)(b) encompasses the processing of personal data "which substantially affects or *is likely to* substantially affect data subjects in more than one Member State."[33] In other words, GDPR applies to any company or individual who collects or processes the personal data of *any* EU resident, regardless of where the entity doing the collecting, or the data subject, are physically located—so long as such activities are merely *likely* to substantially affect residents of more than one EU Member State.[34]

The third set of relevant provisions defines the types of data and individuals about which GDPR is concerned. The Regulation defines "personal data" as "any information relating to an identified or identifiable natural person ('data subject')," [35] and defines "identifiable natural person" as "one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."[36] These definitions are rather broad but nonetheless more instructive than those set forth by the U.S. patchwork of privacy laws discussed in Part I.B.2.a. *infra*.

The fourth set of relevant provisions concerns the rights of individuals afforded GDPR data privacy protections. All data subjects are afforded eight basic rights and are entitled to certain disclosures regarding the use of their data. The eight user rights are: (1) the right to information;[37] (2) the right of access;[38] (3) the right to rectification;[39] (4) the right to

31. *Id.* at arts. 3, 4(1)–(2).
32. *Id.* at arts. 3, 4(1).
33. *Id.* at art. 4(23)(b) (emphasis added).
34. *Id.*
35. *Id.* at art. 4(1).
36. GDPR, *supra* note 3, at art. 4(1).
37. *Id*. at arts. 13–14.
38. *Id.* at art. 15.
39. *Id.* at art. 16.

erasure;[40] (5) the right to restriction of processing;[41] (6) the right to data portability;[42] (7) the right to object;[43] and (8) the right to avoid automated decision-making.[44] Additionally, Article 34 mandates that controllers disclose data breaches "without undue delay," and requires controllers to maintain adequate technical and organizational security measures to prevent, and mitigate the severity of, data breaches.[45]

The fifth set of relevant provisions dictates when a controller is permitted to process a data subject's personal data. Article 6, the most frequently violated provision,[46] permits the processing of personal data to only six legal bases: (1) the controller has obtained the data subject's consent; (2) the data processing is necessary for the performance of a contract to which the data subject is a party; (3) the data processing is necessary for compliance with a legal obligation to which the controller is subject; (4) the data processing protects the "vital interests of the data subject;" (5) the data processing is necessary for the performance of a task carried out in the public interest; or (6) the data processing is necessary for carrying out the controller's legitimate interests, "except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject."[47] If none of these six conditions are met, GDPR prohibits the processing of *any* data subject's personal data by *any* controller. Violators may be subject to painfully high financial penalties.

2. Penalties for Violations

Financial penalties are not just reserved for controllers who violate Article 6, however. Article 77 of GDPR provides every data subject with a private right of action against any processor who touches their personal data.[48] Moreover, Article 83 grants EU regulators with enforcement authority against processors for non-compliance, a two-tiered penalty hierarchy, and provides guidelines for imposing appropriate penalties.[49] Depending on the violation and the "nature, gravity and duration of the infringement," fines can amount to 2% of a violator's global annual turnover or €10 million, whichever is higher; or 4% of global annual

40. *Id.* at art. 17.
41. *Id.* at art. 18.
42. GDPR, *supra* note 3, at art. 20.
43. *Id.* at art 21.
44. *Id.* at art. 22.
45. *Id.* at art. 34.
46. As of February 18, 2023, at least 500 fines for "insufficient legal basis for data processing" under Article 6 have been issued, totaling $489,616,113.21. *GDPR Enforcement Tracker*, *supra* note 9.
47. GDPR, *supra* note 3, at arts. 6(1)(a)–(f).
48. *Id.* at art. 77.
49. *Id.* at art. 86.

turnover or €20 million, whichever is higher.[50] For companies like Meta (previously known as Facebook) or Google, who raked in $117 billion and $257 billion in annual revenue in 2021, respectively fines can be as high as $5.1 billion and $10.2 billion, respectively, *per infringement*.[51]

GDPR's use of private rights of action as a means of enforcement illustrates a notable difference between EU and U.S. approaches to data privacy regulation. Unlike in the EU, U.S. federal law generally does not permit a private right of action in data privacy cases, with a few exceptions.[52] As discussed further below, this is just one of several key differences which exist between EU and U.S. methods of data privacy regulation.

B. *U.S.: Privacy Patchwork*

In addition to the lack of private rights of action for U.S. citizens, there are two clear differences between the EU and U.S. approaches to data privacy. First, the EU explicitly recognizes privacy as a fundamental human right, whereas the U.S. Constitution does not recognize any *explicit* right to privacy.[53] Rather, the Supreme Court has interpreted the overlap of multiple enumerated rights within the Bill of Rights as creating an implied right to privacy.[54] Secondly, as previously mentioned in this Note, the U.S. lacks a comprehensive data privacy framework at the federal level which even remotely resembles GDPR. Instead, a patchwork of several federal and state laws narrowly focuses on data privacy in specific industries.

1. U.S. Common Law on the "Right" to Privacy in the Digital Age

Supreme Court decisions from the last several decades illustrate the United States' evolving attitude toward privacy as a constitutional right. Judicial attitudes toward modern privacy stem largely from Bill of Rights jurisprudence with a particular focus on the Fourth Amendment.

In *Katz v. United States*,[55] the Court established that the standard for determining whether the Fourth Amendment precludes government

50. *Id.* at arts. 77, 83(2).

51. Facebook, Inc, *supra* note 5, at 80; Alphabet, Inc., *supra* note 5, at 50.

52. CHRIS D. LINEBAUGH, CONG. RSCH. SERV., LSB10303, ENFORCING FEDERAL PRIVACY LAW—CONSTITUTIONAL LIMITATIONS ON PRIVATE RIGHTS OF ACTION (2019); *see also* Davidson Lentz, *The Top 9 Federal Data Privacy Laws*, TN CYBERSECURITY LAW (Nov. 14, 2019), http://www.tncyberlaw.com/overview-of-federal-data-privacy-laws/ (noting that the FTC Act, COPPA, GLBA, HIPAA, and FERPA do not contain private rights of action, but FCRA, CFAA, and ECPA do.).

53. Shannon Togawa Mercer, *The Limitations of European Data Protection as a Model for Global Privacy Regulation*, 114 AJIL UNBOUND 20, 22 (2020).

54. Griswold v. Conn., 381 U.S. 479, 484 (1965) ("Various guarantees create zones of privacy.").

55. Katz v. U.S., 389 U.S. 347, 360 (1967) (Harlan, J., concurring).

search and seizure under the Fourth Amendment is whether a criminal defendant had a "reasonable expectation of privacy" of his person, or in the area or item being searched. Two years after *Katz*, the Supreme Court held that privacy is a fundamental right in *Griswold v. Connecticut*,[56] explaining that a constitutional right to privacy can be found when reading several guarantees within the Bill of Rights together.

> The foregoing cases suggest that specific guarantees in the Bill of Rights have penumbras, formed by emanations from those guarantees that help give them life and substance. Various guarantees create zones of privacy. The right of association contained in the penumbra of the First Amendment is one, as we have seen. The Third Amendment, in its prohibition against the quartering of soldiers 'in any house' in time of peace without the consent of the owner, is another facet of that privacy. The Fourth Amendment explicitly affirms the 'right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures.' The Fifth Amendment, in its Self-Incrimination Clause, enables the citizen to create a zone of privacy which government may not force him to surrender to his detriment. The Ninth Amendment provides: 'The enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people.'[57]

The Court's recognition of these penumbras and their associated implied rights was not met without controversy, but nonetheless provided a basis for the Court to recognize that Americans' right to privacy is as worthy of being protected as much as other enumerated constitutional rights. For example, in *United States v. Jones*, the U.S. Supreme Court, in considering an appeal of a Fourth Amendment challenge to law enforcement's warrantless access to a criminal defendants' location data derived from his cell phone provider, gave credence to the "mosaic theory" as a means of establishing a limited right to privacy.[58] Justice Scalia, writing for the five-justice majority in *Jones*, did not adopt the lower court's mosaic theory, but five other justices wrote or joined opinions that echoed the lower court's[59] reasoning. They reasoned that

---

56. 381 U.S. at 485 ("The present case, then, concerns a relationship lying within the zone of privacy created by several fundamental constitutional guarantees.").

57. *Id.* at 484 (internal citation omitted).

58. United States v. Jones, 565 U.S. 400 (2012).

59. United States v. Maynard, 615 F.3d 544, 562 (D.C. Cir. 2010), *aff'd in part sub nom.* United States v. Jones, 565 U.S. 400 (2012) ("A person who knows all of another's travels can deduce whether he is a weekly church goer, a heavy drinker, a regular at the gym, an unfaithful husband, an outpatient receiving medical treatment, an associate of particular individuals or political groups—and not just one such fact about a person, but all such facts.").

while an individualized search (or, for our purposes here, a singular instance of data collection) may not necessarily intrude on an individual's right to privacy, an aggregated collection of seemingly innocuous individual pieces of collected data taken together can, like a mosaic,[60] create a vivid and intrusively detailed picture of how a person lives their life in such a way that effectively strips them of privacy.[61] Thus, while the *Jones* Court did not strike down this particular acquisition of an individual's data, they opened the door to the idea that such permissionless access, in the aggregate, may amount to certain unconstitutional privacy violations.

The Court took the *Jones* ruling a step further in the subsequent *Riley v. California*[62] case. The Court wrestled with the privacy-utility tradeoff in the criminal context, acknowledging that while personal data stored in a defendant's cell phone can be useful to law enforcement, data "cannot itself be used as a weapon to harm an arresting officer or to effectuate the arrestee's escape."[63] The Court held that warrantless searches of the data stored on a criminal suspect's cell phone violates the Fourth Amendment,[64] and the exigency exception[65] to the Fourth Amendment's warrant requirement does not *always* apply. Notably, despite this decision's seeming resemblance to *Katz*, the Court reached this decision under a theory of property law rather than along the privacy-focused lines of *Katz*,[66] evincing the Court's reluctance to address the issue of "privacy" absent federal legislation.

Another victory for data privacy rights would emerge through the Court's eventual challenging of the third-party doctrine. Historically, data privacy in the United States has been limited by the broad application of the third-party doctrine, which holds that "a person has no legitimate expectation of privacy in information he voluntarily turns over to third parties,"[67] "even if the information is revealed on the assumption that it

60. Rob Silvers, *Marilyn* (illustration), *in* Robert S. Silvers, *Photomosaics: Putting Pictures in Their Place*, MASS. INST. OF TECH. 84 (1996), http://dspace.mit.edu/bitstream/1721.1/29135/2/38491951-MIT.pdf [https://perma.cc/QX7A-SMRL].

61. Orin S. Kerr, *The Mosaic Theory of the Fourth Amendment*, 111 MICH. L. REV. 311, 312–15 (Dec. 2012).

62. Riley v. California, 573 U.S. 373 (2014).

63. *Id*. at 387.

64. Katz v. United States, 389 U.S. 347, 360–61 (Harlan, J. concurring) (stating that "a person has a constitutionally protected reasonable expectation of privacy . . . [and] that electronic as well as physical intrusion into a place that is in this sense private may constitute a violation of the Fourth Amendment[.]").

65. Riley, 573 U.S. at 390 ("Moreover, in situations in which an arrest might trigger a remote-wipe attempt or an officer discovers an unlocked phone, it is not clear that the ability to conduct a warrantless search would make much of a difference.").

66. Katz, 389 U.S. at 360–61.

67. Smith v. Maryland, 442 U.S. 735, 743–44 (1979).

will be used only for a limited purpose."[68] In her concurrence in *Jones*, Justice Sotomayor seemed to disfavor the third-party doctrine, stating that "it may be necessary to reconsider the premise that an individual has no reasonable expectation of privacy in information voluntarily disclosed to third parties."[69]

In 2018, the Court attacked the doctrine head-on in *Carpenter v. United States*, holding "[i]n light of the deeply revealing nature of [smartphone location information], its depth, breadth, and comprehensive reach, and the inescapable and automatic nature of its collection, the fact that such information is gathered by a third party does not make it any less deserving of Fourth Amendment protection."[70] In his *Carpenter* opinion, Chief Justice Roberts quoted Justice Brandeis' dissent in *Olmstead v. United States*,[71] writing that "the Court is obligated—as '[s]ublter and more far-reaching means of invading privacy have become available to the Government(sic)'—to ensure that the 'progress of science' does not erode Fourth Amendment protections."[72] As prophetic as Justice Brandeis' concerns in *Olmstead* may seem when read nearly one hundred years later, one explanation for the U.S.'s lack of a federal data privacy framework could be that the American tradition of fostering economic growth through free enterprise tips the scale to favor utility over privacy, discussed in Part II *infra*.

2. Federal Privacy Laws

Fourth Amendment jurisprudence contemplates the idea of privacy as a right,[73] and has grappled with the Privacy-Utility Tradeoff most clearly in the law enforcement context, but adequate privacy legislation suitable for the digital age has not followed.

The Federal Trade Commission (FTC) is the primary agency responsible for enforcing federal consumer protection laws, including data privacy laws. The FTC has the authority to take action against companies that engage in deceptive or unfair practices related to data privacy, guided primarily by the Commission's own guidelines.[74] The Office of Civil Rights under the Department of Health and Human

68. United States v. Miller, 425 U.S. 435, 443 (1976).

69. United States v. Jones, 565 U.S. 400, 417 (2012) (Sotomayor, J., concurring).

70. Carpenter v. United States, 138 S. Ct. 2206, 2223 (2018).

71. Olmstead v. United States, 277 U.S. 438, 473–74 (1928) (Brandeis, J., dissenting).

72. Carpenter, 138 S. Ct. at 2223 (quoting Olmstead, 277 U.S. at 473–74 (Brandeis, J., dissenting)).

73. *See* Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 213 (1890); *see also* Jed Rubenfeld, *The Right of Privacy*, 102 HARV. L. REV. 737, 744–47 (1989).

74. *See* FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE (2010), http://www.ftc.gov/os/2010/12/101201privacyreport.pdf [https://perma.cc/6JT3-CCY3] [hereinafter 2010 FTC Report].

Services enforces data privacy and security matters pertaining to healthcare and patient data under its authority granted in Health Insurance Portability and Accountability Act of 1996 (HIPAA).[75] The Federal Communications Commission (FCC) regulates data privacy related to broadband providers, and the Department of Education (DOE) enforces data privacy regulations under the Family Educational Rights and Privacy Act (FERPA) which applies to educational institutions.[76] Other federal statutes grant individuals a private right of action to sue for limited damages, with some jurisdictions unclear as to whether explicit privacy violations related to one industry-specific categorization of data apply to others.[77]

Professor Steven M. Bellovin and his coauthors summarized the central problem with this federal patchwork approach nicely:

> Protected sectors range from health (HIPAA) to finance (FCRA), and often hinge the statutory shield on the definition of "personally identifiable information" (PII). Put simply, if a fact (i.e., a datum in the database) contains PII, then it is protected and cannot be shared; if the fact does not contain PII, then it is not protected and may be shared freely. The problem comes from delineating PII from non-PII.[78]

To complicate matters further, and central to this Note's thesis, Professor Paul Ohm challenged the idea that one can reliably separate personally identifiable information from the surrounding benign information,[79] as discussed *infra* in Part II. Thus, GDPR's centralized authority, with universal definitions, rights, and obligations, appears to have several advantages over the U.S.'s patchwork structure, which the U.S. could seek to learn from to improve its own data privacy regulation efforts. Nonetheless, CCPA is the closest that American law has to GDPR and may be a step toward regulatory clarity. Common threads between CCPA and other emerging state privacy laws are beginning to appear and may provide insight into what a single, centralized federal privacy statute will look like.

---

75. *See* 42 U.S.C. § 1320.

76. Family Educational Rights and Privacy Act, 20 U.S.C. § 1232g.

77. *See* Omer Tene, *Privacy Law's Midlife Crisis: A Critical Assessment of the Second Wave of Global Privacy Laws*, 74 OHIO ST. L.J. 1217, 1225 (2013).

78. Bellovin et al., *supra* note 11.

79. Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1701 (2020).

a. Consumer Data Privacy: FTC Act

The Federal Trade Commission is the chief data regulator in the U.S. by way of its charge to protect consumers.[80] The FTC's current role in U.S. consumer data privacy law essentially boils down to using its enforcement authority (granted in § 5 of the FTC Act) to take action against companies who fail to, or who deceptively, obtain users' consent for how their data is used.[81] The FTC also enforces industry-specific federal data privacy statutes, including Gramm-Leach-Bliley Act (GLB Act), the Children's Online Privacy Protection Act (COPPA), the CAN-SPAM Act, and the Telemarketing and Consumer Fraud and Abuse Prevention Act ("Do Not Call Rule").[82] Neither the FTC Act nor the others named above contain a private right of action, so the Commission itself, rather than private litigants, is tasked with investigating suspected violations and deciding whether to bring a lawsuit.[83]

In the U.S., each sector-specific privacy law has its own definition of personally identifiable information. Commentators have rightfully pointed out that such an approach essentially renders the term impossible to understand.[84] Generally, definitions of "personal information" or "personally identifiable information,"—the U.S. corollaries to GDPR's "personal data," under federal privacy laws—operate on the assumption that "personal information" worthy of statutory protection is simply information that can be used to identify a person.[85] For example, under

---

80. *See* 15 U.S.C. § 41.

81. *See* 15 U.S.C. § 45 (prohibiting the use of deceptive or unfair trade practices, which has been interpreted broadly to cover data privacy and antitrust).

82. *See* 15 U.S.C. §§ 6801–6809 (consumer financial data); 15 U.S.C. §§ 6501–6506 (children's online privacy); 15 U.S.C. §§ 7701–7713 (unsolicited electronic messages); 15 U.S.C. §§ 6101–6108 (telemarketing calls).

83. The FTC's latest data privacy report from 2010 provides a good history of the FTC's role in data privacy governance. The 2010 Report proposed a legal framework for Congress to impose that features privacy by design, simplified consumer choice, and transparency as core components, though no such law has been passed to date. *See* FED. TRADE COMM'N, 2010 ANNUAL REPORT 38, 39–79 (2010).

The proposed framework in the 2010 FTC Report is based on the FTC's "Fair Information Practice Principles" (FIPPs), published in 2000. The four key components are: notice, choice, access and security. These principles were modeled after contemporaneously emerging European privacy legislation and appear to be similar to a more primitive form of GDPR's core principles, however FIPPs are merely a nonbinding set of guidelines, as Congress failed to codify them into law. *See* FED. TRADE COMM'N, PRIVACY ONLINE: FAIR INFORMATION PRACTICES IN THE ELECTRONIC MARKETPLACE i (2000) [hereinafter FIPPs Report], https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-fair-information-practices-electronic-marketplace-federal-trade-commission-report/privacy2000.pdf [https://perma.cc/5KKJ-KLNX].

84. *See* Bellovin et al., *supra* note 11, at 4; *see also* Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814, 1829–35 (2011); Ohm, *supra* note 79, at 1701; Tene, *supra* note 77, at 1217.

85. Schwartz & Solove, *supra* note 84, at 1819.

the Video Privacy Protection Act (VPPA), personally identifiable information is defined as "information which identifies a person."[86] As Professors Schwartz and Solove point out, this definition's utility is in its openness and flexibility to respond to changing circumstances, but the definition's flaw is that it "simply states that PII is PII."[87] In contrast, the GLB Act, a financial privacy statute, covers "nonpublic personal information," and defines it as "personally identifiable financial information (i) provided by a consumer to a financial institution, (ii) resulting from a transaction or service performed for the consumer, or (iii) otherwise obtained by the financial institution."[88] These are very different definitions which only add to the confusion that is privacy law in the U.S.

When reviewing the definition of "personal information," it is also important to understand what is and what is *not* protected under these federal statutes. In other words, it is important to know when PII becomes non-PII. Most federal statutes reflect the assumption that when certain information that can be directly linked to a person (*e.g.*, full name, social security number, bank account number, IP address, etc.) is removed from a dataset, then it is considered "deidentified" or "anonymized" and the entity in possession is no longer subject to the same privacy and security requirements.[89] The GLB Act, for example, specifically excludes publicly available information and any consumer list attained without using personally identifiable financial information.[90] The FTC's final rule under the GBL Act excludes deidentified data from the statute, classifying it as "[i]nformation that does not identify a consumer, such as aggregate information or blind data that does not contain personal identifiers such as account numbers, names, or addresses."[91] Conversely, HIPAA is much more specific and enumerates eighteen particular identifiers that constitute "protected health information" ("PHI"),[92] the healthcare corollary to GDPR's "personal data." The eighteen identifiers are: names; postal address information other than town or city, state and zip code; telephone numbers; fax numbers; email addresses; social security numbers; medical record numbers; health insurance numbers; account numbers; certificate/license numbers; vehicle identifiers and

86. Video Privacy Protection Act of 1988, 18 U.S.C. § 2710(a)(3) (defining personally identifiable information as "information which identifies a person.").

87. Schwartz & Solove, *supra* note 84, at 1829.

88. Gramm-Leach-Bliley Act of 1999, 15 U.S.C. § 6809(4)(A).

89. *See* Schwartz & Solove, *supra* note 84, at 1828. *See also infra* Part II.A (discussing why de-identification and anonymization are misnomers and how current technology has rendered de-identification as an insufficient method of privacy protection).

90. 15 U.S.C. § 6809.

91. 16 C.F.R. § 313.3(o)(2)(ii)(B) (2001); *see also* Benjamin Charkow, *The Control Over the De-Identification of Data*, 21 CARDOZO ARTS & ENT. L.J. 195, 198 (2003).

92. 45 C.F.R. § 164.514(e) (2013).

serial numbers including license plate numbers; URLs; IP addresses, biometric identifiers including finger and voice prints; and full face photographic images and any comparable images.[93] The statute also provides permissible methods for deidentification that removes PHI from HIPAA's scope.

b. Health Data: HIPAA

HIPAA is one of the only data privacy laws in place in the U.S. that resembles GDPR at the federal level due to its robust and comprehensive nature. HIPAA applies only to "covered entities," which include healthcare providers, insurers and clearinghouses, and to "business associates" that receive data from covered entities.[94] HIPAA's purpose is to protect the privacy and security of patients' sensitive healthcare information, denoted as "protected health information" or "PHI."[95] PHI is defined as:

> any individually identifiable health information that is transmitted or maintained in any form or medium; is held by a covered entity or its business associate; identifies the individual or offers a reasonable basis for identification; is created or received by a covered entity or an employer; and relates to a past, present or future physical or mental condition, provision of health care, or payment for healthcare to that individual.[96]

To address the obvious privacy-utility tradeoff inherent in dealing with private personal health information, Congress passed the Health Information Technology for Economic and Clinical Health Act of 2009 (HITECH)[97] which updated HIPAA to impose strict compliance standards and hefty monetary penalties for unauthorized disclosures of PHI.[98] Detailed, intimate health data is an extremely valuable resource to a pharmaceutical company, for example, and HITECH's regulatory teeth are an attempt to quell abuse.

---

93. *Id.*

94. *Id.* § 160.103; *see also* Centers for Medicare & Medicaid Services, *Are You a Covered Entity?*, https://www.cms.gov/Regulations-and-Guidance/Administrative-Simplification/HIPAA-ACA/AreYouaCoveredEntity.html [https://perma.cc/D6SS-NBB9] (May 26, 2022, 10:37 AM).

95. 45 C.F.R. § 160.103 (2000).

96. *Id.*

97. Health Information Technology for Economic and Clinical Health Act ("HITECH Act"), Pub. L. No. 111-5, 123 Stat. 226 (2009) (*codified at* 42 U.S.C. §§ 300jj, 17901).

98. 45 C.F.R. § 160.404(b)(2)(i)(A), (B) (2009).

Like GDPR, the Privacy Rule[99] and the Security Rule[100] under HITECH and HIPAA set forth minimum necessary privacy and security standards regarding PHI, to which covered entities and business associates must adhere to remain in compliance. Notably, HIPAA is the most robust codification of the four principles set forth in the FTC's Fair Information Privacy Practices: notice, choice, access, and security.[101] The Office of Civil Rights within the Department of Health and Human Services is the primary enforcer of the Privacy Rule, and it can levy civil monetary penalties of up to $1.75 million per calendar year per type of violation.[102] For example, Anthem, Inc., a business associate providing administrative services to a health insurer, paid $16 million—the highest settlement for a HIPAA violation to date—when a series of cyberattacks targeting Anthem exposed the electronic PHI (ePHI) of over 78 million individuals.[103]

While HIPAA is a good start to protecting privacy in a universally sensitive area of an individual's life, its dependency on deidentification as the ultimate line of defense is no longer sufficient due to risks of reidentification.[104]

3. California Consumer Privacy Act (CCPA)

California's CCPA is the first American consumer data privacy statute that resembles GDPR, due to CCPA's comprehensive and centralized nature.[105] CCPA grants California residents specific rights, including the right to notice, the requirement of user consent, the right to erasure, the right to opt out from the sale of personal information, and the right to be free from discrimination if a consumer chooses to opt out of a company's data collection practices.[106] However, unlike GDPR, CCPA does not contain a right to correction, and the penalties imposed by CCPA are much lower than those imposed by GDPR.[107] Broadly, the statute applies

99. 45 C.F.R. §§ 160.101–.552, 164.102–.106, 164.500–.534 (2013).

100. 45 C.F.R. §§ 160.101–.552, 164.102–.106, 164.302–318 (2013).

101. *See* FIPPs Report, *supra* note 83, at 13.

102. 45 C.F.R. § 160.404(b)(2)(i)(A)–(B) (2009). For more information about the enforcement process, *see Enforcement Process*, DEP'T OF HEALTH & HUM. SERV. (Sept. 17, 2021), https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/enforcement-process/index.html [https://perma.cc/2KH7-ZPFD].

103. *Anthem Resolution Agreement*, DEP'T OF HELATH & HUM. SERV. (Oct. 15, 2018), https://www.hhs.gov/sites/default/files/anthem-ra-cap.pdf.

104. Ohm, *supra* note 79, at 1740.

105. California Consumer Privacy Act, CAL. CIV. CODE § 1798.100–199.95 (West 2018).

106. *Id*. § 1798.110.

107. *Id.* § 1798.155(a) ("Any business, service provider, contractor, or other person that violates this title shall be liable for an administrative fine of not more than two thousand five hundred dollars ($2,500) for each violation or seven thousand five hundred dollars ($7,500) for each intentional violation or violations . . . .).

to for-profit businesses that collect the personal information of any California residents, and which: earn $25 million or more in global annual revenue, collect personal information from 50,000 or more consumers, or derive 50% of their revenue from selling data.[108] Like GDPR, CCPA applies to all businesses who collect California residents' data regardless of where in the world the business is located.[109]

CCPA defines "personal information" as "information that identifies, relates to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household."[110] Personal information also includes direct identifiers; commercial information; biometric information; internet activity; geolocation data; audio, electronic, visual, thermal, olfactory or similar information; employment information, education information; and psychographic information.[111]

CCPA enforcement falls primarily under the authority of the California Attorney General, who investigates and takes action against companies who violate any portion of the statute.[112] Individuals also have a private right of action, but only for data breaches in which the individual's unencrypted personal information is disclosed "as a result of a business's failure to implement and maintain reasonable security procedures and practices appropriate to the nature of the information."[113] Statutory damages range from $100 to $750 per consumer, per "incident," or breach.[114]

Many commentators regard CCPA as a step in the right direction for U.S. privacy laws, as it simplifies the definition of "personal information" such that it is both broad enough to keep up with new data collection practices, but specific enough to be administrable.[115] CCPA also places the burden on tech companies to protect user information, rather than permitting them to escape liability through disclaimers and limitations of liability in their terms of service and privacy policies.[116]

C. *Schrems II & Fall of the EU/U.S. Privacy Shield*

While EU law has long offered robust data privacy rights and enforcement capabilities (even prior to GDPR's enactment), similarly

108. *Id.* § 1798.140(c).
109. *Id.* § 1798.80(a).
110. *Id.* § 1798.140(o)(1).
111. *Id.*
112. *Id.* § 1798.155.
113. *Id.* § 1798.150(a)(1).
114. *Id.* § 1798.150(a)(1)(A).
115. *See, e.g.*, Lauren Davis, *The Impact of the California Consumer Privacy Act on Financial Institutions, Across the Nation*, 24 N.C. BANKING INST. 499, 499–501 (2020).
116. California Consumer Privacy Act, Cal. Civ. Code § 1798.100 (West 2018).

robust protections have remained noticeably absent from U.S. law. This gap between U.S. and EU privacy laws has repeatedly complicated bilateral trade as it pertains to data transfers.[117] Before GDPR went into effect in 2018, EU residents had no way to directly control what happened to their data after it was transferred to different countries. Accordingly, to ensure that EU residents enjoyed the "adequate levels" of data protection afforded to them under Articles 25 and 26 of the EU Data Protection Directive (95/46/EC), corrective measures were taken in 2000 and again in 2016: the Safe Harbor Agreement[118] and the EU-U.S. Privacy Shield, respectively.[119] Ultimately, however, neither measure would prove to be sufficient, as both were deemed invalid under EU law[120] thanks to the efforts of a bold, young Austrian lawyer and data privacy activist named Max Schrems.[121] Thus, U.S. tech companies are now back in regulatory limbo with EU data privacy law.

The Safe Harbor Agreement aimed to ensure that U.S. companies provided EU residents' personal data the same levels of protection it would otherwise receive in the EU.[122] Under the Safe Harbor Agreement, American companies subject to the jurisdiction of the FTC and the Department of Transportation (DOT) were eligible to self-certify their inclusion into the Safe Harbor program.[123] Relying on the prohibition against deceptive or unfair trade practices under § 5 of the Federal Trade Commission Act, the FTC would take enforcement actions against any

---

117. *See* Ruth Boardman & Ariane Mole, *Schrems II: Privacy Shield Invalid, SCCS Survive. What Happens Now?*, BIRD & BIRD (July 15, 2020), https://www.twobirds.com/en/insights/2020/global/schrems-ii-judgment-privacy-shield-invalid-sccs-survive-but-what-happens-now [https://perma.cc/TY3B-U4KV]; Davide Szép, *America's Tech Giants: It's Back to the Drawing Board on European Data*, 92 N.Y. ST. BAR ASS'N. J. 45 (Nov. 2020).

118. Commission Decision 2000/520/EC, 2000 O.J. (L 215) [hereinafter Decision 2000/520/EC].

119. European Commission Press Release IP/16/433, Restoring Trust in Transatlantic Data Flows Through Strong Safeguards: European Commission Presents EU-U.S. Privacy Shield (Feb. 29, 2016), http://europa.eu/rapid/press-release_IP-16-433_en.htm [https://perma.cc/T8TK-WTX5].

120. *Data Protection Commissioner v. Facebook Ireland Ltd. Court of Justice of the European Union Invalidates the EU-U.S. Privacy Shield*., 134 HARV. L. REV. 1567, 1569 n.28 (2021).

121. Anne Beade, *Max Schrems, Reluctant Austrian David to Internet Goliaths*, TECH XPLORE (Apr. 21, 2021), https://techxplore.com/news/2021-04-max-schrems-reluctant-austrian-david.html [https://perma.cc/ZS6K-3SVG].

122. *See* Emily Linn, *A Look into the Data Privacy Crystal Ball: A Survey Of Possible Outcomes for the EU-U.S. Privacy Shield Agreement*, 50 VAND. J. TRANSNAT'L L. 1311, 1322–23 (2017); *see also* Sherri J. Deckelboim, *Consumer Privacy on an International Scale: Conflicting Viewpoints Underlying the EU-U.S. Privacy Shield Framework and How the Framework Will Impact Privacy Advocates, National Security and Businesses*, 48 GEO. J. INT'L L. 263, 279–81 (2016).

123. *Id.*

such companies that failed to comply with applicable EU privacy laws.[124] Although the European Commission deemed the substance of the Safe Harbor Agreement to be sufficient to comply with Articles 25 and 26,[125] the decision was not met without public dissent.

In the EU case, *Schrems v. Data Protection Comm'r*, Max Schrems sued Facebook for transferring his personal data from Ireland to the United States.[126] He argued that the Safe Harbor Agreement in its entirety was incompatible with EU privacy law. Specifically, Schrems argued that Edward Snowden's revelation of the U.S.'s domestic surveillance program (PRISM)[127] evidenced the U.S. government's unfettered access to EU data subjects' personal data without requiring a court order and without providing any means of redress.[128] The Court of Justice of the European Union (CJEU), hearing the case on appeal in 2015, agreed with Schrems and held that the Safe Harbor Agreement was invalid because U.S. law fails to ensure an adequate level of protection under Directive 95/46, stating:

> [o]nce the personal data has been transferred to the United States, it is capable of being accessed by the National Security Agency (NSA) and other federal agencies, such as the Federal Bureau of Investigation (FBI), in the course of the indiscriminate surveillance and interception carried out by them on a large scale.[129]

In response to the invalidation of the Safe Harbor Agreement after *Schrems I*, the U.S. Department of Commerce and the European Commission set forth the EU-US Privacy Shield Framework[130] in 2016—the same year GDPR was passed—to provide an alternative legal means to transfer data from the EU to the U.S. While BigTech titans were quick to praise the new Shield,[131] critics like Schrems himself were quick to

---

124. Linn, *supra* note 122, at 1322–23; 15 U.S.C. §§ 41–58.

125. Commission Decision 25/08/2000, 2000 O.J. (L 215) 7.

126. Case C-362/14, Schrems v. Data Prot. Comm'r, ECLI:EU:C:2015:650, ¶¶ 1–2 (Oct. 6, 2015) [hereinafter Schrems I].

127. For further discussion, *see* GLENN GREENWALD, NO PLACE TO HIDE: EDWARD SNOWDEN, THE NSA AND THE U.S. SURVEILLANCE STATE (Picador 2014).

128. Schrems I, ECLI:EU:C:2015:650 at ¶¶ 26–30.

129. *Id.* at ¶ 31. For a cogent synopsis of the Schrems I holding, *see* Linn, *supra* note 122, at 1320–25.

130. *See* European Commission, Directorate-General For Justice And Consumers, *Guide To The Eu-U.S. Privacy Shield* (2016) [https://perma.cc/N2NC-KXWR] (archived Sept. 21, 2017); *Privacy Shield Framework: Overview*, INT'L TRADE ADMIN., https://www.privacyshield.gov/article?id=OVERVIEW [https://perma.cc/Q83N-6VS6] (last visited Apr. 30, 2021).

131. *See* James Titcomb, *Facebook Signs Up to Privacy Shield Data Treaty*, THE TELEGRAPH (Oct. 16, 2016, 8:15 PM), http://www.telegraph.co.uk/technology/2016/10/15/facebook-signs-up-to-privacy-shield-data-treaty/ [https://perma.cc/DSW2-UXQL]; John Frank, *EU-U.S. Privacy*

point out that the Shield would not resolve the privacy concerns, namely the NSA's surveillance program, that led to the fall of the Safe Harbor program.[132] The Shield was substantively the same as the Safe Harbor Agreement, save for a few additions. Namely, the Shield incorporated the latest requirements and rights in the newly passed GDPR, included a U.S. Ombudsperson responsible for providing guidance on redress to EU residents, and was accompanied by a handful of letters from U.S. intelligence officials assuring EU residents that they can sue the NSA if they became subject to unlawful surveillance[133]—that is, *if* they could establish standing in court.[134]

However, the Shield would meet the same fate as the Safe Harbor Agreement in 2020, when Max Schrems struck again. In *Data Protection Commissioner v. Facebook Ireland Ltd*. (*Schrems II*), the CJEU invalidated the EU-US Privacy Shield, finding that the Privacy Shield was incompatible with GDPR and the European Union's Charter of Fundamental Human Rights.[135] GDPR's Articles 46(1) and 46(2)(c) require that EU data subjects whose personal data is transferred to another country are "afforded a level of protection essentially equivalent to that guaranteed within the European Union," with such protection including

---

Shield: Progress for Privacy Rights, MICROSOFT EU POL'Y BLOG (July 11, 2016), https://blogs.microsoft.com/eupolicy/2016/07/11/eu-u-s-privacy-shield-progress-for-privacy-rights/ [https://perma.cc/3ZVA-5LUT].

132. "[T]he replacement [for the Safe Harbor program] that the Commission has proposed right now called 'privacy shield' is basically safe harbor once again." European Parliament, *'Privacy Shield: Safe Harbour with teeny tiny changes'- Max Schrems*, YOUTUBE (Mar. 18, 2016), https://www.youtube.com/watch?v=EdCmpmL1UJk [https://perma.cc/DM6G-7YVG]; *see also* Amar Toor, *EU-US Privacy Shield Agreement Goes Into Effect: Tech Companies Welcome New Data Transfer Agreement, But Activists Say it Doesn't Do Enough to Protect Privacy*, THE VERGE (July 12, 2016), http://www.theverge.com/2016/7/12/12158214/eu-us-privacy-shield-data-transfer-privacy [https://perma.cc/9G68-NVAK].

133. Allison Callahan-Slaughter, Comment, *Lipstick on a Pig: The Future of Transnational Data Flow Between the EU and the United States*, 25 TUL. J. INT'L & COMP. L. 239, 253–54 (2016); *see also* Linn, *supra* note 122, at 1333; *see also* Letter from Robert S. Litt, Gen. Couns. of the Off. of the Dir. of Nat'l Intel., to Justin S. Antonipillai, Couns., U.S. Dep't of Com. & Ted Dean, Deputy Assistant Sec'y, Int'l Trade Admin. (Feb. 22, 2016), https://www.privacyshield.gov/servlet/servlet.FileDownload?file=015t00000004q1F [https://perma.cc/EK9D-JJQX].

134. Establishing Article III standing in challenges to U.S. government surveillance is a notoriously difficult hurdle to meet. *See* Margaret B. Kwoka, *The Procedural Exceptionalism of National Security Secrecy*, 97 B.U. L. REV. 103, 121–24 (2017); Christopher Slobogin, *Standing and Covert Surveillance*, 42 PEPP. L. REV. 517, 532–33 (2015) (arguing that unconstitutional surveillance programs might be allowed to continue without affirmative action from the legislature or the executive branch to discontinue them); *see also* Steven Graziano, *An Unconstitutional Work of Art: Discussing Where the Federal Government's Discrete Intrusions Into One's Privacy Become an Unconstitutional Search Through Mosaic Theory*, 17 MINN. J.L. SCI. & TECH. 977, 992 (2016).

135. Case C-311/18, Data Prot. Comm'r v. Facebook Ireland Ltd., ECLI:EU:C:2020:559, 163–64 (July 16, 2020).

"appropriate safeguards, enforceable rights and effective legal remedies."[136] The CJEU determined that the U.S. surveillance program did not have "appropriate safeguards," pointing to section 702 of the Foreign Intelligence Surveillance Act[137] and Executive Order 12,333.[138] The CJEU additionally found that U.S. law does not offer "enforceable rights and effective legal remedies" for EU residents whose data is transferred to the United States, pointing to Presidential Policy Directive 28.[139]

In the aftermath of these decisions, companies dealing with EU residents' personal data can still perform cross-border data transfers through standard contractual clauses (SCCs) in their terms of service.[140] However, they are now responsible for ensuring that the receiving country has laws in place that meet the requirements of GDPR, unlike the U.S.[141] According to Stockholm-based business and tech law firm, Sharp Cookie Advisors:

> The recipient is obliged to inform the exporter of any impediments to its compliance to the SCC's [sic]. If the existence of local surveillance laws . . . would impede the alignment with the GDPR, then the exporter (read your customers) must stop the transfer and end the contract. If the data exporter fails its obligations under the SCC, the lead supervisory authority must intervene and may prohibit the transfer.[142]

This puts all transnational companies dealing with personal information in quite a bind, as they inherently operate globally by way of

---

136. *Data Protection Commissioner v. Facebook Ireland Ltd. Court of Justice of the European Union Invalidates the EU-U.S. Privacy Shield.*, 134 HARV. L. REV. 1567, 1569 n.28 (2021) (citations omitted) ("The CJEU noted that 'the assessment of the level of protection afforded in the context of such a transfer must,' inter alia, consider the relevant laws of a third country 'as regards any access by the public authorities of that third country to the personal data transferred.' The court found that the GDPR's protections apply to the commercial transfer of data to a third country, regardless of the likelihood that that data will 'be processed by the authorities of [that] third country . . . for the purposes of public security, defence and State security.'").

137. 50 U.S.C. § 1881a.

138. Exec. Order No. 12,333, 3 C.F.R. § 200 (1982), *reprinted as amended in* 50 U.S.C. § 3001.

139. Press Release, Off. of the Press Sec'y, Presidential Policy Directive -- Signals Intelligence Activities (Jan. 17, 2014), http://www.whitehouse.gov/the-press-office/2014/01/17/presidential-policy-directive-signals-intelligence-activities [https://perma.cc/T6U3-WVYK].

140. Leah Shepherd, *EU Adopts New Standard Contractual Clauses for Data Transfers*, SHRM (July 28, 2021), https://www.shrm.org/resourcesandtools/hr-topics/global-hr/pages/eu-standard-contractual-clauses-data-transfers.aspx [https://perma.cc/M336-VGBH].

141. *Id.*

142. Sharp Cookie Advisors, *Schrems II a Summary – All You Need to Know*, GDPR SUMMARY (Nov. 23, 2020), https://www.gdprsummary.com/schrems-ii/ [https://perma.cc/MK9F-LNRS].

being internet-based companies. Companies must now follow convoluted processes for compliant EU-U.S. data transfers.[143]

II. COMPLIANCE CHALLENGES

The vast differences in how data privacy is implemented in the EU and the U.S. have made compliance difficult for data companies operating globally.[144] Additionally, the tradeoff between protecting user privacy and the usefulness of the retained data lies at the core of privacy laws like GDPR, HIPAA and CCPA. BigTech is particularly resistant to compliance with GDPR, as many data companies are accustomed to the U.S.'s lack of a robust omnibus privacy framework (though CCPA is beginning to change that). Enforcement decisions in the U.S. and the EU often highlight the fact that the compliance difficulties these tech companies face are caused by the overarching business model adopted by BigTech. This business model is simply not conducive to user privacy—as it is instead entirely predicated on maximizing data utility.[145] As per the privacy-utility tradeoff, efforts to improve data privacy are often at odds with the goal of prioritizing data utility.[146]

A. *Privacy-Utility Tradeoff*

Conceptually, the term "privacy-utility tradeoff" is used to refer to the incompatibility of the usefulness of data collected from individual users and the privacy which those individual users enjoy.[147] The tradeoff can be succinctly summarized as follows: "perfect privacy can be achieved by publishing nothing at all—but this has no utility; perfect utility can be obtained by publishing the data exactly as received from the respondents, but this offers no privacy."[148] To illustrate, consider the following:

---

143. Caitlin Fennessy, *The EU-US Data Privacy Framework: A new era for data transfers?*, IAPP (Oct. 7, 2022), https://iapp.org/news/a/the-eu-u-s-data-privacy-framework-a-new-era-for-data-transfers/ [https://perma.cc/3UC7-WBKC].

144. *Id.*

145. *See, e.g.*, Martin J. Conyon, *Big technology and data privacy*, 46 Cambridge J. Econs. 1369, 1369 (2023) ("The collection of individually identifiable data is at the heart of the Facebook business model. The large social networking companies use personal data as a resource, store and bundle that data, and sell it to third parties.").

146. *See* Samuel G. Goldberg et al., *Regulating Privacy Online: An Economic Evaluation of The GDPR* 17–19 (Law & Econs. Ctr. Geo. Mason U. Scalia L. Sch., Rsch. Paper Series No. 22-025, 2022) (finding "a reduction of approximately 12% in both EU user website pageviews and website e-commerce revenue . . . after GDPR's enforcement deadline).

147. Amir Tabakovic, *Only a Little Bit Re-Identifiable?! Good Luck With That…*, MOSTLY AI (July 31, 2020), https://mostly.ai/2020/07/31/only-a-little-bit-re-identifiable [https://perma.cc/TP82-DLQZ]; *see also* Ohm, *supra* note 79, at 1752.

148. Shuchi Chawla et al., *Toward Privacy in Public Databases*, *in* 2 THEORY OF CRYPTOGRAPHY CONFERENCE 363, 364 (Joe Kilian ed., 2005).

Figure 1: Privacy-Utility Tradeoff for Small Datasets

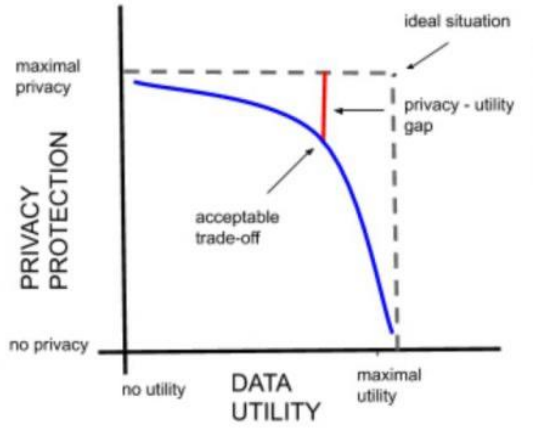
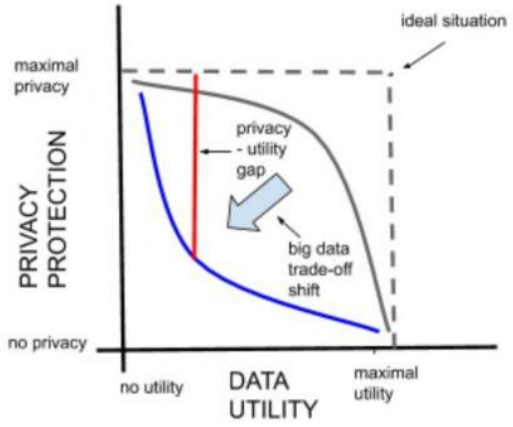


Figure 1[149] illustrates the inverse relationship between privacy and utility. Maximally private data has no utility, and maximally useful data is not private. The "ideal situation" at the dotted intersection, where privacy and utility are maximized, is illusory. An increase in either utility or privacy necessitates a decrease in the other. Figure 2[150] illustrates how the tradeoff is more easily managed in smaller datasets with fewer variables to be deidentified.

Figure 2: Privacy-Utility Tradeoff for Large Datasets



When the privacy-utility tradeoff is applied to larger datasets with hundreds or thousands of attributes, however, Figure 2 demonstrates how quickly privacy can destroy utility and vice-versa. The "big data trade-off shift" differential can be explained by considering the primary

149. Tabakovic, *supra* note 147.
150. *Id.*

shortcoming of deidentification: the possibility of reidentification, or "linkage," of the users underlying deidentified data.[151] Further, Figure 2 demonstrates that the current state of the tradeoff for large datasets is such that even if the dataset is rendered effectively useless by way of deidentification, the deidentified data is *still* not fully anonymous, *i.e.*, it can be linked back to the original data user, thus defeating privacy as well.[152]

Drawing inspiration from HIPAA, data companies currently rely upon deidentification and pseudonymization (the process of replacing personally identifiable information with artificial identifiers to protect individuals' privacy while still allowing the data to be used for specific purposes) as the primary method of compliance with data privacy laws.[153] However, as demonstrated by Figure 2, these methods are not only ineffective at achieving user privacy, but are also ineffective for avoiding GDPR liability because deidentified data can be linked back to the original data subject and falls within the Regulation under Recital 26.[154] In addition to the failure of deidentification and pseudonymization efforts to achieve significant improvements in privacy, such efforts' reduction in data utility run counter to the quintessential BigTech business model.

B. *The BigTech Business Model*

The business model underlying the meteoric rise of BigTech companies like Alphabet (Google's parent company) and Meta (formerly known as Facebook) can be summarized in two words: ad revenue. According to both companies' 10-K filings with the U.S. Securities and Exchange Commission (SEC), "substantially all" of Facebook's $70.7 billion in annual revenue, and Alphabet's nearly $161.9 billion in annual revenue, is earned from advertising.[155] Because advertising requires the attention of potential customers, it follows logically that BigTech's business model revolves around the amount of attention its users give to their devices. The MD&A section of Alphabet's 10-K provides a rather off-putting affirmation of this idea: "[o]ur users are accessing the Internet via diverse devices and modalities, such as smartphones, wearables and smart home devices, and *want to feel connected no matter where they are*

151. *See* Ohm, *supra* note 79, at 1724. Professor Ohm's Article expertly details the numerous ways in which reidentification and linkage between individual people and their data within "anonymous" data sets can occur.

152. Tabakovic, *supra* note 147.

153. 45 C.F.R. §§ 164.502(d), 164.514(a)–(b) (2000).

154. GDPR, *supra* note 3, at Recital 26.

155. Facebook 10-K, *supra* note 5, at 62; Alphabet 10-K, *supra* note 5, at 9.

or what they are doing."[156] The "How we make money" section of Alphabet's 10-K states that "[t]he goal of our advertising products is to deliver relevant ads at just the right time and to give people useful commercial information, regardless of the device they're using."[157] Knowing how "relevant" an advertisement is and what time is "just the right time" requires these companies to collect an immense amount of personal information to accurately characterize their users and predict their behavior in anticipation of the ads they are likely to be most responsive to.

BigTech's public relations teams make these operations sound rather innocuous, however as more information about their operations continues to be revealed, it is perhaps more precise to say that Google's and Facebook's business models depend upon "surveillance capitalism"—a term coined by Harvard professor Shoshana Zuboff, used to describe the practice by which BigTech monitors, monetizes, and subtly influences their users' behavior.[158]

To increase the attractiveness of their advertising capabilities to marketers, Google and Facebook both collect an eerily vast amount of personal data about each of their users—all of it subject to GDPR liability. Both platforms are equipped with biometric recognition features which allow them to identify a user's face and voice.[159] Additionally, these companies can utilize users' search queries and webpage engagement data to identify each user's consumer preferences and religious and political beliefs based on interactions with related webpages.[160] By using scheduling features such as Google Calendar, or Facebook events, these companies can track what users will be doing in the future.[161] These platforms' location services additionally allow them to track where users go, how they get there, how long they spend there, and how often they spend time at specific locations to predict where their

156. Alphabet 10-K, *supra* note 5, at 27 (emphasis added). "MD&A" is an abbreviation for "Management's Discussion and Analysis" in SEC filings. This section's purpose is to provide the public with a more detailed explanation of events and conditions underlying the financial data disclosed.

157. *Id.* at 6.

158. *See* SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER (2019).

159. *See* Ina Fried, *What Facebook Knows About You*, AXIOS (Jan. 2, 2019), https://www.axios.com/facebook-personal-data-scope-suer-privacy-de15c860-9153-45b6-95e8-ddac8cd47c34.html [https://perma.cc/4PE8-P5YG]; *see also* Rob Mardisalu, *What Does Google Know About You: A Complete Guide*, THEBESTVPN (July 9, 2018), https://thebestvpn.com/what-does-google-know-about-you/#infographic [https://perma.cc/AY9N-VS36]. The infographic and the accompanying article provide an illustrative visual about the information Google collects on its users and how the company uses it.

160. Mardisalu, *supra* note 159.

161. *Id.*

users live and work and which locations they otherwise frequent.[162] Such tracking still occurs even if a user turns the location functions off on their device, sign out of their Google or Facebook accounts, or even completely delete their account.[163] In a series of interviews with prominent Silicon Valley BigTech pioneers responsible for designing much of the modern digital world, Netflix's "The Social Dilemma" asserts that social media and BigTech are not just using ads in response to users' digital activity, but are instead using them to influence their behavior on and off the screen.[164]

Illustrative here is the story of Cambridge Analytica. In 2016, the U.K.-based political consulting firm hired by a U.S. presidential contender's campaign was able to target users at an extremely granular level using a technique called "psychographic targeting."[165] By leveraging user data scraped from Facebook's platform, the company was able to send advertisements particularly designed to override humans' innate cognitive defenses by appealing to their most visceral emotions to incite fear or excitement that would theoretically motivate someone to vote for their client in the upcoming election.[166] Although this scandal ultimately ended in a $5.1 billion fine levied against Facebook by the FTC, Facebook CEO Mark Zuckerberg is not yet out of the woods when it comes to data privacy compliance—in fact, he is not even close.[167]

---

162. *Id.*

163. *See* Ryan Nakashima, *AP Exclusive: Google Tracks Your Movements, Like it or Not*, ASSOCIATED PRESS (Aug. 13, 2018), https://apnews.com/article/828aefab64d4411bac257a07c1af0ecb [https://perma.cc/RSE7-6WU7]; *see also* Alfred Ng, *Facebook Still Tracks You After You Deactivate Account: Deactivation does nothing for your privacy*, CNET (Apr. 9, 2019), https://www.cnet.com/news/facebook-is-still-tracking-you-after-you-deactivate-your-account/ [https://perma.cc/X5QE-XLGB].

164. *See* THE SOCIAL DILEMMA (Exposure Labs 2020); *see also* Jonathan Haidt & Tobias Rose-Stockwell, *The Dark Psychology of Social Networks: Why it Feels Like Everything is Going Haywire*, THE ATLANTIC (Dec. 2019), https://www.theatlantic.com/magazine/archive/2019/12/social-media-democracy/600763/ [https://perma.cc/42L7-JT2L]; Tristan Harris, *Our Brains are No Match for Our Technology*, N.Y. TIMES (Dec. 5, 2019), https://www.nytimes.com/2019/12/05/opinion/digital-technology-brain.html [https://perma.cc/JM5T-U484].

165. Sue Halpern, *Cambridge Analytica and the Perils of Psychographics*, NEW YORKER (Mar. 30, 2018), https://www.newyorker.com/news/news-desk/cambridge-analytica-and-the-perils-of-psychographics [https://perma.cc/BF89-NDW6].

166. *Id.*

167. Mike Isaac & Cecilia Kang, *Facebook Expects to Be Fined Up to $5 Billion by F.T.C. Over Privacy Issues*, N.Y. TIMES (Apr. 24, 2019), https://www.nytimes.com/2019/04/24/technology/facebook-ftc-fine-privacy.html [https://perma.cc/6AKE-MG72]; Natasha Lomas, *First Major GDPR decisions looming on Twitter and Facebook*, TECHCRUNCH (May 22, 2020), https://techcrunch.com/2020/05/22/first-major-gdpr-decisions-looming-on-twitter-and-facebook/ [https://perma.cc/5M6W-E8QP]; *see also* Emily Price, *The EU Could Hit Facebook with Billions in Fines Over Privacy Violations*, DIGITAL TRENDS (Aug. 12, 2019), https://www.digitaltrends.com/social-media/facebook-gdpr-decision/ [https://perma.cc/F3R8-2NPG].

When analyzing these practices and revelations in the context of data subjects' rights under GDPR as discussed in Part I.A., one can easily imagine why compliance is so difficult for BigTech. The methods by which these companies strive for compliance (deidentification and pseudonymization) create a disconnect between data and its underlying user, while the lifeblood of the BigTech business model is such a data-to-underlying-user connection. In fact, Facebook cites "decreases in user engagement, including time spent on our products" and "failure to accept our terms of service, as part of changes that we implemented in connection with . . . GDPR" as risks to their financial performance.[168] Alphabet's 10-K contains similar language that indicates that decreased usage and GDPR and CCPA are both major threats to their revenue model as well.[169] Both companies are investing plenty of time and capital into compliance while both U.S. and EU regulators are wasting no time assessing fines for violations.[170]

C. *Current Compliance Methods & the Re-Identification Problem*

Current measures for privacy compliance generally involve some variation of data "anonymization"[171] techniques which essentially come down to stripping data of personal identifiers.[172] "Anonymization" is an umbrella term used to describe a variety of techniques (e.g., deidentification, pseudonymization) aimed at removing personal identifiers from sets of personal data.[173] Each of these techniques may themselves have multiple meanings or involve distinct processes across different laws or jurisdictions.

For example, deidentification under HIPAA refers to stripping a dataset of the eighteen enumerated identifiers.[174] HIPAA's Privacy Rule provides a "safe harbor" for such deidentified data.[175] The statute provides two permissible methods for deidentifying data: (1) removal of the eighteen enumerated identifiers or (2) expert certification.[176] The identifiers include names, telephone numbers, addresses, biometric identifiers, medical record numbers, etc.[177] Expert certification requires that a "person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for

168. Facebook 10-K, *supra* note 5, at 11.
169. Alphabet 10-K, *supra* note 5.
170. Isaac & Kang, *supra* note 167.
171. Ohm, *supra* note 79, at 1706–09.
172. Bellovin et al., *supra* note 11.
173. Ohm, *supra* note 79, at 1706–15.
174. 45 C.F.R. § 164.514(e) (2013).
175. *Id.* § 164.514(e).
176. *Id.* § 164.514.
177. *Id.* § 164.502(d)(2).

rendering information not individually identifiable" applies such knowledge to "determine[] that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information."[178]

This system is far from perfect, however. As Professor Paul Ohm pointed out in 2010, "anonymization" is a bit of a misnomer as it has been used in privacy scholarship.[179] "Deidentification" is a more precise term, because "anonymization" in its colloquially defined use refers to simply stripping raw data of its unique identifiers rather than rendering the data into a state where its origin is indeed truly "anonymous."[180] For a practical example of "deidentification," picture a dataset measuring consumer preferences for Coke or Pepsi which contains a list of respondents' names, sex, dates of birth, ZIP codes, addresses, and email addresses. A deidentified version of the same dataset would likely preserve the indicated beverage preference; ZIP code, to determine geographic variance; and dates of birth, to assess variance among age demographics. It would omit obviously identifiable information like the respondents' names, addresses, and email addresses.

Over ten years ago, however, Ohm and other scholars lifted the veil on how deidentified data still has the potential to be linked back to individual underlying data subjects through the comparison of the deidentified data set with additional information relevant to the data subject (a phenomenon known as "reidentification"), thus defeating the purpose of deidentification. His article[181] illustrated how deidentification is wholly insufficient as a privacy protection measure by highlighting, among other indications, a study conducted by professor of computer science, Latanya Sweeney, in which she was able to identify 87.1 percent of Americans using only their ZIP code, date of birth and sex, each of which are supposedly non-identifying data points which are likely to be found within "deidentified" data sets.[182] Although two subsequent studies were unable to replicate that 87.1 percent finding, they did successfully reidentify 63 and 61 percent of data subjects from 1990 and 2000 U.S. census data, respectively.[183]

178. *Id.* § 164.514.

179. Ohm, *supra* note 79, at 1716.

180. *Id.*

181. Ohm, *supra* note 79, at 1707–08.

182. *Id.* at 1719–20 (citing Latanya Sweeney, *Uniqueness of Simple Demographics in the U.S. Population* (Lab'y for Int'l Data Priv., Working Paper No. 3, 2000)).

183. Philippe Golle, *Revisiting the Uniqueness of Simple Demographics in the US Population*, ASS'N FOR COMPUTING MACH. 77, 78 (2006).

Pseudonymization is effectively analogous to deidentification and carries the same risks of reidentification.[184] The distinction is that instead of merely redacting or omitting identifiers, pseudonymized data renames them with a string of characters and the controller of the dataset preserves a legend that can be used to link the data subject with the unique identifier to which their data has been assigned.[185] In a pseudonymized dataset, "John Smith" becomes "user027462."

The risk of reidentification has even been recognized in GDPR, which provides: "personal data which have undergone pseudonymization, which could be attributed to a natural person by the use of additional information, should be considered as information on an identifiable natural person."[186] In other words, even if Google strips a dataset of its personally identifiable characteristics, it is still treated as personal data under Article 4(1) and Recital 26 of the Regulation if an observer can pair that data with *any* additional information that will allow him to discover its corresponding data subject.

In short, deidentification and pseudonymization both strip the underlying data of any way to attribute the data to an individual user when taken *alone*. However, deidentified and pseudonymized data can all be traced back to an individual with relative ease once an observer gains access to the database, either by cross-referencing dates of birth with sex and ZIP code in a deidentified dataset as Professor Sweeney did, or by accessing a pseudonymized dataset's corresponding legend.[187] With the explosion in the volume of personal data collected in the past several years, reidentification can reasonably be expected to become easier as techniques simultaneously improve.

D. *Recent Enforcement Actions*

Following GDPR's implementation, total fines for violations have reached a staggering $$2,779,699,894, as of February 2023[188] According to a 2020 report[189] by U.K. software company Exonar, "39% (appx. $244

184. Sophie Stalla-Bourdillon & Alison Knight, *Anonymous Data v. Personal Data - False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data*, 34 WIS. INT'L L.J. 284, 286–87 (2016).

185. *See id.*

186. GDPR, *supra* note 3, at Recital 26.

187. Bellovin et al., *supra* note 11, at 9–18; *see also* Ohm, *supra* note 79, at 1744–48.

188. *GDPR Enforcement Tracker*, *supra* note 9.

189. Sead Fadilpašić, *Majority of GDPR Penalties Issued as a Result of These Two Problems*, IT PRO PORTAL (Oct. 16, 2020), https://web.archive.org/web/20201031065033/https://www.itproportal.com/news/majority-of-gdpr-penalties-issued-as-a-result-of-these-two-problems/.

million)[190] of GDPR-related fines were due to insufficient security, [and] 25% of fines (appx. $159 million) were related to unsecured or over-retained data."[191] The GDPR Enforcement Tracker tells a similar story.[192] Data from the Tracker indicates that the two most frequent categories of GDPR violations are: (1) "[i]nsufficient technical and organizational measures to ensure information security," typically related to data breaches where user data is exposed like in the British Airways case[193] (though this category also includes violations whereby a user discovers that a company is storing their personal data with insufficient cybersecurity measures);[194] and (2) "[i]nsufficient legal basis for data processing," which can include processing data without sufficient consent under Article 6.[195]

In the United Kingdom, the British Information Commissioner's Office (ICO) found that British Airways violated Articles 5(1)(f)[196] and 32[197] of the Regulation when a 2018 cyberattack exposed the data of nearly 430,000 customers due to what ICO found to be inadequate security measures.[198] The ICO initially imposed a $200 million fine but reduced it to $25 million on appeal due to the financial hardship the airline had been under due to the devastating economic impact of the COVID-19 pandemic in 2020.[199]

In a seminal example of an enforcement action taken in response to an Article 6 violation, the French data regulator, le Commission Nationale de l'Informatique et des Libertés (CNIL), imposed a $57

190. Fines reflected here were current as of October 16, 2020, but large fines such as the £183 million (approximately $219 million USD) fine levied upon British Airways by the U.K.'s Information Commissioner's Office was reduced to £20 million (approximately $24 million USD) on appeal due to the financial hardship experienced by the airline during the COVID-19 pandemic. Carly Page, *U.K. Privacy Watchdog Hits British Airways With Record-Breaking £20 Million GDPR Fine*, FORBES (Oct. 16, 2020, 5:39 AM), https://www.forbes.com/sites/carlypage/2020/10/16/ico-hits-british-airways-with-record-breaking-fine-for-2018-data-breach/?sh=60b15bd9481a [https://perma.cc/SC3H-BJD7].

191. Fadilpašić, *supra* note 189.

192. *GDPR Enforcement Tracker*, *supra* note 9.

193. Press Release, UK Information Commissioner's Office, *Intention to Fine British Airways £183.39m Under GDPR for Data Breach* (July 8, 2019); *see also* Statement, UK Information Commissioner's Office, *Intention to Fine Marriott International, Inc More Than £99 Million Under GDPR for Data Breach* (July 9, 2019).

194. GDPR, *supra* note 3, at art. 32(1).

195. *See supra* Part I.A.

196. "Personal data shall be . . . processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized(sic) or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational(sic) measures ('integrity and confidentiality')." GDPR, *supra* note 3, at art. 5(1)(f).

197. GDPR, *supra* note 3, at art. 32.

198. *See* Press Release, UK Information Commissioner's Office, *supra* note 193.

199. Page, *supra* note 190.

million fine on Google—and rejected its appeal to reduce it.[200] CNIL found that Google had not provided "sufficiently clear" information to its consumers regarding how it processes their personal data for the purpose of providing targeted ads, and that the company failed to obtain the consumers' informed consent.[201] The ruling on failure to obtain informed consent rested on two separate grounds. The agency court first noted that consumers had to click through five to six pages before they could meaningfully access the settings on how their data was collected and how it would be used.[202] They then ruled that the collected consent is neither "specific" nor "unambiguous" because a user has to agree to Google's terms of service and privacy policy before accessing the platform, therefore giving their consent in full before having a chance to modify the collection options offered.[203] Thus, users' consent was not lawfully obtained, therefore nullifying their "consent" as a lawful basis for processing under Article 6(1)(a).[204] In the U.S., we consider these kinds of clickwrap agreements to be routine. These enforcement actions thus illustrate the contrast between standard industry practices and the practices which GDPR mandates.

In a July 2021 SEC filing, Amazon disclosed to the public that it was issued a fine of approximately $797 million by the Luxembourg National Commission for Data Protection for failing to comply with GDPR regarding the processing of personal data.[205] A French privacy organization called La Quadrature du Net complained to CDNP in 2018, alleging that Amazon's targeted advertising strategies involved undisclosed data collection tactics for which it failed to obtain user consent, as required under GDPR.[206] Little is known about the details of this enforcement action, however, as the CNPD has stated that due to secrecy laws in Luxembourg, it cannot comment on individual cases or complaints.[207] It is uncertain if the CNPD will publish its findings, as they

200. Press Release, Commission Nationale de l'Informatique et des Libertés (CNIL), The CNIL's Restricted Committee Imposes a Financial Penalty of 50 Million Euros Against Google LLC (Jan. 21, 2019) [hereinafter CNIL Decision].

201. *Id.*

202. *Id.*

203. *Id.*

204. Lomas, *supra* note 167.

205. Amazon.com Inc., Quarterly Report (Form 10-Q) (June 30, 2021), https://www.sec.gov/ix?doc=/Archives/edgar/data/0001018724/000101872421000020/amzn-20210630.htm#i5986f88ea1e04d5c91ff09fed8d716f0_103.

206. *Amazon Fined 746 Million Euros Following our Collective Legal Action*, La Quadrature du Net (July 30, 2021), https://www.laquadrature.net/en/2021/07/30/amazon-fined-746-million-euros-following-our-collective-legal-action/ [https://perma.cc/7V5S-N9US].

207. Press Release, Decision Regarding Amazon Europe Core S.À.R.L., Luxembourg National Commission for Data Protection (Aug. 6, 2021), https://cnpd.public.lu/en/actualites/international/2021/08/decision-amazon-2.html [https://perma.cc/E7AF-9P56].

are usually anonymous unless special powers are invoked.[208] According to one U.K. law firm, details from the complaint suggest that the case focused on whether Amazon had a sufficient lawful basis for processing personal data, and Amazon's argument that it could process personal data based on a contract with data subjects.[209]

Privacy enforcement has ramped up in the United States as well, though not as aggressively as in the EU. As the 2010 FTC Report highlights, "[s]ince 2001, the FTC has used its authority under a number of statutes—including the FCRA, the GLB Act, and Section 5 of the FTC Act—to bring 29 cases against businesses that allegedly failed to protect consumers' personal information."[210] However, contrary to the authors' congratulatory tone employed in the 2010 FTC Report, twenty-nine cases in nine years pales in comparison to GDPR's aggressive enforcement, with 611 fines issued between the Regulation's enactment in January 2018 and April 2021.[211] In the FTC's *Privacy & Data Security Update: 2019*, the Commission states that they have brought "more than 130 spam and spyware cases and 80 general privacy lawsuits," and based on figures from the two preceding years, these numbers appear to be cumulative. [212] To be fair, GDPR's enforcement covers a broad array of data privacy and security actions that may not fall within the scope of the FTC's § 5 authority, which may explain the FTC's apparent dearth of enforcement actions when compared to GDPR. Additionally, while Amazon's $797 million fine is the largest GDPR fine issued to date, the FTC has issued *massive* fines.[213] Recently, the Commission fined Facebook $5 billion—

208. Jonathan Armstrong & Katherine Eyres, *Client Alert: Amazon fined €746 million by Luxemburg Data Protection Regulator for GDPR infringements*, Cordery Legal Compliance (Oct. 18, 2021), https://www.corderycompliance.com/amazon-fine-lux-cnpd/ [https://perma.cc/S6UA-55GY].

209. *Id.*

210. Lomas, *supra* note 167, at n.17 (highlighting a list of exemplary cases).

211. *GDPR Enforcement Tracker*, *supra* note 9.

212. *See* FED. TRADE COMM'N, PRIVACY & DATA SECURITY UPDATE: 2019, at 2 (2020), https://www.ftc.gov/system/files/documents/reports/privacy-data-security-update-2019/2019-privacy-data-security-report-508.pdf [https://perma.cc/7HG4-QKCK] (The FTC brought "130 spam and spyware cases and 75 general privacy lawsuits," in 2018, and "130 spam and spyware cases and 50 general privacy lawsuits" in 2017.); *see also* FED. TRADE COMM'N, PRIVACY & DATA SECURITY UPDATE: 2018, at 3 (2019), https://www.ftc.gov/system/files/documents/reports/privacy-data-security-update-2018/2018-privacy-data-security-report-508.pdf [https://perma.cc/24NR-DTHG]; FED. TRADE COMM'N, PRIVACY & DATA SECURITY UPDATE: 2017, at 2 (2018), https://www.ftc.gov/system/files/documents/reports/privacy-data-security-update-2017-overview-commissions-enforcement-policy-initiatives-consumer/privacy_and_data_security_update_2017.pdf [https://perma.cc/VRC5-JA3J].

213. Press Release, Federal Trade Commission, FTC Imposes $5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook (July 24, 2019), https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions

the largest fine ever issued for a consumer data privacy violation.[214] According to the settlement, the FTC found that Facebook misrepresented users' ability to control the privacy of their information and deceptively shared information about users and their friends with third-party applications.[215]

The use of synthetic data may have reduced the severity of, or even fully prevented, the data privacy violations at issue in some of the enforcement actions discussed above. If the data that had been breached in the British Airways case, for example, was synthetic data, the number of customers whose identifiable personal data was exposed could have been reduced. In the Google and Facebook cases above, both companies would still have had to obtain meaningful consent to process and collect user data (which necessarily precedes the generation and use of synthetic data), so the use of synthetic data would not likely have changed the outcome. However, regardless of the consent issue, any company is susceptible to a data breach, and storing synthetic data rather than personal data would be consistent with GDPR's principles of data minimization and related privacy principles, and would thus likely reduce the harm and consequent fines resulting from a potential data breach.

As discussed in Part I *infra*, *Schrems II* upped the ante even more for U.S.-based companies who enjoyed the protections of the EU-U.S. Privacy Shield that insulated them from liability for processing EU residents' data.[216] By contrast, the current state of affairs following the collapse of the EU-U.S. Privacy Shield, with public scrutiny and enforcement actions ramping up, increased interest in privacy, and emergent privacy laws, gives data privacy compliance heightened urgency, with reliable means of limiting data privacy infractions becoming increasingly valuable to data companies. As such, due to its potential to proactively prevent or mitigate such infractions, synthetic data could prove to be a very useful compliance tool at this particular juncture.

(noting that before the Facebook penalty, the FTC fined Uber, British Airways and Equifax $148 million, $230 million, and $275 million, respectively, in what were previously considered to be the largest penalties to date).

214. *Id.*

215. Complaint at 1–5, United States v. Facebook, Inc., Case No. 19-cv-2184 (D.C. Cir. 2019), https://www.justice.gov/opa/press-release/file/1186511/download [https://perma.cc/U9QN-CD5D].

216. *Schrems II Landmark Ruling: A Detailed Analysis*, NORTON ROSE FULBRIGHT (July 2020), https://www.nortonrosefulbright.com/en/knowledge/publications/ad5f304c/schrems-ii-landmark-ruling-a-detailed-analysis [https://perma.cc/9RUV-CH5A]; *Schrems II Confirms Validity of EU Standard Contractual Clauses, Invalidates EU–U.S. Privacy Shield*, JONES DAY (July 2020), https://www.jonesday.com/en/insights/2020/07/schrems-ii-confirms-validity [https://perma.cc/2KPC-5ZXH].

III. TRUE ANONYMIZATION WITH SYNTHETIC DATA TO AVOID EU AND U.S. REGULATORY INFRACTIONS

The multi-layered compliance challenges faced by data companies as a result of shifting global privacy laws, increasingly aggressive enforcement, and conflicts with prevailing business models have created a need, now more so than ever, for innovations in the BigTech space. Synthetic data may represent this much-needed innovation. Synthetic data stands out as a particularly useful compliance tool because it can help companies reliably adhere to data privacy regulations, within both the U.S. and EU, while resolving the privacy-utility tradeoff which plagues traditional compliance methods. [217] Because synthetic datasets are entirely fabricated, they are *truly* anonymous in the sense that the underlying data subjects cannot possibly be identified, thus rendering GDPR, CCPA, HIPAA, and other major privacy laws inapplicable. This Part first describes the process of developing synthetic data, then discusses its efficacy under EU and U.S. privacy law.

A. *Synthetic Data Primer*

AI-created synthetic data might prove to be a potent solution to the compliance issues that more effectively balances the needs of tech companies with the privacy rights of consumers.[218] Synthetic data is essentially "fake" data made from real data that is statistically equivalent to the authentic personal data that it is given.[219] It uses an original dataset comprised of personal data and creates an entirely new "synthetic" dataset in an irreversible one-way hashing process that makes it impossible for hackers or malicious insiders to recreate the original personal data or to identify its source.[220] Unlike pseudonymized data, synthetic data cannot be used to identify original users.[221]

Synthetic data is created using two methods of AI: variational autoencoders (VAE) and generative adversarial networks (GAN).[222] As Unite AI's Dan Nelson explained:

---

217. Randy Koch, *Opinion: TikTok's Data-Privacy Problem Has an East Solution: 'Synthetic Data'*, MARKETWATCH (Oct. 2, 2020, 7:57 AM), https://www.marketwatch.com/story/tiktoks-data-privacy-problem-has-an-easy-solution-synthetic-data-2020-10-02?siteid=yhoof2 [https://perma.cc/WXA3-TVKF].

218. Tordable, *supra* note 14; Stefanie Koperniak, *Artificial data give the same results as real data — without compromising privacy*, MIT NEWS (Mar. 3, 2017), https://news.mit.edu/2017/artificial-data-give-same-results-as-real-data-0303 [https://perma.cc/M4JY-UNGX].

219. Grace Brodie, *Five Compelling Use Cases for Synthetic Data*, HAZY (June 1, 2020), https://hazy.com/blog/2020/06/01/five-use-cases-for-synthetic-data [https://perma.cc/B4VE-4GQ5].

220. Daniel Nelson, *What is Synthetic Data?*, UNITE AI (Sept. 14, 2020), https://www.unite.ai/what-is-synthetic-data/ [https://perma.cc/BWR9-A5CX].

221. *Id.*

222. *Id.*

> VAEs are unsupervised machine learning models that make use of encoders and decoders. The encoder portion of a VAE is responsible for compressing the data down into a simpler, compact version of the original dataset, which the decoder then analyzes and uses to generate a representation of the base data. A VAE is trained with the goal of having an optimal relationship between the input data and output, one where both input data and output data are extremely similar.
>
> ****
>
> When it comes to GAN models, they are called "adversarial" networks due to the fact that GANs are actually two networks that compete with each other. The generator is responsible for generating synthetic data, while the second network (the discriminator) operates by comparing the generated data with a real dataset and tries to determine which data is fake. When the discriminator catches fake data, the generator is notified of this and it makes changes to try and get a new batch of data by the discriminator. In turn, the discriminator becomes better and better at detecting fakes. The two networks are trained against each other, with fakes becoming more lifelike all the time.[223]

"Generators" are thus able to generate increasingly lifelike fake datasets as time goes on.[224]

Stanford researchers have used an apt analogy in explaining GAN-generated synthetic data: counterfeit money.[225] The generator component studies the details of the dollar bill, creates what it believes to be an indistinguishable copy, and the discriminator scrutinizes its details and sends it back to the generator when it finds a distinction. The process repeats until the discriminator cannot separate the authentic from the counterfeit. Not only can synthetic data optimize compliance, but it can foster innovation by creating simulations.[226] In practice, synthetic data has been the key to recent advancements in self-driving cars, development of vaccines to SARS-CoV-2, and it is even being used by Facebook now to train AI algorithms to identify language that resembles bullying to augment their content moderation practices.[227] Thus, there are

223. *Id.*

224. *Id.*

225. Bellovin et al., *supra* note 11, at 5.

226. Nelson, *supra* note 220.

227. *See* Yashar Behzadi, *Why Synthetic Data Could Be the Ultimate AI Disruptor*, TDWI (June 28, 2019), https://tdwi.org/articles/2019/06/28/adv-all-synthetic-data-ultimate-ai-dis

numerous potential benefits available to companies who make use of such synthetic data technology, even beyond those relating to the issues discussed in this Note.

B. *Privacy Law Exceptions for Synthetic Data*

Synthetic data is generally exempt from the provisions of many data privacy regulations. For example, GDPR's Recital 26 provides:

> The principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone *pseudonymisation*, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. . . . The principles of data protection should therefore not apply to *anonymous* information, namely information which does not relate to an identified or identifiable natural person or to *personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable*.[228]

Notably, Recital 26 distinguishes between *pseudonymized* data and *anonymized* data, as discussed in Part II.B. To reiterate, the term "anonymized" in its conventional use (referring to pseudonymized or deidentified data) is a misnomer because such data may be reidentified, and is thus not truly "anonymous" under Recital 26. Synthetic data, however, does fall under the definition of "anonymous" as provided for in Recital 26 because it can never be linked back to the underlying data subject. Thus, as per Recital 26, "[t]he principles of data protection should therefore not apply" to synthetic data, which is not regulated under GDPR.

Similarly, synthetic data does not fall under the several definitions of "personal information," or the equivalent term, under the various U.S. data privacy laws, even HIPAA's definition of PHI, as discussed in Part I. *infra*.[229] U.S. privacy laws generally follow HIPAA's permissible

---

ruptor.aspx [https://perma.cc/E62J-PU95]; *see also* Brandi Vincent, *NIH Partners with Israeli Startup to Generate Synthetic COVID-19 Data*, NEXTGOV (June 18, 2020), https://www.nextgov.com/emerging-tech/2020/06/nih-partners-isreali-startup-generate-synthetic-covid-19-data/166255/ [https://perma.cc/F62K-M7H2]; *The Rise of Synthetic Data to Help Developers Create and Train Algorithms Quickly and Affordably*, INSIDEBIGDATA (May 8, 2018), https://insidebigdata.com/2018/05/08/rise-synthetic-data-help-developers-create-train-ai-algorithms-quickly-affordably/ [https://perma.cc/8AGM-BDU8]; Tordable, *supra* note 14.

228. GDPR, *supra* note 3, at Recital 26 (emphasis added).

229. 45 C.F.R. § 160.103 (2014).

methods of creating deidentified data, meaning data where specific identifiers have been removed.[230]

C. *Synthetic Data as a Compliance Solution*

Because synthetic data escapes the various definitions of "personal data" discussed in this Article, its future as a data privacy compliance tool is quite promising—and investors have taken notice.[231] Its value proposition is simple: a data controller can collect a small, representative dataset, replicate its utility with stunning accuracy to a global scale, all while reducing the associated exposure to data privacy regulations, and at a fraction of the cost of traditional compliance methods.

1. Benefits of Synthetic Data

The principal benefits of synthetic data are statistical equivalence, , ease in achieving regulatory compliance, and cost-effectiveness (given the high value of synthetic data and its relative low cost of generation). Synthetic data startups are raising significant amounts of capital as investors have begun to realize synthetic data's potential in financial technology, healthcare, government, telecom, pharmaceuticals, e-commerce, transportation and logistics, manufacturing, and, as this Note suggests, consumer-based data platforms.[232]

Synthetic data is already being utilized as a HIPAA-compliant, and extremely efficient, replacement for PHI for COVID-19 vaccine researchers.[233] For example, the National Institute of Health partnered with Syntegra, a synthetic data company, to generate a comprehensive synthetic database using over 2.6 million COVID patients' health information.[234] Because patients' healthcare providers validated the data, the synthetic data was accurate. NIH exclusively utilized this synthetic data in their research efforts, without retaining any real patient data. And bea was And because the synthetic data was fed into NIH's database, no

230. *See supra* Part I.B.2.b.

231. *See, e.g.*, Emil Protalinski, *Zapata raises $38 million for quantum machine learning*, VENTUREBEAT (Nov. 19, 2020), https://venturebeat.com/2020/11/19/zapata-raises-38-million-for-quantum-machine-learning/ [https://perma.cc/MA7D-YSXK].

232. Sri Muppidi, *Growing Applications of Synthetic Data*, SIERRA VENTURES (Sept. 22, 2020, 6:00 AM), https://www.sierraventures.com/blog/growing-applications-of-synthetic-data/ [https://perma.cc/Y7DR-3BDP].

233. Julia Evangelou Strait, *Synthetic data mimics real patient data, accurately models COVID-19 pandemic*, Wash. U. Sch. Med. St. Louis (Apr. 27, 2021), https://medicine.wustl.edu/news/synthetic-data-mimics-real-patient-data-accurately-models-covid-19-pandemic/ [https://perma.cc/NJ54-F4AX].

234. Brandi Vincent, *Synthetic Data Engine to Support NIH's COVID-19 Research-Driving Effort*, NEXTGOV (Jan. 14, 2021), https://www.nextgov.com/analytics-data/2021/01/synthetic-data-engine-support-nihs-covid-19-research-driving-effort/171421/ [https://perma.cc/CMX8-F8SS].

patients' privacy was ever at risk of being violated in through NIH's research process. In fact, in January of 2021, the Department of Health and Human Services, recognizing the great potential of synthetic data in healthcare, opened the Synthetic Health Data Challenge, offering $100,000 in prize money for competitive solutions.[235] The goal of the program is to "[e]ngag[e] the broader community of researchers and developers to validate the realism and demonstrate the potential uses of the generated synthetic health records through a challenge," with a focus on synthetic opioid, pediatric, and complex care patient records.[236]

Google's fate in France would have likely been different in 2019 if the company had used synthetic data.[237] If it had used synthetic data that was truly anonymized and fell outside of GDPR's scope, it could have avoided liability and still provided useful consumer trends for advertisers that were statistically equivalent to the data that ended up costing them $57 million. Because Google currently dominates the search engine market (so much so that they are facing antitrust action in the U.S.) with a whopping 88% market share in the U.S. market for general search engines and 70% in the search advertising market, they are in such a powerful position that advertisers will have a hard time finding a better place to take their advertising expenditures.[238]

Moreover, data companies may realize cost savings from synthetic data use through reductions in time and labor costs required to manually deidentify personal data, the expense of manually labeling datasets purchased from a collector for another purpose, or by leveraging synthetic data's predictive capacity to entirely replace tests or surveys needed to collect the data in the first place.[239] Data companies could also be saving shareholders millions, if not billions, per year in foregone enforcement fines for GDPR violations or similar data privacy laws as other jurisdictions start to catch on. Most of the companies generating synthetic data are private, and therefore so are their financials and price points. Amazon Web Services, however, offers access to its own

---

235. Department of Health & Human Services – Office of National Coordinator for Health Information Technology, *Synthetic Health Data Challenge*, CHALLENGE.GOV (accessed Apr. 30, 2021), https://www.challenge.gov/challenge/synthetic-health-data-challenge/.

236. *Synthetic Health Data Generation to Accelerate Patient-Centered Outcomes Research*, DEPT. HEALTH & HUM. SERVS., https://www.healthit.gov/topic/scientific-initiatives/pcor/synthetic-health-data-generation-accelerate-patient-centered-outcomes [https://perma.cc/2FNX-DZCR] (July 21, 2021, 5:00PM).

237. *See* CNIL Decision, *supra* note 200.

238. Katherine Kemp, *The US is taking on Google in a huge antitrust case—it could change the face of online search*, TECH XPLORE (Oct. 21, 2020), https://techxplore.com/news/2020-10-google-huge-antitrust-caseit-online.html [https://perma.cc/FQ4R-ALR7].

239. Tordable, *supra* note 14 ("Acquiring and storing all of this data [for autonomous vehicle development] from live tests of real cars on real roads would have been too expensive and cumbersome.").

synthetic data generator for $995 per year.[240] Such a cost represents a significant decrease from the typical costs associated with current deidentification processes.

2. Drawbacks

While promising, synthetic data is not entirely foolproof. Much like any algorithm or dataset, the outcomes are only as good as the underlying data. Synthetic data (just like traditional, identifiable data) that is not effectively controlled for racial bias can exacerbate discriminatory outcomes.[241] Some researchers claim that they can eliminate bias in GAN-generated data by employing weak supervision and weighing input variables susceptible to bias,[242] though some commentators remain skeptical of this claim.[243]

Additionally, there is a minimal, but still present, possibility of "leakage" of PII if synthetic data is not paired with additional privacy preserving features like differential privacy.[244] At a high level, differential privacy is a technique whereby a statistician includes enough noise into a dataset to induce a sufficient level of deniability so that an entry of "yes" or "no" into a dataset becomes "maybe."[245] This reduces the data's utility by design to make it less useful for hackers, but also reduces utility for a lawful custodian.[246]

In the consumer data context, integrating synthetic data into the BigTech advertising model will likely reduce the precision with which these companies can bring advertisements. Because Google currently dominates the search engine market (so much so that they are facing antitrust action in the U.S.) with an 88% market share in the U.S. market for general search engines and 70% in the search advertising market,[247] they are in such a powerful position that advertisers will have a difficult

240. *See Synthetic Data Generator*, AMAZON WEB SERVS., https://web.archive.org/web/20210116190558/https://aws.amazon.com/marketplace/pp/Synthetic-Data-Generator-Synthetic-Data-Generator/B07XPJ8Z7M [https://perma.cc/MQ54-G3X7] (last visited Nov. 29, 2020).

241. Todd Feathers, *Fake Data Could Help Solve Machine Learning's Bias Problem—if We Let It*, SLATE (Sept. 17, 2020, 9:00 AM), https://slate.com/technology/2020/09/synthetic-data-artificial-intelligence-bias.html [https://perma.cc/QUY2-E497].

242. *Id.* (citing Choi et al., *Fair Generative Modeling Via Weak Supervision*, ASS'N FOR COMPUT. MACH. (July 13, 2020), https://dl.acm.org/doi/pdf/10.5555/3524938.3525114 [https://perma.cc/7P97-KC8N].

243. Sage Lazzaro, *AI experts refute Cvedia's claim its synthetic data eliminates bias*, Venture Beat (July 26, 2021, 2:20PM) https://venturebeat.com/ai/ai-experts-refute-cvedias-claim-its-synthetic-data-eliminates-bias/ [https://perma.cc/A5UQ-SVSS].

244. *See* Bellovin et al., *supra* note 11, at 18–21.

245. *Id.*

246. *Id.*

247. Tiago Bianchi, *Worldwide desktop market share of leading search engines from January 2015 to December 2022*, Statista (Jan. 6, 2023), statista.com/statistics/216573/worldwide-market-share-of-search-engines/ [https://perma.cc/4L29-5P5R].

time finding a better place to take their advertising expenditures. Even if synthetic data is not as useful as genuine user data, GDPR applies significant legal and regulatory risk evenly to competitors.[248] Thus BigTech's dominance in the attention market is highly unlikely to change solely due to the use of synthetic data.

Further, as discussed in Part I *infra*, psychographic targeting may present significant risks, as illustrated in the Cambridge Analytica affair. Synthetic data's hampering of such targeted advertisement efforts may thus be viewed as a positive by some.[249] Additionally, there is a valid concern regarding the validation of the underlying dataset from which synthetic data is generated. Without adequate validation methods, like the NIH's use of PHI validated by hospitals,[250] synthetic data could be used nefariously to mislead people who rely on it if the data or validation methods are not available for independent scrutiny.

CONCLUSION & RECOMMENDATIONS

The statistical and functional equivalence between synthetic and authentic data alleviates the tension between BigTech's enormous appetite for personal data and the privacy requirements of current data privacy laws. Synthetic data, by definition, is anonymized data under GDPR's Recital 26 and similarly falls outside of the scope of the U.S.'s several definitions of "personal information," as there is no way for an outside observer to identify the original data subjects underlying the synthetic dataset. Yet, despite this lack of identifiability, synthetic data preserves the statistical outcomes, and thus the utility, of the underlying authentic data. As one commentator aptly noted, "[i]f an organization can identify all of its personal data, take it out of the data security and compliance equation completely—rending it useless to hackers, insider threats, and regulation scope—it can eliminate a huge amount of risk, and drastically reduce the cost of compliance."[251] Thus, an innovation such as synthetic data could help privacy-conscious data subjects and anxious BigTech CEOs alike sleep better at night knowing that the big data engines are still humming while user privacy is being protected.

By incorporating truly anonymous, privacy-compliant, synthetic data into the BigTech business model, companies like Google and Facebook

248. *See supra* Part I.A.2.

249. Rebecca Walker Reczek et al., *Targeted Ads Don't Just Make You More Likely to Buy — They Can Change How You Think About Yourself*, HARV. BUS. REV. (Apr. 4, 2016), https://hbr.org/2016/04/targeted-ads-dont-just-make-you-more-likely-to-buy-they-can-change-how-you-think-about-yourself [https://perma.cc/F9SP-G34T].

250. Vincent, *supra* note 234.

251. Randy Koch, *GDPR, CCPA and Beyond: How Synthetic Data Can Reduce the Scope of Stringent Regulations*, HELP NET SEC. (Apr. 14, 2020), https://www.helpnetsecurity.com/2020/04/14/synthetic-data/ [https://perma.cc/4KW8-7NMZ].

could continue to operate in their current, highly successful fashion while resolving the challenges presented by the Privacy-Utility Tradeoff by protecting their users' privacy while continuing to profit off of mass data collection. Facebook and Google could provide advertisers with synthetic datasets that reflect unique consumer consumption trends that, while not as specific and granular as they are currently, are effective enough for them to track changes in market trends and advertise to potential consumers. Consumers who are unbothered by the amount of personal data currently collected by BigTech companies can opt into data collection from their use of the platforms or their devices, decreasing controllers' threshold for the volume of useful data. In exchange, advertisers can offer discounts for users who opt in. This would provide BigTech enough seed data for a synthetic dataset to accurately replicate and would give advertisers a means to continue to reach their target audience.

Even if this practice were not as effective for targeted advertisements, causing advertisers and political campaigns to gripe at the decline in the return on investment from marketing expenditures, data controllers and their shareholders can avoid hefty blows to their bottom lines caused by violations of GDPR and similar forthcoming privacy regulations and mitigate reputational damage as users have begun to prioritize privacy. Ultimately, the potential decline in advertising effectiveness is vastly outweighed by the substantial public policy interest in protecting individuals' rights to privacy and providing users with a way to escape the invasive and Orwellian digital world we have found ourselves in.[252]

252. *See supra* Part II.A.2.