

GEORGETOWN LAW TECHNOLOGY REVIEW

VOLUME 4, ISSUE 1

FALL 2019

LETTER FROM THE EDITORS

Dear Reader:

Thank you for picking up the first issue of Volume 4 of the *Georgetown Law Technology Review* (GLTR). Over the past four years, GLTR has established itself as one of the premier technology law journals in the country, and this issue represents our continuing development in two ways.

First, we are extremely excited about the articles featured in this issue. In *Online Manipulation: Hidden Influences in a Digital World*, Daniel Susser, Beate Roessler, and Helen Nissenbaum wrestle with the definition of online manipulation and how to distinguish that harmful practice from other forms of influence. We think that this article will be essential reading for those debating platform regulation and thinking about the future of information technology. Then, in *The Global Last Mile Solution: High-Altitude Broadband Infrastructure* by Snezhana Stadnik Tapia and *Health Data at Your Fingertips: Federal Regulatory Proposals for Consumer-Generated Mobile Health Data* by Jianyan Fang, the issue confronts current topics in technology policy. Both articles offer unique solutions to difficult problems.

Second, we are proud to present the top-three papers submitted to GLTR's First Annual Writing Competition. This year's topic invited writers to consider law and policy questions related to artificial intelligence, machine learning, the use of data analytics, or algorithmic decision-making. We received over fifty submissions. Congratulations to Lauren Renaud for winning the competition with her piece, *Will You Believe It When You See It? How and Why the Press Should Prepare for Deepfakes*. Thanks to our sponsors at the BSA Foundation for helping to make this competition a reality.

In addition to these major developments, we are also pleased to present two Notes in this issue, both of which have been written by Georgetown University Law Center students and both of which concern free speech in the contemporary public squares provided by large tech platforms. These authors confront issues that have been at the forefront of technology law and policy, and we hope that you enjoy their contributions.

Last but certainly not least, we have included a selection of Technology Explainers for your reference. We hope that you find these short pieces helpful in understanding the technology of the moment.

Sincerely,
The Editorial Board
Fall 2019

GENERAL INFORMATION

Subscriptions: The *Georgetown Law Technology Review* is primarily an online law review, and as such, GLTR encourages you to visit its website at <https://georgetownlawtechreview.org/>. There, you can find information about how to order print-on-demand copies of this and other issues of GLTR. You may also download GLTR's material free of charge at that address.

Publication Timeline: As an online law review, GLTR publishes articles throughout the year to its website. Twice per year—once in the winter and once in the summer—the journal collects these articles into an issue for print publication. Upon compilation of each print issue, GLTR will post information concerning the purchase of print issues.

Submissions: GLTR accepts unsolicited materials through multiple submissions platforms as well as the email address gltr-submissions@georgetown.edu. The most current submissions guidelines can be found on GLTR's website.

Copyright: All contents of this publication are copyrighted by the *Georgetown Law Technology Review* (GLTR) except where otherwise expressly indicated. For all pieces copyrighted by GLTR, GLTR permits copies to be made for classroom use as long as proper citations are provided and notification of use is given to GLTR by emailing gltr-editor@georgetown.edu. For those pieces where another copyright holder has been expressly indicated, classroom use is likewise permitted provided that proper citations are used and the identified copyright holder is notified.

Contact: GLTR may be contacted by emailing gltr@georgetown.edu.

ADVISORY BOARD

Julie Cohen

MARK CLASTER MAMOLEN PROFESSOR OF LAW AND TECHNOLOGY
GEORGETOWN UNIVERSITY LAW CENTER

Alexandra Reeve Givens

EXECUTIVE DIRECTOR OF THE INSTITUTE FOR TECHNOLOGY LAW AND POLICY
GEORGETOWN UNIVERSITY LAW CENTER

Paul Ohm

ASSOCIATE DEAN FOR ACADEMIC AFFAIRS, PROFESSOR OF LAW
GEORGETOWN UNIVERSITY LAW CENTER

Tanina Rostain

PROFESSOR OF LAW
GEORGETOWN UNIVERSITY LAW CENTER

EDITORIAL BOARD

EDITOR-IN-CHIEF
Joshua Banker

MANAGING EDITOR
Harsimar Dhanoa

MANAGING EDITOR
Laura Hillsman

SENIOR ARTICLES
EDITOR
Aaron Scheinman

SENIOR SOLICITATIONS
EDITOR
Sarah Koslov

SENIOR NOTES
EDITOR
Sherlyn Abdullah

SENIOR
CASE COMMENTS
EDITOR
Michael Rose

SENIOR
TECHNOLOGY
EXPLAINERS EDITOR
Rachel Wehr

SENIOR
LEGAL NEWS
EDITOR
Temesgen Woldezion

DIRECTOR
OF
DEVELOPMENT
Nolan Fargo

DIRECTOR
OF
OUTREACH
Alexandra Coyle

DIRECTOR
OF
TECHNOLOGY
Avi Ginsberg

John Black
Jeffrey Brown
Raymond Coscia
Jake DeBacher
Allison Elkman

ASSISTANT EDITORS
Adam Gerchick
Shreya Kundur
Kelly McCluer
Alex Nealon
Daniel Passon
Peter Pyatigorsky

Alex Rhim
Leetal Weiss
Eric Westerhold
Ryan Whittington
Christina Wing

STAFF

Ryan Abercrombie
Niki Arakelian
Joseph Baillargeon
Daniel Barabander
Zev Beeber
Richard Bernache
Evan Burroughs
Joseph Cahill
Daniel Carlen
Jordan Cohen
Caitlyn Cook
Laura Cummings
Drew Diedrich
Andrew Do
Joseph Ehrenkrantz
Corey Fitzpatrick
Nicole Fulk
Tyler Gerstein
Andres Gonzalez
Clinton Greub
Brittany N. Griffin
Rebecca Iafrazi
Samuel Hanks
Grace Harter
Isabella Havas
John Heflin
Laura Hernandez

Tyler Kaufman
Gabriel Khoury
Gabriela Larralde
Kristen Logan
Ladan Mohaddes
Jessica Monsell
Michelle Mount
Florence Noorinejad
Jenevieve Nutovits
Sofia Panero
Josh Pereira
Sam Pickerill
Molly Rosen
Jay Schuffenhauer
David Seidman
Shelby Smith
Kendall Spencer
Lyle Stewart
Priyanka Surapaneni
Haris Vrahliotis
Yangbeini Wang
Mary Weaver
Ruiqiao Wen
Matthew Wells
May Yang
Jeff Liji Zhou
Yifan Zhu

TABLE OF CONTENTS

ARTICLES

Online Manipulation: Hidden Influences in a Digital World	1
<i>Daniel Susser, Beate Roessler, Helen Nissenbaum</i>	
The Global “Last Mile” Solution: High-Altitude Broadband Infrastructure	47
<i>Snezhana Stadnik Tapia</i>	
Health Data at Your Fingertips: Federal Regulatory Proposals for Consumer-Generated Mobile Health Data.....	125
<i>Jianyan Fang</i>	

NOTES

Incitement and the Geopolitical Influence of Facebook Content Moderation	183
<i>Sarah Koslov</i>	
<i>Red Lion Broadcasting Co. v. FCC</i> and the Rise of Speech-Enhancing Regulations of Social Media Platforms	215
<i>Connor J. Suozzo</i>	

STUDENT WRITING COMPETITION

Will You Believe It When You See It? How and Why the Press Should Prepare for Deepfakes.....	241
<i>Lauren Renaud</i>	
Getting On Board with Robots: How the Business Judgment Rule Should Apply to Artificial Intelligence Devices Serving as Members of a Corporate Board	263
<i>Thomas Belcastro</i>	
CFIUS and A.I.: Defending National Security While Allowing Foreign Investment.....	279
<i>Theodore Bruckbauer</i>	

TECHNOLOGY EXPLAINERS

Search Engine Optimization: What We See and Why We See It.....	299
<i>Joseph Baillargeon</i>	
Social Spambots.....	307
<i>Richard Bernache</i>	
Data Breaches	315
<i>Drew Diedrich</i>	
5G Wireless Connectivity: The Next Step.....	325
<i>Matthew Wells</i>	

ARTICLES

ONLINE MANIPULATION: HIDDEN INFLUENCES IN A DIGITAL WORLD

Daniel Susser, Beate Roessler, Helen Nissenbaum*

CITE AS: 4 GEO. L. TECH. REV. 1 (2019)

TABLE OF CONTENTS

I. INTRODUCTION	2
II. SOME CASES OF TROUBLING INFLUENCE.....	4
A. Targeting Advertisements at Vulnerable Teenagers.....	5
B. Algorithmically-Nudged Labor	7
C. Psychographic Profiling and Election Influence.....	9
III. DEFINING MANIPULATION	12
A. Persuasion, Coercion, and Manipulation	13
B. The Means of Manipulation.....	18
C. Deception and Manipulation.....	21
D. Nudges and Manipulation.....	22
E. Manipulation and Manipulative Practices Defined.....	26
IV. MANIPULATION THROUGH INFORMATION TECHNOLOGY	29
A. Surveillance.....	29
B. Digital Platforms.....	31
C. Mediation	33
V. HARMS OF MANIPULATION: THE SIGNIFICANCE OF AUTONOMY	34
A. Autonomy at a Glance	35
B. Autonomy and Manipulation	38
1. <i>Autonomy, Manipulation and Choice Architecture</i>	38
2. <i>Manipulation and Vulnerabilities</i>	40
C. Objections and Responses.....	41
VI. CONCLUSION.....	44

* We are very grateful for comments from participants in the 2018 Privacy Law Scholars Conference, 2018 Amsterdam Privacy Conference, 2018 Association of Internet Researchers Conference, and 2018 Information Ethics Roundtable, from faculty colloquia at MIT, San Jose State University, Penn State University, and the University of Amsterdam, and especially from Kiel Brennan-Marquez, Yafit Lev-Aretz, Paul Ohm, Marijn Sax, and Tal Zarsky. Thanks to Corinne Su for providing outstanding editorial assistance and Katherine Magruder for in-depth research assistance on our case studies. We are deeply grateful for research support from the National Science Foundation (Grants CNS-1704527 and SES-1650589) and the John D. and Katherine T. MacArthur Foundation.

I. INTRODUCTION

Privacy and surveillance scholars increasingly worry that data collectors can use the information they gather about our behaviors, preferences, interests, incomes, and so on to manipulate us.¹ Consider: investigative journalists recently discovered that Facebook allows advertisers to target vulnerable teenagers at moments when they feel “worthless” and “insecure.”² “Sharing economy” firms like the ride-hailing company Uber have explored ways to influence not only the behavior of their customers but also that of their drivers, raising concerns about potential manipulation in the workplace.³ And recent elections in the United States, the United Kingdom, Germany, France, and elsewhere have raised questions about the use of similar techniques to manipulate democratic political processes.⁴

Charges that some practices are manipulative are a strong caution, even a rallying cry to protest. But what it means, exactly, to manipulate someone and how we might systematically distinguish cases of manipulation from other forms of influence—such as persuasion and coercion—has not been thoroughly enough explored in light of the unprecedented capacities that information technologies and digital media enable. This Article endeavors to meet this challenge—to develop a definition of manipulation that addresses these enhanced capacities, to investigate how information technologies facilitate manipulative practices, and to describe the harms to individuals and social institutions that flow from such practices. We use the term “online manipulation” to highlight the particular class of manipulative practices

¹ For example, Tal Zarsky offers an early discussion of the problem in *Online Privacy, Tailoring, and Persuasion*, in *PRIVACY AND TECHNOLOGIES OF IDENTITY: A CROSS-DISCIPLINARY CONVERSATION* (Katherine J. Strandburg & Daniela Stan Raicu eds., 2006). Frank Pasquale points to it throughout his book, *THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION* (2015). Frederik Zuiderveen Borgesius provides a helpful treatment of manipulation questions, especially as they relate to European privacy law in *IMPROVING PRIVACY PROTECTION IN THE AREA OF BEHAVIOURAL TARGETING* (2015). The most far-reaching treatment is Ryan Calo’s discussion of “digital market manipulation” and consumer protection law in *Digital Market Manipulation*, 82 *GEO. WASH. L. REV.* 995 (2014).

² Sam Machkovech, *Facebook Helped Advertisers Target Teens Who Feel “Worthless,”* *ARS TECHNICA* (May 1, 2017), <https://arstechnica.com/information-technology/2017/05/facebook-helped-advertisers-target-teens-who-feel-worthless/> [<https://perma.cc/BPD9-6NKP>].

³ See generally Ryan Calo & Alex Rosenblat, *The Taking Economy: Uber, Information, and Power*, 117 *COLUM. L. REV.* 1623 (2017).

⁴ See generally Zeynep Tufekci, *Engineering the Public: Big Data, Surveillance, and Computational Politics*, 19 *FIRST MONDAY* 7 (July 7, 2014), <https://journals.uic.edu/ojs/index.php/fm/article/view/4901/4097> (charting the questions of manipulating democratic political processes) [<https://perma.cc/4XVR-KRKL>].

enabled by a broad range of information technologies. We aim to contribute to philosophical accounts of manipulation with our conceptual and normative work by clarifying the nature of manipulative practices and drawing attention to the new universe of manipulation cases that information technology raises. Beyond philosophical accounts, however, our work engages with law and policy debates and aims to guide efforts to combat the corrosive impacts of manipulation.

In Part II of the Article, we describe cases that have been cited as instances of manipulation. Then, drawing on the nature of these cases and on discussions about manipulation in moral philosophy, we present our own account of manipulation in Part III, defining manipulation and distinguishing it from neighboring terms. We argue that at its core, manipulation is hidden influence—the covert subversion of another person’s decision-making power. In contrast with persuasion, which is the forthright appeal to another person’s decision-making power, or coercion, which is the restriction of acceptable options from which another person might choose, manipulation functions by exploiting the manipulee’s cognitive (or affective) weaknesses and vulnerabilities in order to steer his or her decision-making process towards the manipulator’s ends. Manipulation, therefore, shares certain features with nudging, though we argue that only some nudges are manipulative. In addition, some have argued that manipulation is merely a species of deception.⁵ We argue that while deception is often a tool of manipulation, manipulating someone does not necessarily require instilling false beliefs.

In Part IV, we describe the particular forms manipulation takes in a world where digital technologies pervade everyday life. We argue that information technology, for a number of reasons, makes engaging in manipulative practices significantly easier, and it makes the effects of such practices potentially more deeply debilitating. First, widespread digital surveillance makes it easy for data collectors and aggregators to identify our weaknesses. The information we volunteer and shed about our interests, preferences, desires, emotional states, beliefs, habits, and so on, provides everything a would-be manipulator needs to know about how to subvert our decision-making.⁶ Second, digital platforms offer a perfect medium through which to leverage those insights. They are dynamic, interactive, intrusive, and incisively personalizable choice architectures—decision-making contexts that can be specifically designed to adapt to and exploit each individual user’s

⁵ For example, see our discussion of Robert Goodin’s theory of manipulation, *infra* Section II.B.

⁶ Karen Yeung describes the problem in terms of “hypernudging.” Karen Yeung, *Hypernudge: Big Data as a Mode of Regulation by Design*, 20 INFO. COMM. & SOC’Y 118 (2017).

particular vulnerabilities.⁷ Finally, the reach of digital tools is enormous. Because digital interfaces mediate so much of so many people's lives, they have the potential to affect far more people far more deeply than their analogue counterparts. Social media services, like Facebook, with millions or even billions of users can be leveraged as tools of massive and hyper-targeted manipulation.

In Part V, we turn to the harms of online manipulation. Subverting another person's decision-making power undermines his or her autonomy. Given that respect for individual autonomy is a bedrock principle of liberal democracy, the potential for massive online manipulation is a cause for grave concern. We flesh out the notion of "vulnerability," mapping its various types and the ways they are exploited as well as the places where different kinds of vulnerabilities interact to reinforce or exacerbate one another. We show how altering people's choice architecture to exploit their vulnerabilities affects them. And we look to philosophical accounts of autonomy—specifically to accounts of socially-situated, relational autonomy—to bring these strands together, showing where and how manipulative influences thwart people's capacity to form decisions they can recognize and endorse as their own.

We conclude by considering directions for future research, suggesting that it is especially important to consider the effects of online manipulation in different social contexts. Although we might be willing to tolerate some amount of external influence on our decision-making processes in consumer contexts, we likely want less interference in political contexts. We argue that the role autonomy plays in political decision-making is more fundamental and consequential than it is in consumer decision-making. Threats to the former therefore demand a more vigorous response. As we work to combat the new forms of manipulative practice made possible by information technology, we will need to distinguish carefully among the contexts in which they operate.

II. SOME CASES OF TROUBLING INFLUENCE

To provide concrete vehicles for our analysis, we introduce a few well-known cases to our discussion. First, we consider targeted advertising in the commercial sphere, examining worrying reports that Facebook has the ability to target advertisements at teenagers during moments they are perceived to be especially vulnerable to influence. Second, we consider cases of so-called "algorithmic management"—strategies that gig platforms such as Uber use to influence worker behavior.⁸ Some of these strategies may verge on

⁷ See generally Marjolein Lanzing, "Strongly Recommended" Revisiting Decisional Privacy to Judge Hypernudging in Self-Tracking Technologies, 32 PHIL. & TECH. 549 (2019).

⁸ See generally Alex Rosenblat & Luke Stark, *Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers*, 10 INT'L J. COMM. 3758 (2016).

manipulative, raising difficult questions about appropriate influence in the workplace. Third, we consider Cambridge Analytica and claims that the political advertising firm engaged in voter manipulation. Together, these three cases exemplify concerns about online manipulation across social spheres—in the market, in the workplace, and in the realm of politics.

A. Targeting Advertisements at Vulnerable Teenagers

Consumer advertising has long been contested ethical terrain. Its defenders argue that advertising plays an important role in an ideal competitive free market, informing consumers about products so they may effectively select among the alternatives and thus ensuring supply meets demand and prices adjust accordingly.⁹ Critics see the opposite. They identify trends in advertising that are increasingly aimed at provoking action against reason or circumventing reason altogether. Vance Packard famously charged the advertising industry with utilizing “motivation analysis,” psychological and psychoanalytical means to exploit “hidden weaknesses and frailties,” to appeal to non-rational and subconscious mental processes in service of marketing ends.¹⁰ And, of course, some stake a middle ground by suggesting that advertising can serve a useful function even though certain forms and outcomes are deeply problematic.¹¹ The answer is neither to abandon nor outlaw but to divine criteria to determine when advertising is performing a useful service (such as informing consumers of and about products, promoting an active marketplace, or even promoting brand loyalty) and when it is not.¹²

With the commercialization of the Internet, it was inevitable that advertising would migrate online, and alongside it, detractors who regretted any such incursions.¹³ “Sponsored search,” which delivers advertisements tailored to specific search terms, was initially the norm and was developed by Web search services that could easily perform this match.¹⁴ This was quickly followed by advertising targeted at individuals, with DoubleClick (now owned

⁹ See, e.g., JERRY KIRKPATRICK, IN DEFENSE OF ADVERTISING: ARGUMENTS FROM REASON, ETHICAL EGOISM, AND LAISSEZ-FAIRE CAPITALISM (1994).

¹⁰ See generally VANCE PACKARD, THE HIDDEN PERSUADERS (1957).

¹¹ Some in the advertising industry take this position, recognizing that there is such a thing as unethical advertising but rejecting the notion that all advertising is unethical. See, e.g., Wallace S. Snyder, *Ethics in Advertising: The Players, the Rules, and the Scorecard*, 22 BUS. & PROF. ETHICS J. 1 (2003).

¹² See, e.g., MICHAEL SCHUDSON, ADVERTISING, THE UNEASY PERSUASION: ITS DUBIOUS IMPACT ON AMERICAN SOCIETY (Routledge 2013) (1984).

¹³ For a historical account of the emergence of online advertising, see generally TIM WU, THE ATTENTION MERCHANTS: THE EPIC SCRAMBLE TO GET INSIDE OUR HEADS (2016).

¹⁴ See generally Bernard J. Jansen & Tracy Mullen, *Sponsored Search: An Overview of the Concept, History, and Technology*, 6 INT’L J. ELEC. BUS. 2 (2008).

by Google) taking advantage of the online advertising industry's victory in the Internet Engineering Task Force (IETF) debate over the Web cookie standard, which allowed third-parties to set and retrieve cookies across numerous sites.¹⁵ Who these third parties are and how they utilize the intelligence they glean from user surveillance is not exactly known, but the dogged research and publication of academics, activists, and the popular press has gone a long way to expose both technical methods and a dizzying array of commercial actors, including those in direct contact with users and those behind the scenes that now populate this space.

A recent case that captured much public attention offers a hint at how these practices are evolving. In May 2017, an Australian newspaper reported that it had obtained a leaked internal Facebook strategy document describing how advertisers could use the company's platform to target advertisements at teenagers as young as fourteen years old at moments when they feel vulnerable. "By monitoring posts, pictures, interactions and internet activity in real-time, Facebook can work out when young people feel 'stressed', 'defeated', 'overwhelmed', 'anxious', 'nervous', 'stupid', 'silly', 'useless', and a 'failure,'" the report claims.¹⁶ Facebook responded that the report is misleading, and the features described are simply meant to "help marketers understand how people express themselves on Facebook," not to target ads.¹⁷ They did not deny, however, that their insights into teenagers' emotional states *could* be used to influence the vulnerable, leaving many to wonder if all that stands between us and this kind of purported manipulation is Facebook's company policies.¹⁸

The marriage of advertising and information technology has thus rendered urgent questions about what Cass Sunstein calls the "ethics of influence."¹⁹ There is a growing sense that the ways of influencing consumers

¹⁵ See generally David M. Kristol, *HTTP Cookies: Standards, Privacy, and Politics*, 1 ACM TRANSACTIONS ON INTERNET TECH. 151 (2001); John Schwartz, *Giving Web a Memory Cost Its Users Privacy*, N.Y. TIMES (Sept. 4, 2001), <https://www.nytimes.com/2001/09/04/business/giving-web-a-memory-cost-its-users-privacy.html> [<https://perma.cc/8S8L-MPTC>].

¹⁶ Darren Davidson, *Facebook Targets 'Insecure' Young People to Sell Ads*, AUSTRALIAN (May 1, 2017), <https://www.theaustralian.com.au/business/media/facebook-targets-insecure-young-people-to-sell-ads/news-story/a89949ad016eee7d7a61c3c30c909fa6/> [[perma.cc link unavailable](https://perma.cc/link-unavailable)].

¹⁷ *Comments on Research and Ad Targeting*, FACEBOOK NEWSROOM (Apr. 30, 2017), <https://newsroom.fb.com/news/h/comments-on-research-and-ad-targeting/> [<https://perma.cc/K3RM-NBMU>].

¹⁸ See Nitasha Tiku, *Get Ready for the Next Big Privacy Backlash Against Facebook*, WIRED (May 21, 2017), <https://www.wired.com/2017/05/welcome-next-phase-facebook-backlash/> [<https://perma.cc/REZ9-QDEA>].

¹⁹ See CASS SUNSTEIN, *THE ETHICS OF INFLUENCE: GOVERNMENT IN THE AGE OF BEHAVIORAL SCIENCE* (2016).

we once (perhaps reluctantly) tolerated have become intolerable—that what some call “persuasive” strategies are in fact manipulative. Ryan Calo claims that “digital market manipulation is a problem, if at all, because it constitutes a form of persuasion that is dangerous to consumers or society.”²⁰ Karen Yeung points to the “persuasive, manipulative qualities” of nudges driven by big data.²¹ Anthony Nadler and Lee McGuigan suggest that “persuasive communication can manipulate consumer attitudes and behaviors.”²² Tal Zarsky argues that online surveillance “might facilitate the manipulation of subjects.”²³

We applaud these efforts to bring renewed scrutiny to consumer advertising practices in light of the powerful digital socio-technical systems that now convey them. What is lacking in this emerging literature, however, is conceptual clarity about what exactly online manipulation is, how it differs from other forms of influence, and how, normatively, it ought to be addressed. These are the questions we take up below.

B. Algorithmically-Nudged Labor

On April 21, 2017, Noam Scheiber wrote for *The New York Times*:

And yet even as Uber talks up its determination to treat drivers more humanely, it is engaged in an extraordinary behind-the-scenes experiment in behavioral science to manipulate them in the service of its corporate growth—an effort whose dimensions become evident in interviews with several dozen current and former Uber officials, drivers and social scientists, as well as a review of behavioral research.²⁴

Drawing on academic research, Scheiber cites several practices supporting these assertions. The backdrop is Uber’s insistence that drivers are not employees. Instead, they are independent contractors taking advantage of Uber’s software platform, which connects drivers with people needing rides and for which Uber charges commission.²⁵ Although both Uber and its drivers

²⁰ Calo, *supra* note 1, at 1020.

²¹ Yeung, *supra* note 6, at 119.

²² Anthony Nadler & Lee McGuigan, *An Impulse to Exploit: The Behavioral Turn in Data-Driven Marketing*, 35 CRITICAL STUD. MEDIA COMM. 151, 161 (2018).

²³ Tal Z. Zarsky, *Privacy and Manipulation in the Digital Age*, 20 THEORETICAL INQUIRIES L. 157, 158 (2019).

²⁴ Noam Scheiber, *How Uber Uses Psychological Tricks to Push Its Drivers’ Buttons*, N.Y. TIMES (Apr. 21, 2017), <https://www.nytimes.com/interactive/2017/04/02/technology/uber-drivers-psychological-tricks.html> [<https://perma.cc/4UU8-YUV6>].

²⁵ *Id.*

make money through rides, incentives are not neatly aligned. To compete successfully with other companies, generally, Uber must serve riders within the shortest timeframe and at the lowest cost. Purportedly lacking the coercive clout of an employer, Uber has devised a range of features to encourage driver behaviors that are profitable for the company, though, some have argued, not optimal for drivers.

Because conditions of high rider demand and low driver supply yield price surges, they are favored by drivers. Uber's interests, however, are best served by many drivers serving as many riders as possible, with prices that beat those of their competitors. Accordingly, Uber barrages drivers with texts, emails, popups, and carefully designed graphics to keep them behind the wheel and to direct them, ostensibly, to areas of highest demand. One of the techniques cited by critics is the carefully curated graphic representations of predicted needs.²⁶ For example, the inclusion of sporting events or bar-closing times hint that both demand and the likelihood of surge pricing will be high.²⁷ In other words, drivers are presented with vague promises and enticed by heat map estimates, which conflate real-time and predictive demand. They are thus nudged toward performing a service for highly uncertain rewards in a manner Ryan Calo and Alex Rosenblat liken to a "bait and switch."²⁸

A second feature is to urge drivers to continue working as they reach the end of a shift and try to log out of the system. They may receive push notifications reading, "Are you sure you want to go offline? Demand is very high in your area. Make more money, don't stop now!" and accompanied by a surge icon.²⁹ Or the app may indicate that they are approaching some arbitrary earnings level for that shift, e.g., "you're \$10 away from making \$330!"³⁰ These are attempts to exploit a well-known decision-making vulnerability—"people's preoccupation with goals—to nudge them into driving longer"—evidenced in robust findings from behavior research.³¹ Similar effects are imputed to the gamification of Uber's interactive app, which shows work status (hours, earnings, rides, etc.) in game-like formats known for their power to hold players at the game console, and, presumably, drivers at the wheel. And Uber utilizes automatic queuing, a strategy familiar to those who subscribe to streaming services such as Netflix or Amazon Video. Before a ride ends Uber cues up the next ride request, making it just a bit more

²⁶ Calo & Rosenblat, *supra* note 3, at 1662.

²⁷ Rosenblat & Stark, *supra* note 8, at 3769 ("We also want to remind you that we predict that New Year's Eve will be the busiest time of the year. *With such high demand, it will be a great night to go out and drive!*").

²⁸ Calo & Rosenblat, *supra* note 3, at 1662.

²⁹ Rosenblat & Stark, *supra* note 8, at 3768.

³⁰ Scheiber, *supra* note, 24.

³¹ *Id.*

difficult to refuse than to take on another ride.³² Although Uber responded to concerns by allowing automatic queuing to be disengaged, the feature is engaged by default and reloads after breaks.³³ Uber is not alone in employing these practices.³⁴

In describing Uber's practices, critics regularly cite to behavioral science research that reveals human vulnerability to forms of irrational, biased, and bounded thinking, as well as tendencies toward compulsive and even addictive ruts.³⁵ These cognitive vulnerabilities range from well-studied phenomena,³⁶ such as loss aversion and preoccupation with goals, to what Natasha Schüll, in her study of addictive gambling machines, calls "ludic loop," a compulsion to keep playing.³⁷ It even extends to the practice of Uber employees taking on female personas based on experimental findings that suggest the majority male driver population is more likely to engage with messages emanating from female rather than male communicators.³⁸ While these critics have raised a number of ethical concerns, including coercion, excessive control through punitive measures, exploitation, and power and information asymmetries, importantly for this discussion they also claim these practices can be manipulative, raising deep ethical questions about which forms of influence are appropriate in the workplace.³⁹

C. Psychographic Profiling and Election Influence

In March 2018, sudden public interest arose in whether or not the political services firm, Cambridge Analytica, had improperly used Facebook's advertising platform to exert influence in the 2016 U.S. presidential election. While many questions about the case remain unanswered, many commentators have pointed to the possibility that the purported influences were manipulative.

Piecing together a story from public media, and conceding that it includes contradictions and denials, it appears that Cambridge Analytica

³² *Id.* ("It requires very little effort to binge on Netflix; in fact, it takes more effort to stop than to keep going.")

³³ *Id.*

³⁴ *Id.*; see also Tae Wan Kim & Kevin Werbach, *More Than Just a Game: Ethical Issues in Gamification*, 18 ETHICS & INFO. TECH. 157, 157–73 (2016).

³⁵ See, e.g., Calo & Rosenblat, *supra* note 3; Scheiber *supra* note 24; Luke Stark, *Algorithmic Psychometrics and the Scalable Subject*, 48 SOC. STUD. SCI. 204 (2018).

³⁶ See RICHARD H. THALER & CASS R. SUNSTEIN, *NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS* (2008).

³⁷ NATASHA DOW SCHÜLL, *ADDICTION BY DESIGN: MACHINE GAMBLING IN LAS VEGAS* (2014).

³⁸ Scheiber, *supra* note 24.

³⁹ *Id.*

employed a company called Global Science Research (GSR) to generate vast repositories of digital user profiles. These repositories were initially seeded by the results of a personality quiz, “thisisyourdigitallife” (2014), administered by Aleksandr Kogan, a lecturer in the Department of Psychology at the University of Cambridge and the head of GSR.⁴⁰ Kogan distributed the quiz through a Facebook app after being refused access to a dataset Michael Kosinski had assembled with colleagues at Cambridge and Microsoft Research for their widely cited study published in the *Proceedings of the National Academy of Science*.⁴¹ Subjects, limited to U.S. voters with Facebook accounts, each received a few dollars for taking the quiz and providing access to their Facebook accounts. From the few hundred thousand quiz-takers, Cambridge Analytica accumulated tens of millions of Facebook user accounts through a feature (no longer active) that allowed developers to gain access to the accounts of “friends” of quiz-takers—a number possibly as high as 87 million.⁴²

We can only surmise that Kogan’s work with Cambridge Analytica was inspired by the Kosinski, et al. study, which demonstrated that a tremendous amount can be inferred about individuals from Facebook “likes” alone. Such inferences include gender, sexual orientation, race, religion, political views, relationship status, substance use, and size and density of friendship networks.⁴³ The study further claimed to uncover correlations between “like” patterns, psychological traits (such as intelligence as measured by Raven’s Standard Progressive Matrices), and personality profiles (such as openness, agreeableness, emotional stability, and conscientiousness as measured by the International Personality Item Pool).

This provides context for the claims made by Christopher Wylie, one of Cambridge Analytica’s co-founders-turned-whistleblower: “We exploited Facebook to harvest millions of people’s profiles and built models to exploit what we knew about them and target their inner demons. That was the basis

⁴⁰ Carole Cadwalladr & Emma Graham-Harrison, *Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach*, GUARDIAN (Mar. 17, 2018), <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> [<https://perma.cc/X58U-SJ8Z>].

⁴¹ Michael Kosinski et al., *Private Traits and Attributes are Predictable from Digital Records of Human Behavior*, 110 PROC. NAT’L ACAD. SCI. 5802 (2013).

⁴² Cecilia Kang & Sheera Frankel, *Facebook Says Cambridge Analytica Harvested Data of Up to 87 Million Users*, N.Y. TIMES (Apr. 4, 2018), <https://www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html> [<https://perma.cc/VFZ8-3W2Q>].

⁴³ Kosinski et al., *supra* note 41.

the entire company was built on.”⁴⁴ According to news reports, Cambridge Analytica claims to have built “psychographic” profiles on approximately 220 million U.S. voters based on 5,000 pieces of data.⁴⁵ In a 2016 speech, then-CEO Alexander Nix boasted about Cambridge Analytica’s ability to personalize messages in a range of media from direct mail, to online cookie-driven ad targeting, to social media banners, and even to set-top televisions.⁴⁶ Cambridge Analytica’s personalized approach was different from past advertising, Nix claimed, because “we don’t need to guess at what creative solution may or may not work. We can use hundreds or thousands of individual data points on our target audiences to understand exactly which messages are going to appeal to which audiences.”⁴⁷ Cambridge Analytica’s strategies also differed from traditional mass advertising—even advertising that is targeted at defined demographic groups, such as age, race, socioeconomic class, etc. “[W]e’ve rolled out a long form quantitative instrument to probe the underlying traits that inform personality,” Nix claimed, which measures:

openness—how open you are to new experiences; conscientiousness—whether you prefer order and habits and planning in your life; extraversion—how social you are; agreeableness—whether you put other people’s needs and society and community ahead of yourself; and finally neuroticism—a measurement of how much you tend to worry. By having hundreds and hundreds of thousands of Americans undertake this survey, we were able to form a model to predict the personality of every single adult in the United States of America.⁴⁸

Compiling profiles is one thing. But even assuming one accepts the premises upon which claims about these profiles are based—that personality tests, augmented with inferences based on online measures, such as Facebook “likes,” produce sound profiles—it is another thing to go from psychographic

⁴⁴ *Cambridge Analytica and Facebook: The Scandal So Far*, AL JAZEERA (Mar. 28, 2018), <https://www.aljazeera.com/news/2018/03/cambridge-analytica-facebook-scandal-180327172353667.html> [https://perma.cc/PU4C-KNSN].

⁴⁵ Carole Cadwalladr, *Google, Democracy and the Truth About Internet Search*, GUARDIAN (Dec. 4, 2016), <https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook> [https://perma.cc/5ZJ4-GAWV].

⁴⁶ Alexander Nix, *Cambridge Analytica—The Power of Big Data and Psychographics*, Presentation at the 2016 Concordia Summit (Sept. 27, 2016), in YOUTUBE, <https://www.youtube.com/watch?v=n8Dd5aVXLcC> [https://perma.cc/N9C6-HTMJ].

⁴⁷ *Id.*

⁴⁸ *Id.*

profiles to targeted messaging. This step requires tailoring messages to correspond to the specific personality traits of their recipients. Matz et al. claim to have shown experimentally that “targeting people with persuasive appeals tailored to their psychological profiles can be used to influence their behavior as measured by clicks and conversions.”⁴⁹ The business literature is filled with studies that claim to “prove” the effectiveness of conventional ad targeting, and there are possibly industry studies, not revealed publicly, showing that targeting works. As to pinpointing advertisement recipients, we learned that Facebook, while it does not allow marketers to target advertisements based on psychological traits directly, “it does so indirectly by offering the possibility to target users based on their Facebook Likes.”⁵⁰ In a 2018 paper, computer scientists Irfan Faizullahoy and Aleksandra Korolova demonstrated that one can get around technical enforcement of policies disallowing overly narrow targeting and successfully target particular messages down to the individual.⁵¹

In any event, the popular embrace of claims about the efficacy of these methods in the broadest terms very likely stirs its continuing appeal even in the absence of hard evidence. For the purposes of this article, we do not need to demonstrate that Cambridge Analytica’s efforts had a deep and meaningful impact on the U.S. presidential election. While it is unlikely that its effects were dispositive, that they worked, at some level, seems incontrovertible. More importantly, there is every reason to believe efforts like these will continue to evolve, and worries about online manipulation in political contexts will continue to grow.⁵² Our goal, then, is to reveal the dimensions of these practices that make them manipulative, and in so doing, expose why they are deeply disturbing.

III. DEFINING MANIPULATION

Manipulation is a tricky term, much like the behavior it describes. Colloquially, to manipulate something is to steer or control it. One often speaks of manipulating complex technical instruments, devices that would do nothing at all without human direction. For example, cars are manipulated via steering wheels and pedals, and computers are manipulated via keyboards and

⁴⁹ S.C. Matz et al., *Psychological Targeting as an Effective Approach to Digital Mass Persuasion*, 114 PROC. NAT’L ACAD. SCI. U.S. 12714, 12715 (2017).

⁵⁰ *Id.*

⁵¹ Irfan Faizullahoy & Aleksandra Korolova, *Facebook’s Advertising Platform: New Attack Vectors and the Need for Interventions*, (Workshop on Technology & Consumer Protection, arXiv:1803.10099, 2018), <https://arxiv.org/abs/1803.10099> [<https://perma.cc/39Z8-7P45>].

⁵² For a sense of the broad contours of debates about voter microtargeting, see Frederick J. Zuiderveen Borgesius et al., *Online Political Microtargeting: Promises and Threats for Democracy*, 14 UTRECHT L. REV. 82 (2018).

mouse devices. Similarly, one can manipulate living, animate things by changing the way they would behave absent any interventions. Gardeners manipulate tomato plants by fastening them to a trellis, and trainers manipulate dogs by enticing them with treats. To manipulate a *person* is likewise to steer or control them, as though they were a car or a tomato plant.⁵³ Manipulators are often described as “puppet masters” who pull their targets’ strings.

It is worth distinguishing at the outset between manipulation understood in this way—as steering or controlling a person—and the manipulation of institutions or systems. Given the recent rise of online disinformation campaigns, especially targeted at voters, there is growing concern about what some have termed “media manipulation” and its effects on election outcomes.⁵⁴ Obviously, the ultimate goal of influencing the media is to influence the people consuming it. Likewise, influencing people can be a means of altering the institutions they participate in—for example, when they vote. At the heart of these worries, though, are concerns about individuals and the independence of their decision-making processes. It is this sense of manipulation—as influence over individuals—that occupies us here.

Manipulation, then, is a kind of influence—an attempt to change the way someone would behave absent the manipulator’s interventions. The question is what kind of influence it is. In the following section, we show how manipulation differs from related concepts. We consider what distinguishes manipulation from persuasion, as well as other familiar forms of influence—coercion and deception. In general, persuasion is thought to be perfectly acceptable, while deception and coercion are not. If manipulation is like persuasion in relevant ways, then we might decide it is not worth worrying about. If it is like deception or coercion, we might worry indeed. Finally, we consider how manipulation differs from nudging, a form of influence about which there is considerably less agreement.

A. Persuasion, Coercion, and Manipulation

“Persuasion” has both broad and narrow meanings. In the broad sense, to persuade simply means to change someone’s mind—it is an umbrella term

⁵³ As Allen Wood argues, “The manipulative person ‘steers’ the other as a driver steers an automobile. The automobile is already moving through its own internal combustion engine and momentum, but its direction is influenced by the one who steers it.” Allen Wood, *Coercion, Manipulation, Exploitation*, in *MANIPULATION: THEORY AND PRACTICE* 17, 33–34 (Christian Coons & Michael Weber eds., 2014).

⁵⁴ See generally Yochai Benkler et al., *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (2018); Alice Marwick & Rebecca Lewis, *Media Manipulation and Disinformation Online*, DATA & SOC’Y RES. INST. (May 15, 2017), https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf [<https://perma.cc/H4TY-446Q>].

covering nearly all forms of influence, from argumentation and rhetoric to pointing a gun at someone's head. In this sense, persuasion includes both making someone an offer and making someone *an offer they can't refuse*. By contrast, persuasion in the narrow sense (i.e., rational persuasion) means changing someone's mind by giving reasons he or she can reflect on and evaluate. Distinguishing between these two meanings of the term is important because they carry significantly different normative connotations. Persuasion in the narrow sense is generally thought to be perfectly morally acceptable, whereas persuasion in the broad sense denotes both behaviors we accept and behaviors we do not accept.

Many authors writing about these issues use the term "persuasion" in the broad sense. Conceptually, there is nothing wrong with that. Normatively, however, in discussions about manipulation, it muddies the waters, raising questions about how and why a good thing (persuasion) becomes a bad thing (manipulation). For this reason, throughout this Article we use "persuasion" in the narrow sense, meaning rational persuasion, and the generic term "influence"—a term without normative baggage—to describe the category of which rational persuasion and manipulation are both members.

At an abstract level, there are two ways of influencing another person's decision-making: change the options available to the other person (their decision space) or change how they understand their options (their internal decision-making process).⁵⁵ Persuasion, the paradigmatic form of respectful influence, works by operating either of these two levers. When we persuade someone to do something (or to refrain from doing it), we appeal openly to their capacity for conscious deliberation and choice. We offer arguments or incentives. We assume, in other words, that they pursue their own ends, and we try to demonstrate that doing things our way would advance them. Take, for instance, the manager of a car dealership. To persuade her sales team to work longer hours, she might argue that because shoppers stay out later during long summer days, working late during the summer would increase their odds of earning extra commissions. In this way, she appeals to her team members' own ends—earning an income—and tries to show them that behaving the way she wants is a means to achieving those ends. She attempts to influence her team by changing the way they understand their options—i.e., by appealing to their internal decision-making process.

If her argument fails, the manager can shift strategies. Instead of offering an argument, she might offer an incentive, such as overtime pay. Rather than change how her team members understand their situation, the manager changes the situation itself. She alters their decision-space by reconfiguring the options available to them. This is equally a form of

⁵⁵ We will complicate this dichotomy later. See *infra* Section V.B.

persuasion. In either case, the attempted influence is forthright and transparent, and the decision is ultimately left entirely up to the staff. The manager successfully influences her team's behavior only when she *convinces* them to behave the way she wants, motivating them with abstract arguments or concrete incentives.

This example highlights what Joel Rudinow would call “resistible incentives.”⁵⁶ Although the incentives change the terms being deliberated over, the change is not determinative; it is easy to imagine some of the team members deciding, on balance, that the overtime pay is not worth it. In that case, the manager has yet another option: she can coerce them. To coerce someone is to offer “*irresistible* incentives.”⁵⁷ Or, to put the same point slightly differently, as Allen Wood does, coercing someone means eliminating all of the “acceptable alternatives.”⁵⁸ The manager could threaten her team members’ jobs, or she could put a gun to their heads. She could, in other words, restrict her team members’ decision-spaces, arranging them such that they only have one acceptable option—behaving the way the manager wants.⁵⁹ If we face irresistible incentives (or lack acceptable alternatives) we are *forced* to act as our coercers would have us act; we are deprived of choice.

Persuading someone leaves the choice of the matter entirely up to them, while coercing someone robs them of choice. At the same time, although coercing someone deprives them of choice, in an important sense, it leaves their capacity for conscious decision-making intact. After all, recognizing that some incentive is irresistible, or that an alternative is unacceptable, requires having our wits about us. If one did not understand that the only acceptable option available to them was to do as their coercer instructed, or if they could not act on that understanding, then they would have no motivation or no means to go along with the coercer’s plan. Coercing someone forces them to act the

⁵⁶ Joel Rudinow, *Manipulation*, 88 ETHICS 338, 342 (1978).

⁵⁷ An irresistible incentive, Rudinow argues, is one which could only be avoided through “heroism, madness, or something similarly extraordinary.” *Id.* at 341.

⁵⁸ Wood, *supra* note 53, at 21–23.

⁵⁹ What counts as irresistible or unacceptable might differ somewhat from person to person (the threat of being fired might be resistible/acceptable for someone with other job prospects, but irresistible/unacceptable for someone without them). Which is not to say that they are determined by the agent’s own beliefs or feelings about them. Rather, it is to say that what counts as an acceptable or unacceptable option can be determined, in part, by context. *Id.* As Wood writes, “there is sometimes an objective fact of the matter that certain alternatives are (or are not) acceptable to a given agent *under specific circumstances.*” *Id.* (emphasis added).

way the coercer wants, not by undermining or circumventing their decision-making faculties, but by making the coercer's way the only acceptable one.⁶⁰

In one respect, persuasion and coercion are, therefore, opposites, since persuading someone leaves the choice of the matter entirely up to the target, while coercing someone deprives them of choice. But what persuasion and coercion have in common is that they attempt to influence without undermining the target's decision-making powers. In cases of persuasion and coercion, the agent is steering the ship. When coerced, a person is forced to abandon their self-chosen ends (the destination, say), but it is still the coerced person who does the abandoning. They understand what is happening; they recognize it as the only acceptable option. They direct themselves to do as their coercer demands. If asked later what happened, why they acted the way that they did, the coerced person has no trouble explaining it. "My hands were tied," they say, regretfully, "It was the only available option. I was forced to act against my will." Or perhaps not: rare though they may be, there are occasions when a person with a gun to their head defies their would-be coercer anyway. There are heroic stories of German gentiles, under threat of death, refusing to give up their Jewish neighbors to the Nazis. That such a thing is even possible demonstrates that no matter how forceful the attempted coercion, it is ultimately the coerced person who makes the decision to act.

Manipulation, by contrast, means taking hold of the controls. To manipulate people is to displace them as the decider, to "subvert," as Wood puts it, their capacity for self-government, to "undermine or disrupt the ways of choosing that they themselves would critically endorse if they considered the matter in a way that is lucid and free of error."⁶¹ It is to deprive them of

⁶⁰ Christian Coons and Michael Weber write: "the instruments of coercion (threats, incarceration, and other penalties) are attempts to alter the context of choice, making it rational for you to comply. In this way, the coercer typically treats the coerced as rational. In fact, coercion depends on the target's being rational." Christian Coons & Michael Weber, *Introduction: Investigating the Core Concept and Its Moral Status*, in *MANIPULATION: THEORY AND PRACTICE* 1, 15 (Christian Coons & Michael Weber eds., 2014). We agree. Note in the next sections, though, that ultimately, we do not argue that manipulation subverts *rationality*; rather, we argue that it subverts conscious, self-aware decision-making.

⁶¹ We should note that we only follow Wood part of the way. For Wood, self-government is coextensive with *rational* decision-making—to self-govern is to decide rationally. Wood, *supra* note 53, at 35 ("What is characteristic of manipulative behavior is that it influences people's choices in ways that circumvent or subvert their rational decision-making processes, and that undermine or disrupt the ways of choosing that they themselves would critically endorse if they considered the matter in a way that is lucid and free of error."). We do not tie self-government or its disruption so closely to rationality. It is possible, on our account, to be irrationally influenced without being manipulated, just as it is possible to be manipulated and still decide rationally. The salient issue, to our minds, is not whether the influence appeals to the target's rational faculties, but, rather, whether it appeals to their *conscious* decision-making process. Which is to say, the issue for us is not rationality but *awareness*.

authorship over their actions.⁶² That is what it means to feel like someone else's puppet: when a person is coerced that person feels used, when a person is manipulated that person feels *played*.

Of course, manipulation is rarely so thorough as to *totally* deprive its target of self-control or self-government. That is why many people intuitively believe that manipulation involves less control than coercion (and consequently, that people should almost always be excused for doing things they were coerced to do, but only sometimes be excused for things they were manipulated into doing). Rather than entirely displacing the target as the decision-maker, the manipulator insinuates himself in his target's decision-making process. In Sarah Buss's words, he interferes with "the self-governed (and self-governing) activity we call 'making up one's own mind about how to act.'"⁶³ To say that you "feel manipulated" is to say that you do not fully understand why you acted the way that you did, or whether your actions served your own or someone else's ends.

Whereas persuasion and coercion work by appealing to the target's capacity for conscious decision-making, manipulation attempts to subvert that capacity. It neither convinces the target (leaving all options open) nor compels the target (eliminating all options but one). Instead, it interferes with the target's decision-making process in order to steer them toward the manipulator's ends. Importantly, this is not meant to suggest that coercion is acceptable or less bad than manipulation, but that coercion and manipulation threaten an individual's capacity to choose and pursue their own ends in different ways. Coercion is blunt and forthright: one almost always *knows* one is being coerced. Manipulation is subtle and sneaky. Rather than simply depriving a person of options as the coercer does, the manipulator infiltrates their decision-making process, disposing it to the manipulator's ends, which may or may not match their own.

⁶² Or more precisely, it deprives them of *part* authorship. In thinking about how we influence others, it is tempting to use as one's baseline an imaginary *uninfluenced* person—a perfectly unencumbered decision-maker. This person is the imagined subject of much liberal social and political discourse. She is rational, deliberative, and fully autonomous, an individual in every sense of the word. She is the author of her own life story. It is important to recognize that this subject is a fiction. Although we are individuals, we are also social beings who exist inexorably with others, and the people we live and interact with affect us. Thus, as Joseph Raz argues, "an autonomous person is *part* author of his own life. His life is, *in part*, of his own making." JOSEPH RAZ, *THE MORALITY OF FREEDOM* 204 (1986) (emphasis added). Autonomy is not the absence of influence, but the presence of self-government. To be autonomous—to be part-author of one's own life—is to know that one's decision-making is conditioned, and yet, still, to take one's own reasons for acting as authoritative. It is this authority of one's own reasons that manipulation subverts, as we discuss in Part IV, below.

⁶³ Sarah Buss, *Valuing Autonomy and Respecting Persons: Manipulation, Seduction, and the Basis of Moral Constraints*, 115 *ETHICS* 195, 195 (2005).

B. The Means of Manipulation

The question, then, is how this is done—how are manipulators able to alienate their targets from their own decision-making powers and interfere with the way they make up their minds about how to act? Philosophers and political theorists have described a wide variety of purportedly manipulative techniques. Marcia Baron points to lies, false promises, applying pressure, and playing on people’s emotions.⁶⁴ Wood adds encouraging false assumptions, fostering self-deception, and appealing to “character flaws.”⁶⁵ Kate Manne discusses guilt trips.⁶⁶ Buss considers certain kinds of seduction.⁶⁷

The most systematic account is Robert Noggle’s influential theory that one manipulates another by causing their beliefs, desires, or emotions to deviate from certain ideals.⁶⁸ Beliefs, for instance, are ideally true. Desires, according to Noggle, are ideally directed toward things we have reason to want. Emotions are ideally appropriate to the situation at hand. When we make decisions according to right-functioning rational decision-making processes, our beliefs, desires, and emotions approximate these ideals.⁶⁹ Manipulating someone, Noggle argues, means corrupting rational decision-making by “adjusting” these “psychological levers”—belief, desire, and emotion—away from their ideal settings. One might deceive their target (causing them to have false beliefs), tempt them (creating a desire for what they lack reason to want), incite them (causing an inappropriate emotional response), and so on.

While there is much to like about this story, it fails to capture what is distinctive about manipulation—that it undermines our sense of authorship over our decisions. Causing someone to make non-ideal decisions is sometimes manipulative, but it is not manipulative in every case. Suppose you are a great lover of martinis but really ought not to drink. And knowing this, but not wanting to feel like lushes themselves, your friends parade the finest martinis before you, tempting you to have one. In this case your friends have

⁶⁴ See Marcia Baron, *Manipulateness*, 77 PROC. & ADDRESSES AM. PHIL. ASS’N 37 (2003).

⁶⁵ See Wood, *supra* note 53, at 35.

⁶⁶ See generally Kate Manne, *Non-Machiavellian Manipulation and the Opacity of Motive*, in MANIPULATION: THEORY AND PRACTICE 221 (Christian Coons & Michael Weber eds., 2014).

⁶⁷ See Buss, *supra* note 63, at 195.

⁶⁸ Which ideals exactly are relevant here is up for some debate. For Noggle, manipulation involves inducing someone to deviate from what the manipulator believes are the ideals for beliefs, desires, and emotions. See Robert Noggle, *Manipulative Actions: A Conceptual and Moral Analysis*, 33 AM. PHIL. Q. 43, 47–48 (1996). Anne Barnhill endorses all of Noggle’s account save this. In Barnhill’s view, the relevant ideals are those objectively in the target’s self-interest. Anne Barnhill, *What is Manipulation?*, in MANIPULATION: THEORY AND PRACTICE 51, 65–72 (Christian Coons & Michael Weber eds., 2014).

⁶⁹ Noggle, *supra* note 68, at 44–47.

caused you to desire what you ought not to have (or to feel particularly drawn to it). But have they manipulated you?

Surely not. If you really cannot resist the temptation, then your friends have offered you an irresistible incentive—they have coerced you. If you *can* resist it, and yet you do not, then you have simply been persuaded by bad reasons (e.g., “If a thing you love is paraded before you then you really ought to have it, regardless of the consequences”). Either way, you have not been deprived of authorship over your decision. You have chosen to drink; you have willed it. Or, at the very least, you were unable to will yourself not to drink. Indeed, this last possibility—that you suffered from weakness of will—illustrates the point most clearly. If parading martinis in front of you caused an internal struggle, one you were aware of and engaged in with intention, there can be no doubt that it was you (in all your complexity) who made the final choice.⁷⁰

To sharpen the point even further, manipulation cannot simply mean causing someone to make less-ideal decisions than they otherwise would, since it is possible to manipulate someone into making more ideal decisions. Imagine your friends (different ones) know you love martinis, but they also know you worry about gaining weight. Having heard that you intend to spend the evening drinking—which you agree you really should not do—your friends share articles on Facebook and Twitter about the high calorie content of alcohol and the correlation between drinking and weight gain, knowing you are likely to see them. You are unaware that your friends have plotted to influence you, but as a result of seeing the articles you decide not to drink. In this case, you have been manipulated since your friends have insinuated themselves covertly into your decision-making process and redirected it to their own ends. This is despite the fact that their ends (and the decision-making process that led you to them) are—by your own admission—more ideal.

Apart from the direction of the influence (toward or away from ideals), the salient difference between these two cases is that the influence in the first case is overt, while the influence in the second case is hidden. You would feel manipulated in the second case—if you later found out why your friends had posted those articles—because you would realize that you had been motivated by someone else’s reasons. You made what seemed at the time like a decision that was yours, only to find out that it was infected by external machinations. In the first case, by contrast, you knew while you were deciding that your friends were trying to influence your decision. Because you were aware of their influence, you could treat it as you would any other decision—you could

⁷⁰ Alan Ware makes a similar argument to this, but he frames his analysis in terms of opportunity for consent, rather than awareness of and authorship over the influence. See Alan Ware, *The Concept of Manipulation: Its Relation to Democracy and Power*, 11 BRITISH J. POL. SCI. 163, 169–70 (1981).

contemplate it, weigh it against other considerations, attempt to mount defenses against it, and so on. The decision you reached was therefore your own (to whatever extent our decisions are ever our own). It was no different, in this respect, from just happening to see a perfect martini at the next table over at dinner. You would be tempted by it, sure, but not manipulated.

The hiddenness of manipulative influences explains how it is possible to alienate someone from their own decision-making powers. In order to get someone to act the way you want without realizing *why* they are acting that way, they must be unaware of the influence. As soon as we become conscious of outside influence, of someone else's plans and how we are implicated in them, we incorporate that influence into our own decision-making. Once you know someone else is trying to get you to do something, that fact becomes a regular part of how you make up your mind. It becomes one of the reasons that helps you explain your actions to yourself. Since we are never totally free of outside influence, what gives us (part) authorship over our own actions is that we regard our own reasons for acting as authoritative.⁷¹ Manipulation thwarts that.

Robert Goodin places the notion of hidden influence (or what he calls “deceptive influence”) at the center of his theory of manipulation. “One person manipulates another,” he writes, “when he deceptively influences him, causing the other to act contrary to his putative will.”⁷² Alan Ware agrees, arguing that for A to manipulate B it must be true that (among other things) “B either has no knowledge of, or does not understand, the ways in which A affects his choices.”⁷³ Although nearly everyone else concedes that hidden influence is an effective means of manipulation, not everyone agrees that hiddenness is necessary. To demonstrate this point, a number of theorists have proposed cases of supposedly overt manipulation.⁷⁴ Noggle invokes the image of Satan tempting Christ as he fasts in the wilderness.⁷⁵ Anne Barnhill claims that blatant guilt trips are manipulative.⁷⁶ Moti Gorin describes an employee who tries openly to get his co-worker—a recovering alcoholic—to relapse in order

⁷¹ See RAZ, *supra* note 62, at 204; Buss, *supra* note 63, at 195.

⁷² ROBERT E. GOODIN, *MANIPULATORY POLITICS* 19 (1980). We distinguish between manipulation and deception, below, and use the term “hidden influence” rather than “deceptive influence,” in order to keep the two concepts apart. Furthermore, as we argue in Part IV, in our view, it is not strictly necessary for manipulation to result in the target acting contrary to their putative will. It is possible, in other words, to be manipulated and still behave as one would have, absent the manipulation.

⁷³ Ware, *supra* note 70, at 165.

⁷⁴ Wood, *supra* note 53, at 38. Wood leaves the door open to the possibility that there might be cases of overt manipulation, but he does not develop the point.

⁷⁵ Noggle, *supra* note 68, at 44.

⁷⁶ Barnhill, *supra* note 68, at 60.

to beat her out for a promotion.⁷⁷ In our view, these are not cases of manipulation, since they share all the relevant features of the martini case above. If you know you are being guilted (in Barnhill's case) or tempted (in Noggle's and Gorin's cases), then you have not been deprived of authorship over your decision. Either you cannot resist the influence and have therefore been coerced, or you can resist it and do not, in which case you have simply been moved by bad reasons.

These arguments give us little reason to think manipulation can operate out in the open. It is only because the target of influence is unaware of or does not understand how they are being influenced that the manipulator's intervention can infiltrate and disrupt the target's decision-making process, steering them without force.

C. Deception and Manipulation

With this in mind, one can begin to see how manipulation is related to deception. To deceive someone is to cause them to hold false beliefs. This, of course, can be a powerful tool of manipulation. Imagine, once again, the manager of our car dealership. If she told her team they would be fired for refusing to work late but in fact lacked the authority to carry out the threat, or if she held a convincing, but fake, gun to their heads, then she would be deceiving them. She would be inducing them to act under false pretenses. Instead of changing or restricting the options open to her team members (i.e., persuading or coercing them), she would be disposing them to act the way she wanted by undermining their ability to understand their options. Here we have an example of deception in the service of manipulation. Although the team members know their manager is influencing them, they are misled about *how* she is influencing them. They are fed false beliefs about their options, which induces them to decide in ways they likely would not endorse if they had access to all the facts.

Deception is thus an important tool in the manipulator's toolkit, but it is not the only one. While instilling false beliefs is a blunt way of controlling another person's decision-making process, there are subtler means of shaping a person's beliefs. Social scientists have shown that human reasoning is "bounded" in significant ways. When evaluating information, people use a variety of unreliable shortcuts and heuristics, which researchers describe as "cognitive biases." Famously, Daniel Kahneman and Amos Tversky demonstrated that people are often influenced by irrelevant information (so-called "anchoring effects") and give more weight to evidence they can easily

⁷⁷ Moti Gorin, *Towards a Theory of Interpersonal Manipulation*, in *MANIPULATION: THEORY AND PRACTICE* 73, 80–81 (Christian Coons & Michael Weber eds., 2014).

recall (the “availability heuristic”).⁷⁸ People draw different conclusions from the same information depending on how it is presented (“framing effects”), and so on.⁷⁹ Because these cognitive biases are widespread and predictable, manipulators can easily treat them as vulnerabilities to exploit. Manipulators can remind targets of unimportant facts so that they give them undue weight. They can point out that their targets’ friends believe certain things in hopes that they will believe them too. They can frame information in misleading ways. Manipulation, therefore, need not involve outright deception; the truth can also be used to control our decision-making.

In fact, manipulators need not influence beliefs at all. To use Noggle’s metaphor, there are other “psychological levers” a manipulator can “adjust.”⁸⁰ Some of the most common examples of manipulation involve leveraging emotions and desires. If you and your partner are having a disagreement and you know they still feel guilty about something unrelated from the week before, you can subtly mention the source of guilt in conversation to wear down their resolve. Rather than create straightforward advertisements, where products are center-stage, marketers try to subliminally connect their products to our fantasies and aspirations by paying celebrities to appear with them or having them casually turn up on television and in films.

Though people like to imagine themselves as reliable and independent decision-makers, in fact they have many vulnerabilities to exploit. Beliefs, desires, and emotions form in response to a wide variety of factors, many of which operate outside conscious awareness. If a manipulator wants to influence someone without their knowledge, they need only intervene in a way that flies under the metaphorical radar. Deception exploits one kind of vulnerability—lack of perfect information—and can therefore be understood as a species of manipulation. But exploiting vulnerabilities in the way people form and manage desires and emotions is manipulative too.

D. Nudges and Manipulation

We have so far distinguished manipulation from persuasion and coercion, demonstrating that the former is hidden and works by exploiting vulnerabilities in the way we make decisions, while the latter operate out in the open, appealing directly to our capacity for conscious decision-making. We have also posited the relationship between manipulation and deception,

⁷⁸ Amos Tversky & Daniel Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124, 1128 (1974). For Kahneman’s more recent, popular treatment, see DANIEL KAHNEMAN, THINKING, FAST AND SLOW (2011).

⁷⁹ See Amos Tversky & Daniel Kahneman, *Rational Choice and the Framing of Decisions*, 59 J. BUS. 251, 257 (1986).

⁸⁰ Noggle, *supra* note 68, at 44–47.

arguing that people may deceive in order to manipulate, but that manipulation does not require instilling false beliefs. It remains to distinguish manipulation from nudging.

Now well-known, the term “nudge” was coined by the behavioral economist Richard Thaler and legal scholar Cass Sunstein, and refers to “any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives.”⁸¹ By “choice architecture,” Thaler and Sunstein mean the context in which people make decisions. To take one of their familiar examples, when someone is deciding what to buy at a cafeteria, the arrangement of options—some at eye-level, others slightly out of reach—influences the choices they make as people are more likely to reach for what’s closest.⁸² Thus, in order to shape other people’s decision-making, one can alter their choice architecture to nudge them in a preferred direction.

The notion of choice architecture is useful for understanding manipulation in general, and online manipulation in particular; we carry this notion with us through the discussion that follows. Like manipulation, intentionally shaping someone’s choice architecture to nudge them in a particular direction involves influencing their decision-making without force. What’s more, nudges often work by leveraging cognitive biases. Those familiar with the growing literature on nudges will recognize many of their key features in our characterization of manipulation as hidden influence and our description of exploiting vulnerabilities as the primary means of manipulation. If the way options are selected, arranged, and presented (i.e. choice architecture) affects the way people understand and respond to them, and if such effects operate outside conscious awareness, the question arises as to where nudging ends and manipulation begins.⁸³

Recognizing an overlap, Sunstein argues that while some nudges might be manipulative, most are not. To support this claim he offers a definition of manipulation very much in the spirit of our discussion. An influence is manipulative, he argues, “to the extent that it does not sufficiently engage or

⁸¹ THALER & SUNSTEIN, *supra* note 36, at 6.

⁸² *Id.* at 1–4.

⁸³ See generally, e.g., Robert Noggle, *Manipulation, Salience, and Nudges*, 32 *BIOETHICS* 164, 166–68 (2018); Thomas RV Nys & Bart Engelen, *Judging Nudging: Answering the Manipulation Objection*, 65 *POL. STUD.* 199, 203–08 (2016); Gérard Reach, *Patient Education, Nudge, and Manipulation: Defining the Ethical Conditions of the Person-Centered Model of Care*, 10 *PATIENT PREFERENCE & ADHERENCE* 459, 460–61 (2016); Luc Bovens, *The Ethics of Nudge*, in *PREFERENCE CHANGE: APPROACHES FROM PHILOSOPHY, ECONOMICS AND PSYCHOLOGY* 207 (Till Grüne-Yanoff & Sven Ove Hansson eds., 2009); Evan Selinger & Kyle Whyte, *Is There a Right Way to Nudge? The Practice and Ethics of Choice Architecture*, 5 *SOC. COMPASS* 923, 928–930 (2011); Daniel M. Hausman & Brynne Welch, *Debate: To Nudge or Not to Nudge*, 18 *J. POL. PHIL.* 123, 130–35 (2010).

appeal to [the target's] capacity for reflection and deliberation."⁸⁴ He disagrees, however, that manipulative influences must be hidden. Unlike the challenges to the hiddenness condition that we have already considered, Sunstein argues that while manipulation is usually hidden, we ought not to consider hiddenness its defining criterion because the notion of "hiddenness" is too difficult to spell out.⁸⁵ What about an influence needs to be exposed, he asks, in order for it not to be hidden? When the government orders that graphic health warnings be presented on cigarette packs, must it also—in order not to be manipulative—disclose information about the particular psychological mechanisms that make such warnings effective?⁸⁶

Although the question Sunstein raises is a good one, the answer is not as difficult as he suggests. Sunstein is surely right that when people make everyday decisions, much of what conditions their outcomes is not wholly apparent to them. People are not entirely self-transparent. They do not—and perhaps cannot—reflect on every belief, desire, emotion, and habit that impacts the decisions they reach. Moreover, they cannot know everything about how the world of choices they confront—whether in the natural world or the human-built one—came to be precisely the way it is, i.e., how the decision-making contexts they face were formed. Yet the reason this does not, generally, thwart a person's capacity to act competently is that they are able to act on the basis of reliable assumptions. Upon encountering graphic health warnings, for example, the average person has no trouble intuiting why they are there or how—roughly—they work: the health bureaucracy is encouraging people not to smoke, both by providing information and by appealing to a visceral sense of fear and disgust. Hidden influences thwart such assumptions. If people learned after some time that the *real* reason graphic health warnings were placed on cigarette packages is that the alcohol lobby paid off government officials, in an attempt to drive people away from smoking and toward drinking, the influence would be hidden in the relevant sense and the public would rightly feel manipulated.

Interestingly, Sunstein's definition raises an even more difficult question. Namely, what is "sufficient" engagement of a target's capacity for reflection and deliberation? The notion of sufficiency here is a normative one—what constitutes sufficient engagement is a question about values. When attempting to influence someone, how much *ought* the influencer to engage their target's capacity for reflection and deliberation? This normative component of Sunstein's definition creates internal tensions in his account. He argues that "from the standpoint of welfare, there might . . . be a justification

⁸⁴ SUNSTEIN, *supra* note 19, at 82.

⁸⁵ *Id.* at 102–05.

⁸⁶ *Id.*

for hidden manipulation in . . . extreme circumstances—as, for example, when people are trying to stop a kidnapping or to save a kidnapping victim.”⁸⁷ We agree that manipulation might be justified in such circumstances, but it is not clear why that situation is manipulative under Sunstein’s theory. If the hidden influence is justified, then it would seem that the level of engagement with the target’s (i.e., the kidnapper’s) capacity for reflection and deliberation was decidedly sufficient. Yet it is important to recognize that even in situations where manipulation might, ultimately, be justified, it is still manipulation; and, as a result, it carries with it a real harm, which must be weighed against potential benefits.⁸⁸

Thus, while we agree with Sunstein’s conclusion that some nudges are manipulative and others are not, we believe that our account of manipulation as hidden influence better delineates between them. Sunstein suggests, for example, that disclosures (purely informational nudges) are not manipulative since they represent “an effort to appeal to [people’s] deliberative capacities” rather than bypass them.⁸⁹ But there are different types of disclosures. Nutrition labels are straightforward—they simply provide consumers with some of the information they need to make an informed choice about what foods to buy. They are not manipulative because they are not designed to exploit a cognitive bias, and they instead encourage and strengthen the consumer’s capacity for conscious deliberation. And there is nothing hidden about them—people understand why they are there and what effects they are meant to have. Another case that Thaler and Sunstein cite approvingly is less clear. Deploying a so-called “social nudge,” the state of Minnesota informed its citizens that 90% of Minnesotans pay their taxes in order to encourage the 10% who do not pay to comply.⁹⁰ Depending on how this information was conveyed, this strategy is plausibly manipulative because, although the information it conveyed was true, its purpose may have been hidden.

In our view, nudges are manipulative if they are hidden and exploit vulnerabilities. Many nudges—indeed, most of the nudges that Thaler and Sunstein recommend—are transparent and work to *correct* cognitive vulnerabilities rather than exploit them. As such, they are not manipulative at all. Another distinction between nudges and manipulation, which we discuss at length in the sections that follow, stems from the fact that manipulation is usually *targeted*.⁹¹ In order to exploit someone’s vulnerabilities, one must know something about what those vulnerabilities are and how precisely to

⁸⁷ *Id.* at 104.

⁸⁸ We discuss the precise nature of that harm in Part IV, below.

⁸⁹ SUNSTEIN, *supra* note 19, at 82.

⁹⁰ THALER & SUNSTEIN, *supra* note 36, at 66–65.

⁹¹ See Patrick Todd, *Manipulation*, in 5 INT’L ENCYCLOPEDIA ETHICS 3139, 3140 (Hugh LaFollette ed., 2013).

leverage them. Most nudges, by contrast, are not targeted to particular individuals. Like the cafeteria example, above, the sorts of nudges Thaler and Sunstein advocate for are meant to be applied in the same way to everyone.

As we discuss below, part of what makes information technology particularly well-suited to facilitating manipulation is that it allows for fine-grained microtargeting, making it possible for potential manipulators to engage in what Karen Yeung calls “hypernudging.”⁹² Hypernudges are not only hidden, they precisely target and exploit individual vulnerabilities, making them much more difficult to resist. Although we can imagine mass manipulation, which attempts to influence large groups of people all in exactly the same way, the more targeted manipulation is the more we ought to worry about it.

E. Manipulation and Manipulative Practices Defined

At this point, we can pull these different strands together and define manipulation as *imposing a hidden or covert influence on another person’s decision-making*. That means influencing someone’s beliefs, desires, emotions, habits, or behaviors without their conscious awareness, or in ways that would thwart their capacity to *become* consciously aware of it by undermining usually reliable assumptions. This definition captures what is essential about manipulation—namely, that it disrupts the target’s capacity for self-authorship. Which is to say, it explains why, upon learning they have been manipulated, people feel like puppets. It differentiates manipulation from persuasion and coercion, which are forthright efforts to alter a person’s decision-making process. It reveals that nearly all instances of deception are also instances of manipulation, but not all manipulation is deceptive. And it helps distinguish between nudges that are manipulative and those that are not. Furthermore, we have considered *how* manipulators can impose hidden influences on their targets: by targeting and exploiting their cognitive, emotional, or other decision-making vulnerabilities.

An issue we have not considered in depth is whether or not a hidden influence must be *intentionally* imposed to be manipulation. We believe that it does. The notion of manipulation naturally evokes the image of a manipulator, and influences hidden by natural or accidental means would likely not undermine a person’s sense of self-authorship. It is known and expected that some of the underlying causes of experience lay outside conscious awareness. What is not expected is intentional meddling in other people’s decision-making. People assume—and ought to be able to assume—that they do not live in the company of Descartes’ evil demon or on the set of

⁹² Yeung, *supra* note 6.

the *Truman Show*. Thus, when we discuss manipulation as “hidden or covert influence,” we mean such cases where hiddenness is intended.

Even with this definition in hand, however, identifying cases of manipulation is difficult because manipulation is a “success concept.”⁹³ Which is to say, the claim that someone has been manipulated refers not only to the strategies employed by the influencer but also to the effects of those strategies on the influenced. Analogously, consider the distinction between lying and deceiving. To lie is to assert a falsehood. To deceive, by contrast, is to instill false beliefs. In order to determine whether or not someone has lied, one only needs to investigate the actions of the alleged liar. In order to determine whether or not someone has deceived, one needs to investigate the actions of the deceiver *as well as the resulting beliefs of the deceiver’s target*. The same is true of manipulation. Manipulation only exists where the attempt to manipulate succeeds and the manipulee’s decision-making is affected by the influence. Investigating cases of manipulation, therefore, requires a focus on the status of two parties—the intentions of the suspected manipulator and the beliefs and actions of potential manipulees. In many cases, establishing impacts on targets of manipulation may require far-flung empirical findings that are difficult, if not impossible, to access.

For this reason, we focus in what follows on the concept of *manipulative practices*—strategies that a reasonable person should expect to result in manipulation—and not on the success concept of manipulation, in toto. The preceding discussion points to three key characteristics of manipulative practices: they involve influences that (1) are hidden, (2) exploit cognitive, emotional, or other decision-making vulnerabilities, and (3) are targeted. Strictly speaking, the only necessary condition of manipulation is that the influence is hidden; targeting and exploiting vulnerabilities are the means through which a hidden influence is imposed. Indeed, targeting is best understood as an exacerbating condition: the more closely targeted a strategy is to the specific vulnerabilities of a particular manipulee, the more effective one can expect that strategy to be. Hidden influences not targeted at all still count, therefore, as manipulative, but we ought to worry more about manipulative practices the more targeted they are. As we will see in the next section, this is especially important when considering manipulation online.

Drawing on this analysis, let us briefly return to the cases we discussed earlier in this article.⁹⁴ Did Facebook, Uber, or Cambridge Analytica engage in manipulative practices? The Facebook case is the clearest. Its strategy memo described identifying moments when teenagers felt emotionally vulnerable and deploying advertisements to leverage that vulnerability. Such

⁹³ See Wood, *supra* note 53, at 11.

⁹⁴ See *supra* Part II.

practices are targeted to the individual level, exploit vulnerabilities by design, and are hidden—i.e., there is reason to believe few people targeted would know why they were seeing that particular advertisement at that particular moment. On our account, this is manipulative beyond any doubt. The only question, in this case, is whether Facebook engaged in the practice; the company denies that it did.

The Cambridge Analytica case has much in common with the Facebook case. By its own account, Cambridge Analytica attempted to “exploit what we knew about [voters]” and “target their inner demons”⁹⁵—i.e., Cambridge Analytica aimed to frame political messages in ways each targeted individual was most inclined to accept and internalize. This practice is targeted and exploits individual vulnerabilities by design, and as in the case of Facebook above, there is little reason to believe that individuals seeing Cambridge Analytica’s advertisements would know either that the ads were tailored to them specifically or how the tailoring was accomplished. These practices are manipulative under our framework. Unlike with Facebook, we know that Cambridge Analytica actually engaged in these practices; it remains a mystery however, the extent to which they were effective. Even if Cambridge Analytica’s manipulative practices only had minor impacts on recent elections, we ought to treat this case as a cautionary tale. As we discuss in the next Part, these practices are likely to increase in power and sophistication.

Finally, Uber: this case is complicated, not least because there are so many potentially manipulative practices in which the company is known to have engaged. Some strategies seem clearly manipulative, such as intentionally misleading drivers with surge pricing “heat maps” that conflate real-time and predictive demand. But other strategies may not necessarily be manipulative. For example, automatically queuing the next ride request before the driver has time to decide whether or not to continue working might *feel* manipulative because it exploits the fact that we need time to make decisions. However, there is nothing hidden about this strategy. Drivers likely understand perfectly well what is happening, and though they may feel pressured, they likely do not feel tricked. Thus, on our account, automatic queuing is not an obviously manipulative practice, though there may be other reasons for questioning its ethical standing.

One last Uber practice is worth considering—namely, the notifications designed to nudge drivers to stay on the road. Much like automatic queuing, drivers may experience this pressure as manipulative, but as it is neither hidden nor significantly targeted, we ultimately judge it not to be so. As a nudge, it is designed to exploit drivers’ desire to accomplish goals, even if the goals are

⁹⁵ *Cambridge Analytica and Facebook: The Scandal So Far*, *supra* note 44.

functionally meaningless, but it is not a manipulative practice. Tweak some aspects of this scenario, however, and one can imagine future versions of this practice that are indeed manipulative. Consider if notifications were timed to appear right when drivers were desperate—say, if the earnings goals were not arbitrary, but instead were indexed to bills coming due. If the apparatus of such influence were hidden from the driver, manipulation would be a worry.

That possibility is important as we move into the next section, where we discuss more systematically how information technology can facilitate manipulation. Sometimes, as we have seen, technologically-mediated practices are manipulative; sometimes they are not. Our goal in the next section is to make clear how information technology and manipulation intersect.

IV. MANIPULATION THROUGH INFORMATION TECHNOLOGY

Having defined manipulation and developed the concept of manipulative practices, we are now in a position to address this article's next question: what is it about information technology that, together with manipulation, makes such a worrying combination? Building on our definition of manipulation, generally, we define online manipulation as *the use of information technology to covertly influence another person's decision-making*. And, accordingly, we define online manipulative practices as *applications of information technology that impose hidden influences on users, by targeting and exploiting decision-making vulnerabilities*.

We argue, below, that as digital technologies are incorporated into all aspects of people's everyday lives, they become increasingly susceptible to this kind of manipulation. Widespread digital surveillance means it takes little effort to identify people's vulnerabilities. Digital platforms are the perfect medium through which to leverage those insights. And because information technology mediates so much of so many people's lives, there is virtually limitless opportunity to invisibly influence.

A. Surveillance

Living in an "Information Age" means that nearly everything people do is tracked.⁹⁶ Both the information individuals knowingly disseminate about themselves (e.g., when they visit websites, make online purchases, and post

⁹⁶ See, e.g., SHOSHANA ZUBOFF, SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER (2019); Julia Angwin, WHAT THEY KNOW (2019), <http://juliaangwin.com/the-what-they-know-series/> [<https://perma.cc/88QA-4LTU>]. There is, of course, an enormous academic literature on surveillance. For a helpful introduction, see DAVID LYON, SURVEILLANCE STUDIES: AN OVERVIEW (2007).

photographs and videos on social media) and the information they unwittingly provide (e.g., when those websites record data about how long they spend browsing them, where they are when they access them, and which advertisements they click on) reveals a great deal about who each individual is, what interests them, and what they find amusing, tempting, and off-putting.⁹⁷ Moreover, one need not “go online” in the traditional sense to be digitally tracked. Credit card purchases log what people buy in brick-and-mortar stores,⁹⁸ law enforcement license plate readers track where they drive,⁹⁹ and facial recognition software identifies them as they move through public spaces.¹⁰⁰ The workplace has also become a site of intense digital surveillance. Private firms carefully monitor and analyze how their employees conduct themselves throughout the workday and beyond, leading to what some have called a new “digital Taylorism.”¹⁰¹

Tal Zarsky has warned that information technology could potentially be used to manipulate people by harnessing this wealth of information to precisely tailor advertisements that exploit their vulnerabilities.¹⁰² Ryan Calo has deepened Zarsky’s account by describing these practices in terms of what Jon Hanson and Douglas Kysar call “market manipulation”—exploiting cognitive biases in order to influence consumer behavior.¹⁰³ Calo updates

⁹⁷ As Julie Cohen argues, this information is a source of tremendous value and is treated by private firms as a repository of resources to be mined. Understood through the lens of political economy, Cohen suggests that we think of this information as a “biopolitical public domain”: “Personal information processing has become the newest form of bioprospecting, as entities of all sizes compete to discover new patterns and extract their marketplace value.” Julie E. Cohen, *The Biopolitical Public Domain: The Legal Construction of the Surveillance Economy*, 31 PHIL. & TECH. 213 (2017).

⁹⁸ Michael Reilly, *Google Now Tracks Your Credit Card Purchases and Connects Them to Its Online Profile of You*, MIT TECH. REV. (May 25, 2017), <https://www.technologyreview.com/s/607938/google-now-tracks-your-credit-card-purchases-and-connects-them-to-its-online-profile-of-you/> [<https://perma.cc/K33S-UBDB>].

⁹⁹ American Civil Liberties Union, *YOU ARE BEING TRACKED: HOW LICENSE PLATE READERS ARE BEING USED TO RECORD AMERICANS’ MOVEMENTS* (2013), https://www.aclu.org/sites/default/files/field_document/071613-aclu-alprreport-opt-v05.pdf [<https://perma.cc/2QKK-4UWZ>].

¹⁰⁰ Jennifer Lynch, *FACE OFF: LAW ENFORCEMENT USE OF FACE RECOGNITION TECHNOLOGY* (2018), <https://www.eff.org/files/2018/02/15/face-off-report-1b.pdf> [<https://perma.cc/MS78-QTR8>].

¹⁰¹ See, e.g., Christian Parenti, *Big Brother’s Corporate Cousin: High-Tech Workplace Surveillance is the Hallmark of a New Digital Taylorism*, NATION (July 27, 2001), <https://www.thenation.com/article/big-brothers-corporate-cousin/> [<https://perma.cc/7X7B-E8X8>]; Richard Salame, *The New Taylorism*, JACOBIN (Feb. 20, 2018), <https://www.jacobinmag.com/2018/02/amazon-wristband-surveillance-scientific-management> [<https://perma.cc/FR49-GT5V>].

¹⁰² See Zarsky, *supra* note 1, at 219–20.

¹⁰³ See Calo, *supra* note 1, at 1000.

Hanson and Kysar's theory to reflect the increased capacities for market manipulation that information technology furnishes (what Calo calls "digital market manipulation").¹⁰⁴ For example, he points to the use of "disclosure ratcheting"—using behavioral nudges to dispose individuals to reveal more information about themselves than they intend (information which is then used to optimize further nudges)—and "means-based targeting"—using online experiments like A/B testing¹⁰⁵ to tailor the presentation of each advertisement.¹⁰⁶

On the whole, information that was once difficult and expensive to uncover is now available easily and cheaply. People reveal their vulnerabilities constantly, and as the Facebook case suggests, many parties are eager to accrue information in order to leverage it for manipulative ends. Importantly, ubiquitous digital surveillance means that people reveal information about their vulnerabilities not only through what some might judge frivolous or unnecessary disclosures—e.g., on social media. Rather, as the Uber case suggests, people are rendered susceptible to this kind of tracking merely in the routine, day-to-day activities of everyday life: going to work, commuting, and communicating, or even as the Cambridge Analytica case highlights, by simply being associated with others who make disclosures about themselves.

B. Digital Platforms

In addition to providing insight into vulnerabilities that advertisers, employers, and political campaigns may want to exploit, information technology also makes it increasingly easy to leverage those insights. As Karen Yeung argues, digital platforms facilitate "Big Data-driven decision-guidance techniques," which constitute a kind of "hypernudging."¹⁰⁷ Unlike traditional advertisements, which were static and disseminated en masse, digitally-mediated platforms, such as websites and social media applications, constitute dynamic, interactive, intrusive, and personalized choice

¹⁰⁴ *Id.*

¹⁰⁵ In website and app design, A/B testing is a method for determining the relative effectiveness of messaging strategies or user interface elements by presenting different variations to different groups of users and measuring user response.

¹⁰⁶ Calo, *supra* note 1, at 1012–14; *see also* Alessandro Acquisti et al., *Privacy and Human Behavior in the Age of Information*, 347 *Sci.* 509 (2015).

¹⁰⁷ Yeung, *supra* note 6, at 121; *see also* Lanzing, *supra* note 7.

architectures.¹⁰⁸ Websites can alter the presentation of information depending on the specific things they know about each individual visitor. And if the websites do not know much, they can learn as users interact with them. If an advertiser knows that someone is more likely to buy things at a particular time of the day or a particular day of the week—as the Facebook case demonstrated—the advertiser does not have to wait, passively, for that person to browse past their campaigns; the advertiser can send emails or text notifications, or present themselves in the target’s social media feeds at optimal moments. And because this tailoring process is automated, it is fully personalized for each individual target. As Calo puts it, “firms will increasingly be in the position to *create* suckers, rather than waiting for one to be born.”¹⁰⁹

Crucially, as Yeung points out, the insights into people’s vulnerabilities that digital platforms utilize in shaping their choice architectures are not limited to what can be gleaned from information collected about each individual. Platforms also enable “monitoring and refinement of the individual’s choice environment in light of *population-wide* trends identified via population-wide Big Data surveillance analysis.”¹¹⁰ In other words, the digital systems people interact with study both their individual idiosyncrasies *and* the patterns that emerge amongst demographic groups to which they belong, potentially revealing weaknesses and dispositions that individuals themselves cannot see. As Frank Pasquale argues, “the real basis of commercial success in Big Data-driven industries is likely the quantity of

¹⁰⁸ By digital platforms, we mean the built online environments, such as websites and smartphone applications, that facilitate online interaction, and through which individuals communicate, access information, and engage in a wide range of other activities, from reading and writing, conducting research, and playing games to finding work, shopping, and arranging services. For a critical analysis of the concept of platforms, see Tarleton Gillespie, *The Politics of “Platforms,”* 12 *NEW MEDIA & SOC’Y*. 347 (2010).

¹⁰⁹ Calo, *supra* note 1, at 1018. These worries are connected to growing discussions around so-called “dark patterns”—user interface design strategies that drive users to behave in ways beneficial to online services and potentially harmful to users themselves. A number of scholars have suggested these design strategies can be manipulative. *See, e.g.,* Arunesh Mathur et al., *Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites*, in 22ND ACM CONFERENCE ON COMPUTER-SUPPORTED COOPERATIVE WORK AND SOCIAL ENGINEERING 81 (2019); Jamie Luguri & Lior Strahilevitz, *Shining a Light on Dark Patterns* (U. Chicago, Poverty Law Working Paper No. 719, Aug. 7, 2019). While a thorough analysis of dark patterns is outside the scope of this paper, the characteristics of manipulative practices highlighted above (i.e., that they are hidden, exploit decision-making vulnerabilities, and are targeted) might be applied productively to debates in this area, helping to distinguish between manipulative and non-manipulative patterns. We discuss this briefly in related work, see Susser et al., *Technology, Autonomy, and Manipulation*, 8 *INTERNET POL’Y REV.* 7 (2019).

¹¹⁰ Yeung, *supra* note 6, at 122 (emphasis added).

relevant data collected *in the aggregate*—something not necessarily revealed or shared via person-by-person disclosure.”¹¹¹

C. Mediation

Finally, there is a tendency to treat information technology as the kind of thing people approach, direct their attention toward, and engage with intention. But it is more appropriate to understand information technology as a set of tools that increasingly mediates everyday experience. Digital platforms are more like eyeglasses than magnifying glasses—technologies one wears and forgets about rather than those one picks up and puts to use. As Don Ihde argues, when people use such technologies, the technologies recede from view.¹¹² They attend not to the eyeglasses themselves, but to what they can see *through* them.¹¹³ It is only when something goes wrong—when one *cannot* see through the glasses because they get dirty or break—that we notice the technology itself. Likewise, people pay very little attention to smartphones and computers themselves, focusing instead on the information they can access through them, the pictures they can see and videos they can watch, the things they can buy, and directions they can follow.¹¹⁴ Even the very tech-savvy, who understand what is going on behind-the-scenes of the technologies they use—for example, the way data is collected and stored, analyzed and transmitted—likely spend little time considering all that takes place under the hood. For the average person, information technology sits almost entirely outside conscious awareness; few ponder its nature and inner-functioning while they are actively engaged in using it.¹¹⁵

¹¹¹ PASQUALE, *supra* note 1, at 153; *see also*, Solon Barocas & Helen Nissenbaum, *On Notice: The Trouble with Notice and Consent*, in PROCEEDINGS OF THE ENGAGING DATA FORUM: THE FIRST INTERNATIONAL FORUM ON THE APPLICATION AND MANAGEMENT OF PERSONAL ELECTRONIC INFORMATION (2009) (“Aggregated user data can be subject to computerized analyses to produce predictive models that can be used to estimate other like users’ propensity to respond to certain ads.”).

¹¹² *See generally* Don Ihde, *Technology and the Lifeworld: From Garden to Earth* (1990).

¹¹³ *Id.*

¹¹⁴ Robert Rosenberger & Peter-Paul Verbeek, *A Field Guide to Postphenomenology*, in POSTPHENOMENOLOGICAL INVESTIGATIONS: ESSAYS ON HUMAN-TECHNOLOGY RELATIONS 9, 37–38 (Robert Rosenberger & Peter-Paul Verbeek eds., 2015).

¹¹⁵ For more on what postphenomenologists call “transparency” (not to be confused with the notion of transparency as it is used in policy contexts) and its relation to using information technology, *see* Daniel Susser, *Transparent Media and the Development of Digital Habits*, in POSTPHENOMENOLOGY AND MEDIA: ESSAYS ON HUMAN-TECHNOLOGY-WORLD RELATIONS 27 (Yoni Van Den Eede et al. eds., 2017); *see also*, Diane Michelfelder, *Postphenomenology with an Eye to the Future*, in POSTPHENOMENOLOGICAL INVESTIGATIONS: ESSAYS ON HUMAN-TECHNOLOGY RELATIONS 237, 237–46 (Robert Rosenberger & Peter-Paul Verbeek eds., 2015).

This is important to recognize in the context of worries about online manipulation because (as we argued in the previous section) lack of awareness is the defining feature of manipulative strategies. Beyond the fact that information technology provides insight into people's vulnerabilities and furnishes platforms that can leverage those insights, it is designed to be *seen through* and is thus already, in a real sense, hidden. A world increasingly structured by information technology is a world increasingly removed from view—a world of screens people look through, cameras they walk past, and sensors they unknowingly impress upon. A determined manipulator could not dream up a better infrastructure through which to carry out his plans.

Furthermore, because information technology mediates so much of so many people's lives, the reach of online manipulation is virtually limitless. Just as information technology increases the scale of other activities, it increases the potential scale of manipulative influence. Facebook, for instance, currently has more than 2 billion monthly active users.¹¹⁶ If its platform were leveraged to manipulate users—as it allegedly was in the Cambridge Analytica case—the impact could be massive.¹¹⁷ Uber has 75 million monthly active drivers.¹¹⁸ The strategies it devises to influence its drivers can affect what happens in millions of cars around the world. Unlike “offline manipulation,” which is constrained by the manipulator's ability to understand and influence a finite number of other people, online manipulation is practically unbounded.

On the whole, information technology has made people's vulnerabilities both easy to detect and easy to exploit. We have become more manipulable, and—there is reason to believe—more manipulated.

V. HARMS OF MANIPULATION: THE SIGNIFICANCE OF AUTONOMY

One might agree with everything we have claimed thus far but not be concerned. What harm does manipulation cause and why should anyone care that the insurgence of digital technologies into everyday life exacerbates it? A plausible answer might point to the motivations of manipulators and suggest that the reason for resorting to manipulation is generally to induce targets to

¹¹⁶ *Facebook Company Information*, FACEBOOK, (accessed Nov. 13, 2019), <https://newsroom.fb.com/company-info/> [<https://perma.cc/4Z3B-Q4HH>].

¹¹⁷ In 2012, Facebook conducted its infamous “emotional contagion” experiment to gauge whether or not altering the content it showed users affected their emotional states, which involved more than 689,000 (unwitting) user-subjects. See Adam D.I. Kramer et al., *Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks*, 111 PROC. NAT'L ACAD. SCI. 8788, 8788 (2014).

¹¹⁸ Johana Bhuiyan, *Uber Powered Four Billion Rides in 2017. It Wants To Do More—And Cheaper—in 2018*, RECODE (Jan. 5, 2018), <https://www.recode.net/2018/1/5/16854714/uber-four-billion-rides-coo-barney-harford-2018-cut-costs-customer-service> [<https://perma.cc/X9JJ-XYR8>].

act against their interests. If, for example, Facebook targeted advertisements at vulnerable teenagers, the suspicion would be that they did so to exploit moments of weakness, to incline teenagers to buy something they did not need or to pay more for it than they otherwise would. As Calo argues, it is not these forms of online influence alone that are cause for concern; rather, it is the fact that they are “*coupled with divergent interests* that should raise a red flag.”¹¹⁹ According to such accounts, the incorporation of digital technologies into all aspects of people’s lives, and the attendant threat of manipulation, is worrying because the designers and operators of these technologies are often incentivized to subordinate others’ interests to theirs.

The subordination of interests, in certain cases, may be sufficient reason to condemn manipulative practices. But it fails to reach the heart of the matter. First, as we saw in the martini examples above, it is easy to imagine manipulative practices that advance, rather than diminish, a target’s interests. More importantly, beyond the direct, material harms that result from manipulation, such as exploitation, impoverishment, unfairness, and the deprivation of benefits, the deeper harm is infringement of individual autonomy. Since autonomy lies at the normative core of liberal democracies, the harm to autonomy rendered by manipulative practices extends beyond personal lives and relationships, reaching public institutions at a fundamental level.

In the following Section, we first clarify the conception of autonomy that motivates our analysis. Second, we look closer at the key characteristics of manipulation to expose how manipulative practices undermine autonomy. Our main focus is the hidden changes to choice architectures that exploit individual vulnerabilities. Finally, we consider possible objections to our account.

A. Autonomy at a Glance

Personal autonomy is the capacity to make one’s own choices, with respect to both existential and everyday decisions. As Joseph Raz, writes, “[t]he ruling idea behind the ideal of personal autonomy is that people should make their own lives.”¹²⁰ This notion of autonomy is rooted in the view that people can (mostly) rationally deliberate on the different options they are faced

¹¹⁹ Calo, *supra* note 1, at 1023.

¹²⁰ RAZ, *supra* note 62, at 369; *see also, e.g.*, Marilyn Friedman, *Autonomy, Social Disruption, and Women*, in RELATIONAL AUTONOMY: FEMINIST PERSPECTIVES ON AUTONOMY, AGENCY, AND THE SOCIAL SELF 35, 37 (Catriona Mackenzie & Natalie Stoljar eds., 2000) (“Autonomy involves choosing and living according to standards or values that are, in some plausible sense, one’s own.”).

with, that they know (mostly and roughly) what they believe and desire, and that they can act on the reasons they think best.¹²¹

Autonomy theorists often distinguish between competency conditions and authenticity conditions.¹²² Autonomous persons have the cognitive, psychological, social, and emotional *competencies* to deliberate, to form intentions, and to act on the basis of that process. And autonomous persons can (at least in principle) critically reflect on their values, desires, and goals, and act for their *own* reasons—i.e., endorse them *authentically* as their own. Importantly, this conception of autonomy is not overly rationalistic—autonomous persons deliberate on the basis of their beliefs and desires, as well as on the basis of their emotions, convictions, experiences, and commitments. While broad, this understanding of autonomy and its value is indispensable to a normative theory of manipulation: without recourse to some notion of autonomy there is no basis for holding people responsible for their actions, no way to ascribe authorship (or part-authorship) of behavior, to oneself or to others.¹²³

But let us steer clear of two possible misunderstandings: first, autonomy is sometimes still understood entirely in terms of solipsistic “rational choosers,” without taking social contexts (and social others) into account.¹²⁴ We agree with criticisms of this conception of autonomy and develop our own conception from theories of relational autonomy. Such accounts understand autonomy as a capacity of socially-situated persons, whose decisions issue from deliberative processes rich with emotion, feeling, imagination, and reason, and which are conditioned in part by their contexts.¹²⁵

Second, in our view, autonomy is not a capacity of ideal choosers, but rather of ordinary people making decisions in their everyday lives. When we say that to act autonomously people have to be able to act *for their own reasons*, we do not mean that everyone must always and necessarily critically reflect on their every step. We mean that when questioned (by others or by themselves) people are generally able to identify some reasons for their

¹²¹ See JOHN CHRISTMAN, *THE POLITICS OF PERSONS: INDIVIDUAL AUTONOMY AND SOCIO-HISTORICAL SELVES* 149–56 (2009); see generally *AUTONOMY, OPPRESSION, AND GENDER* (Andrea Veltman & Mark Piper eds., 2014); *PERSONAL AUTONOMY AND SOCIAL OPPRESSION: PHILOSOPHICAL PERSPECTIVES* (Marina A.L. Oshana ed., 2015).

¹²² See CHRISTMAN, *supra* note 121, at 155.

¹²³ See RAZ, *supra* note 62.

¹²⁴ See JULIE E. COHEN, *CONFIGURING THE NETWORKED SELF: LAW, CODE, AND THE PLAY OF EVERYDAY PRACTICE* 107–26 (2012).

¹²⁵ See Catriona Mackenzie & Natalie Stoljar, *Introduction: Autonomy Refigured, in RELATIONAL AUTONOMY: FEMINIST PERSPECTIVES ON AUTONOMY, AGENCY, AND THE SOCIAL SELF* 3, 4 (Catriona Mackenzie & Natalie Stoljar eds., 2000).

actions, reasons they themselves can endorse.¹²⁶ In other words, we are describing a hypothetically valid condition: if someone were to reflect on an action, that person could give a rough account of the reasons that motivated it.

On this account, autonomy is exercised by persons with particular identities and who always already live in social, cultural, political, and other contexts. Idealized conceptions of autonomy forget the “intersubjective and social dimensions of selfhood and identity for conceptions of individual autonomy and moral and political agency.”¹²⁷ Social contexts and relations both *enable* and *constrain* autonomous choices.¹²⁸ On one hand, autonomy requires living with others who teach one to act with reflection and deliberation, and it requires social contexts that enable autonomous choices by providing horizons of options. On the other hand, social contexts also *condition* choices because (often stereotypical) societal expectations influence the ways people choose.

Finally, autonomy is not only an individual good; it is also a social and political good. After all, people are not merely consumers, and markets are not the only institutions meant to respond to the preferences their decisions express; people also act as citizens, and democratic institutions are designed (ideally) to reflect autonomous decisions reached in the political sphere. When manipulators aim to elicit votes rather than purchases, it is the integrity of this realm—the political realm—that is called into question. As we saw in the Cambridge Analytica case, political advertisers have attempted to use psychographic profiling to create campaigns that exploit the decision-making vulnerabilities of individual voters. Such practices threaten the autonomy of citizens, and in doing so, they threaten democracy. Interfering with people’s autonomy is thus both an ethical concern and a political one.

¹²⁶ There is an interesting debate about which sorts of reasons should be acceptable here—any reasons, even seemingly delusional or irrational ones? Such questions are outside the scope of this Article. See, e.g., THOMAS NAGEL, *THE VIEW FROM NOWHERE* 164–88 (1986) (discussing difference between subjective and objective reasons); SUSAN WOLF, *MEANING IN LIFE AND WHY IT MATTERS* 37 (2010) (discussing delusionary and irrational activities-cum-reasons).

¹²⁷ Mackenzie & Stoljar, *supra* note 125, at 4. But one has to be careful here: although persons are always socially situated and have to realize their autonomy in social contexts, this does not imply that autonomy is only possible in (social and political) egalitarian relations supporting self-respect, self-esteem, and self-trust—and thereby autonomy. See Joel Anderson & Axel Honneth, *Autonomy, Vulnerability, Recognition, and Justice*, in *AUTONOMY AND THE CHALLENGES TO LIBERALISM: NEW ESSAYS* 127 (John Christman & Joel Anderson eds., 2005).

¹²⁸ See Friedman, *supra* note 120, at 35.

B. Autonomy and Manipulation

Throughout this article, we have argued that manipulative practices are characterized by the hiddenness of their influence on decision-making. By fine-tuning the contexts in which individuals choose—their choice architectures—a manipulator can exploit each person’s decision-making vulnerabilities, and in doing so interfere with their beliefs, desires, emotions, and behavior. But why is it that these interferences undermine autonomy? How do they differ from, say, overt guilt trips or transparent forms of pressure—influences we deem non-manipulative?

1. *Autonomy, Manipulation and Choice Architecture*

As we have seen, manipulators intentionally alter the contexts in which their targets make decisions. But, unlike those who persuade or coerce, manipulators do so without the target’s conscious awareness. The hiddenness of manipulation challenges both conditions of autonomy—competency and authenticity. Because manipulees are unaware that features of their choice environments have been intentionally designed to influence them, their capacity to (competently) deliberate is undermined, yielding decisions they cannot endorse (authentically) as their own.

In many cases of behavioral advertising that is obviously the point. Without realizing that their choice architectures were constructed to influence them, subtle cues are used to tempt or seduce people to buy things. So-called “native advertising”—advertisements designed to mimic the appearance of non-advertising content—provides an especially clear example of a manipulative online practice, since it intentionally conflates information and advertising. While reading the news or browsing social media, people contextualize and understand the information they absorb within that frame. It requires special vigilance to identify and re-contextualize native advertisements, in order to competently evaluate them.

The vast stores of information and detailed behavioral profiles compiled about each individual greatly facilitate manipulative practices. The more that is known about each person’s personality, preferences, habits, and vulnerabilities, the easier it is to construct choice environments that will guide their decision-making in the desired direction. In addition, as we argued in the previous section, the less users know about the inner-workings of the technologies that facilitate online manipulation, the easier it is for these mechanisms to function, making online manipulation especially worrying and its potential threat to autonomy so grave.

The relationship between manipulation, autonomy, and the construction of choice architectures raises another issue worth mentioning. We

pointed out in Part II that, at an abstract level, there are two ways to influence a person's decision-making: changing the options available to them and changing how they understand their options. Some argue that the term "manipulation" has two distinct senses, each corresponding to one of these strategies.¹²⁹ Manipulating a *person* involves interfering directly with their psychological processes, whereas manipulating a *situation* involves interfering with the options available to them. In our view, the preceding discussion reveals the limits of this distinction. Manipulators often change the choices available to someone—their situation—precisely to interfere with their decision-making processes. The construction of online choice contexts makes this clear. While it can be helpful in certain contexts to draw attention to the differences between situations and persons, explaining manipulation requires a holistic understanding. In autonomous actions, the person and the situation are always interconnected; one cannot be analyzed without the other.

Crucial to the self-understanding of autonomous actors is the sense that the choices they make are *their* choices, that they *know* what they are doing, and that they can, in principle, endorse the terms of their decisions. Hidden adjustments to people's choice environments fundamentally interfere with their autonomy, threatening their ability to act for reasons of their own.¹³⁰ As we have seen, this works so well online because online environments are especially well-suited to identifying the adjustments that most effectively shape people's choices.

¹²⁹ See Claudia Mills, *Politics and Manipulation*, 21 SOC. THEORY & PRAC. 97 (1995).

¹³⁰ This point can also be expressed through a traditional Kantian vocabulary. When manipulated, people are treated purely as means, and not, at the same time, as ends in themselves. In this way, manipulation can be understood as reification: people are being made into things; or, as Kant would put it, they are treated as not having dignity, but a price. Examining the problem through this lens is useful as well, as it reveals another worrying effect of manipulation: when people are treated solely as means, and not as ends in themselves, it likely has an impact on their self-understanding. See Jeremy Waldron, *It's All for Your Own Good*, N.Y. REV. BOOKS (Oct. 9, 2014), <http://www.nybooks.com/articles/2014/10/09/cass-sunstein-its-all-your-own-good/> ("What becomes of the self-respect we invest in our own willed actions, flawed and misguided though they often are, when so many of our choices are manipulated to promote what someone else sees (perhaps rightly) as our best interest?") [perma.cc link unavailable]. Here, Waldron criticizes nudges exercised for paternalist aims, but it is easy to see how much more damaging nudges are when people are manipulated not for the sake of paternalist aims but for the sake of sheer acquisitiveness.

2. *Manipulation and Vulnerabilities*

The concept of vulnerability has recently received a great deal of philosophical attention.¹³¹ Scholars have focused on different aspects of vulnerability, with some exploring the distinction between ontological and contingent vulnerabilities, others questioning whether vulnerability is a special source of moral obligations, and so on.¹³² Following Catriona Mackenzie and others' example,¹³³ we distinguish ontological vulnerabilities—vulnerabilities all human beings share in virtue of their embodied condition—from situated, socially constructed, or contingent vulnerabilities, which will be our primary focus.

Contingent vulnerabilities can result from structural conditions or individual differences.¹³⁴ Structural vulnerabilities derive from the fact that individuals are also members of groups, which enjoy varying degrees of privilege and experience varying levels of discrimination. For example, one might be vulnerable because of economic disadvantage, or on account of one's gender or sexual identity. Individual vulnerabilities, on the other hand, are those that exist irrespective of group memberships and may be the result of a person's particular history, circumstances, personality, habits, or practices. For instance, one person might be more vulnerable to online manipulation than another person who is similarly socially situated if the latter is more technologically savvy. Moreover, contingent vulnerabilities can overlap and exacerbate each other. Structural conditions can make a person more susceptible to having their individual weaknesses targeted and exploited. For example, many communities of color in America are more closely surveilled than white communities.¹³⁵ As a result, there is more information available

¹³¹ See generally VULNERABILITY, AUTONOMY, AND APPLIED ETHICS (Christine Straehle ed., 2016); VULNERABILITY: NEW ESSAYS IN ETHICS AND FEMINIST PHILOSOPHY (Catriona Mackenzie et al. eds., 2013).

¹³² See VULNERABILITY, AUTONOMY, AND APPLIED ETHICS, *supra* note 131, at 1–10. Some authors see an incompatibility between autonomy and vulnerability; we do not share this view but cannot go into detail here. The form of vulnerability that we analyze in the following does not follow precisely the taxonomies that these authors suggest. See, e.g., *id.*; VULNERABILITY: NEW ESSAYS IN ETHICS AND FEMINIST PHILOSOPHY, *supra* note 131.

¹³³ Catriona Mackenzie et al., *Introduction*, in VULNERABILITY: NEW ESSAYS IN ETHICS AND FEMINIST PHILOSOPHY 1 (Catriona Mackenzie et al. eds., 2013).

¹³⁴ *Id.*

¹³⁵ Alvaro M. Bedoya, *The Color of Surveillance: What an Infamous Abuse of Power Teaches Us About the Modern Spy Era*, SLATE (Jan. 18, 2016), <https://slate.com/technology/2016/01/what-the-fbis-surveillance-of-martin-luther-king-says-about-modern-spying.html> [<https://perma.cc/U27Z-H5YR>].

about the individual vulnerabilities of people of color living in those communities, potentially rendering them more vulnerable to manipulation.¹³⁶

Thus, vulnerabilities are not monolithic—they represent a variety of pressure-points, which can be exploited in a number of ways. Manipulative practices can reinforce and be reinforced by structural vulnerabilities, but they take advantage of the individual vulnerabilities of particular persons. Consequently, and as we have seen throughout, the more information a would-be manipulator has about a person’s specific vulnerabilities, the more capably they can exploit them.

Distinguishing between the various kinds of vulnerabilities helps emphasize a significant difference between online manipulation and the old-fashioned manipulative practices characteristic of some pre-digital advertising. Rather than aiming only to exploit vulnerabilities almost all of us share, as television advertisements and static billboards often attempt to do, online manipulation targets individuals, exploiting vulnerabilities specific to them.

C. Objections and Responses

There are at least two potential objections to our account of the autonomy harms manipulation threatens. First, one might object that it is difficult, if not impossible, to determine whether a person’s choices are made autonomously. As such, how could one determine in any particular case whether or not someone else has been manipulated?

In response, we emphasize our focus on manipulative *practices*, rather than manipulation per se. While it is indeed difficult to determine if someone has been manipulated in a particular instance (requiring information about the target’s beliefs, the intentions of the alleged manipulator, and so on), much can be surmised by investigating the influence strategy itself. Rather than attempting to determine whether the target of influence was moved or whether the influence was successfully hidden, we should attempt to determine whether the influencer was trying to conceal their efforts, whether the influence was intended to exploit the manipulee’s vulnerabilities, and to what extent the influence was targeted. Manipulative practices—characterized, as we have argued, by concealment, exploitation of vulnerabilities, and targeting—are cause for concern, regardless of whether they succeed in every instance. Indeed, adopting this perspective allows one to reframe the issue at

¹³⁶ The exploitation of structural vulnerabilities thus has discriminatory consequences. This problem has been discussed in the literature several times, but not with a focus on manipulation. *See, e.g.,* PASQUALE, *supra* note 1, at 38; *see generally* CATHY O’NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY 70 (2016).

a macro level: as we have argued throughout, information technology increasingly pervades and mediates nearly all aspects of our lives. If online manipulation becomes standard practice and thus an integral and indistinguishable part of our lives, it will become ever more difficult to identify specific cases of manipulation, and thus to think and talk about autonomous individuals who are responsible for their actions at all.

A second objection asks whether some alleged manipulees might not have behaved in the same way even if they hadn't been the subjects of manipulation. In other words, it points to a counterfactual component of claims to manipulation—that absent the manipulative intervention, the target would have decided differently.¹³⁷ Although some argue that this condition must be fulfilled for an influence to count as manipulative, i.e., the influence must change the way the target would have acted absent the influence, on our account manipulation interferes with the decision-making *process*, irrespective of its outcome, and one can therefore be manipulated even without such a change. While this may seem counterintuitive, it points to an issue of deep significance. The distinguishing feature of autonomy is that people want *to make choices for themselves*, and they would feel estranged from their choices if they knew they were being secretly driven toward a certain result. Autonomous choices and the decision procedures that yield those choices express who the decider is. As T.M. Scanlon argued:

I want to choose the furniture for my own apartment, pick out the pictures for the walls, and even write my own lectures despite the fact that these things might be done better by a decorator, art expert, or talented graduate student. For better or worse, I want these things to be produced by and reflect my own taste, imagination, and powers of discrimination and analysis. I feel the same way, even more strongly, about important decisions affecting my life in larger terms: what career to follow, where to work, how to live.¹³⁸

Scanlon shows that it is not just the result that matters when someone acts autonomously, but rather the choice-making, the processes of making up one's own mind. Autonomous choosers want to be in a position, at least in principle, to stand behind their choices with reasons of their own.

Finally, a third objection concerns a problem mentioned earlier: since autonomy is always socially situated, choices are always framed and conditioned by features of their social contexts. Indeed, social, cultural,

¹³⁷ See generally GOODIN, *supra* note 72.

¹³⁸ Thomas Scanlon, *The Significance of Choice*, in 7 THE TANNER LECTURES ON HUMAN VALUES 149, 180 (1986).

economic, and political contexts determine the sets of options from which one chooses in the first place. Thus, it can be difficult to draw a line between contextual influences and the influences of manipulators.

Imagine deciding to pursue a university education. In addition to determining whether or not someone is even given the option to pursue a university education, their social, cultural, political, and economic contexts may frame the decision in a particular way. For instance, girls may be discouraged from pursuing a technical education or dissuaded from training to become doctors because they are told that women are “better suited” to other careers. While this kind of treatment is sometimes called manipulative, it is not the kind of manipulation we have in mind.¹³⁹ And there are decisive differences between this kind of influence and online manipulation. Most important to our analysis, online manipulation is aimed precisely at *individual* choosers, and it is the specific information about each target that enables online manipulators to exploit that target’s vulnerabilities. As discussed in the previous section, information about groups can reveal individual vulnerabilities, and individual vulnerabilities can be exacerbated by structural conditions, which discriminate against groups. However, for the purposes of this analysis, these disparate sources of vulnerability can and must be distinguished from one another.

Furthermore, while it is sometimes difficult to define the specific beneficiaries of structural conditions (such as gender-stereotyping), the parties benefiting from manipulative online practices are easy to find. Critically analyzing structural conditions thus requires different terms and different theoretical tools than the forms of manipulation at issue here.¹⁴⁰

On the whole, we have seen that using autonomy as the normative lens through which to understand the harms of online manipulation helps foreground its individual, social, and political effects. The consequences are manifold. On one hand, manipulative practices undermine individual autonomy—people’s capacity for self-government, their ability to pursue their own goals. This is troubling in itself. But perhaps more worrying are the threats to *collective* self-government. When citizens are targets of online manipulation and voter decisions rather than purchase decisions are swayed by hidden influence, democracy itself is called into question. Add to this the fact that the tools of online manipulation are concentrated in only a few hands,

¹³⁹ See, e.g., Ann E. Cudd, *Adaptations to Oppression: Preference, Autonomy and Resistance*, in *PERSONAL AUTONOMY AND SOCIAL OPPRESSION: PHILOSOPHICAL PERSPECTIVES* 142 (Marina A.L. Oshana ed., 1st ed. 2015); see generally SERENE KHADER, *ADAPTIVE PREFERENCES AND WOMEN’S EMPOWERMENT* (2011).

¹⁴⁰ For literature on analyzing discrimination in liberal democracies, see, for instance, SALLY HASLANGER, *RESISTING REALITY: SOCIAL CONSTRUCTION AND SOCIAL CRITIQUE* (2012), and ANN CUDD, *ANALYZING OPPRESSION* (2006).

and it is easy to see how the nexus of influence and information technology stands to make already problematic power dynamics far worse.

VI. CONCLUSION

Our aim in this Article has been to put forward a systematic account of the nature of online manipulation, how it differs from other forms of influence, and what harms it threatens, to both individuals and society at large. The question that remains unanswered is, of course, how these threats can be addressed. Our Article offers a starting point and guide for answering that question.

First, online manipulation poses a threat across a wide variety of social contexts—from commercial contexts, to the workplace, to the political realm. Importantly, normative commitments regarding the types and degree of influence vary with each sphere. In the United States, for example, consumer decisions are generally considered less worthy of protection from outside influence than political decisions. With the rise of algorithmically mediated labor, protection for decision-making in the workplace will be the subject of much future debate. Therefore, we must pay close attention to where—in which spheres of life—influence is exerted, and what the effects of that influence mean for the individuals inhabiting them.

Second, because manipulation is, by our definition, *hidden*, combating it requires extra vigilance. The effects will often only become apparent after the harm has already been done. Further, the threat of online manipulation presents additional challenges to the predominant model of data regulation in the United States, which places the full burden of managing information flows and data practices on individuals.¹⁴¹ This model assumes that people are aware of the ways data about them is flowing and the risks and benefits associated with the data practices that implicate them. The emergence and proliferation of hidden, manipulative online practices pushes beyond the outermost limits of this approach. Individuals, unaware of the ways data is collected, aggregated, and used to influence them, simply cannot be left alone to fend off these incursions into their everyday decision-making. Bringing meaningful regulation to the digital sphere is no easy task. However, given that the tools

¹⁴¹ See, e.g., Daniel Solove, *Privacy Self-Management and the Consent Dilemma*, 126 HARV. L. REV. 1880 (2013). In Europe, data protection law works differently and follows slightly different principles, certainly since the implementation of the GDPR in May 2018. See Natali Helberger et al., *The Perfect Match? A Closer Look at the Relationship Between EU Consumer Law and Data Protection Law*, 54 COMMON MKT. L. REV. 1427, 1431 (2017) (“Data protection law operates on the basis of a number of central principles: lawfulness, fairness and transparency; purpose limitation; data minimization; accuracy; storage limitation; integrity and confidentiality; and accountability.”) (internal citations omitted). They also point out the potentially fruitful connection between the GDPR and European consumer law.

of online manipulation are largely controlled by a few powerful actors, we should begin our efforts by holding them to account.

Finally, many of the manipulative practices causing concern today evolved from advertising practices that have long been tolerated. If we are right about the nature of harm wrought by manipulative practices, we should no longer concede this ground. Although some forms of consumer targeting might be acceptable, we have suggested that even within the advertising sphere, there is a line that ought not to be crossed. Moreover, advertising may have received disproportionate attention from the research community because it is the visible tip of the iceberg. Understanding manipulative advertising provides insight into what might lie beneath. The use of information technology to facilitate manipulative practices greatly enhances their ability to shape our decision-making, raising anew questions about their ethical and political legitimacy.

THE GLOBAL “LAST MILE” SOLUTION: HIGH-ALTITUDE BROADBAND INFRASTRUCTURE

Snezhana Stadnik Tapia*

CITE AS: 4 GEO. L. TECH. REV. 47 (2019)

TABLE OF CONTENTS

I. INTRODUCTION	48
II. HISTORICAL SOLUTIONS TO THE LAST MILE PROBLEM.....	52
A. Addressing the Market Efficiency Gap: Liberalization in Fixed Wireline Services	54
B. Addressing the Access Gap: National Universal Service Mechanisms	58
1. <i>Universal Service: History, Priorities, and Modes of Expansion</i>	59
2. <i>Universal Service Financing Mechanisms</i>	64
3. <i>Failure of Universal Service Funds</i>	67
III. MOVING BEYOND LIBERALIZATION AND UNIVERSAL SERVICE FUNDS: INNOVATIVE GLOBAL BROADBAND INFRASTRUCTURE PROJECTS	70
A. High-Altitude Internet Infrastructure Projects and Applicable Regulatory Frameworks.....	73
1. <i>Airspace Connectivity: Internet-beaming Balloons and Drones</i>	73
a. Loon’s Connectivity Project and the Regulatory Framework	76
2. <i>Outer Space Connectivity: The Global Broadband Space Race</i>	81

* Associate at Sidley Austin LLP, focusing on international arbitration and privacy and cybersecurity matters; J.D., New York University School of Law, 2018. I would like to thank Professor Benedict Kingsbury and Thomas Streinz for their help in developing this paper, as well as participants in the Spring 2017 "International Law of Google" colloquium for their valuable insights and feedback. Additionally, thank you to my husband, Josue Tapia, for his encouragement and unwavering support during the writing process.

a. Connectivity via Mega-Constellations and the Regulatory Framework	84
B. Disruptive Potential of the Innovations: Changing the Economics of Broadband Access?.....	90
C. Leapfrogging Stages of Development?.....	96
IV. EVALUATING THE POTENTIAL IMPACT OF BROADBAND INFRASTRUCTURE USING THE CAPABILITY APPROACH.....	101
A. Exploring the Relationship Between Infrastructure and Socioeconomic Development	103
B. Applying the Capability Approach to Broadband Infrastructure Projects.....	108
1. <i>Key Concepts</i>	109
2. <i>From Mere Availability to Genuine Access</i>	111
3. <i>How ICTs Impact Capabilities</i>	116
4. <i>Community-led Development under the Capability Approach</i> . ..	118
V. CONCLUSION.....	122

I. INTRODUCTION

High-speed Internet access, otherwise known as broadband,¹ is considered essential for partaking in the 21st-century economy; the Internet today is considered as important as road and energy infrastructure in terms of its potential to enhance socioeconomic development.² Broadband, a subset of telecommunications infrastructure that includes wire-based and wireless communications networks, is currently a priority in most countries aiming to bridge the “digital divide,”³ or the phenomenon of being excluded from the information society, usually for lack of availability and affordability.⁴ Well-documented in both developed and developing countries, the digital divide is undeniably global, even after factoring in the increased rates of access in developing countries to less costly wireless-based mobile broadband services,

¹ What constitutes high-speed is usually determined by individual countries and the International Telecommunication Union. In the United States, the Federal Communications Commission broadly defines broadband as high-speed Internet access that is constantly on and faster than traditional dial-up. See BERND HOLZNAGEL ET AL., STRATEGIES FOR RURAL BROADBAND AN ECONOMIC AND LEGAL FEASIBILITY ANALYSIS 15 (2010); *Types of Broadband Connections*, FCC (June 23, 2014), <https://www.fcc.gov/general/types-broadband-connections> [<https://perma.cc/3BW7-U3F5>].

² See HOLZNAGEL ET AL., *supra* note 1, at 7.

³ See Dwayne Winseck, *The Geopolitical Economy of the Global Internet Infrastructure*, 7 J. INFO. POL’Y 228, 256–57 (2017). The number of national broadband plans has increased from 38 in 2008 to 151 in 2016. See *id.*

⁴ See HOLZNAGEL ET AL., *supra* note 1, at 17 (noting that the term was initially used to describe the gap “between information ‘haves’ and ‘have-nots’”).

as opposed to fixed wireline services, such as DSL or fiber.⁵ In 2016, the proportion of the population covered by a mobile broadband network reached eighty-four percent globally and sixty-seven percent in rural areas; LTE or faster networks covered about one half of the global population.⁶ Unsurprisingly, broadband is an important item on the global agenda. The United Nation's (UN) Sustainable Development Goals (SDGs) acknowledge the importance of universal access to broadband, encouraging public and private actors alike to bridge the global digital divide by tackling Information and Communications Technology (ICT) infrastructure underdevelopment.⁷

Historically, the lack of universal access to telecommunications services was addressed nationally—via legislation that enacted pro-competitive policies in the telecommunications sector and universal service mechanisms administered by national communications agencies to spur infrastructure development in high-cost areas. These public interventions were considered warranted to address the “last mile problem,” an expensive investment for telecommunications providers also referred to as the “local loop.”⁸ The local loop entails connecting every home to a network provider's local office and in turn, the backhaul network, which is “large-scale wireline infrastructure” that “permits interconnection between the carrier's network

⁵ See Andrew Perrin, *Digital Gap Between Rural and Nonrural America Persists*, PEW RESEARCH CENTER (May 31, 2019), <http://www.pewresearch.org/fact-tank/2017/05/19/digital-gap-between-rural-and-nonrural-america-persists/> [https://perma.cc/AM36-SNUW]. One study notes at least a ten percent gap when comparing fixed broadband connections between rural and urban American households. *Id.* In the Least Developed Countries (LDCs), currently comprising 47 countries, an entry-level fixed-broadband subscription is about 2.6 times more expensive than an entry-level mobile-broadband subscription. See INT'L TELECOMM. UNION, *ICT FACTS AND FIGURES 5* (July 2017), <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf> [https://perma.cc/MGE8-B9LS]; *Least Developed Countries (LDCs)*, UNITED NATIONS, <https://www.un.org/development/desa/dpad/least-developed-country-category.html> [https://perma.cc/XR2G-WAHC].

⁶ INT'L TELECOMM. UNION, *MEASURING THE INFORMATION SOCIETY REPORT 77* (2016), <http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2016/MISR2016-w4.pdf> [https://perma.cc/XAY8-GJQK].

⁷ See generally, *About the Sustainable Development Goals*, UNITED NATIONS, <https://www.un.org/sustainabledevelopment/sustainable-development-goals/> [https://perma.cc/F8TC-ZL6M]. In 2015, the UN General Assembly adopted 17 SDGs as an integral part of the 2030 Agenda for Sustainable Development, and two of the SDGs relate to bridging the global digital divide. *Id.* The first applicable SDG is 9, focusing on infrastructure, industrialization, and innovation. The other SDG that aims to bridge the technological divide is 17, emphasizing the strengthening of global partnerships for sustainable development. *Id.*; see also *infra* Part IV.A.

⁸ See Krishna Jayakar, *Universal Service*, in . . . AND COMMUNICATIONS FOR ALL: A POLICY AGENDA FOR A NEW ADMINISTRATION 181, 194 (Amit M. Schejter ed., 2009).

and other networks.”⁹ For a local loop project to be financially viable, the prospective revenues must be greater than the costs, making population density a crucial factor in infrastructure provision.¹⁰ Faced with chronic underprovision in rural areas, extending communications networks to the last mile historically required public intervention and continues to do so today.

Although pro-competitive policies and universal service mechanisms boosted communications infrastructure deployment worldwide, the broadband access gap persists in many countries. Today, telecommunications providers are especially leery of investing in networks in the developing world where “most governments lack the financial resources on their own to make the diffusion of the Internet a major priority.”¹¹ On a global scale, this makes rural communities in the developing world the ultimate last mile. Nevertheless, the last mile problem has motivated some private actors to innovate, hoping to take advantage of the tremendous opportunity to serve unreached markets. With national interventions falling short, some companies are unintentionally following the SDGs; many non-traditional “telecommunications” companies are deploying innovative broadband infrastructure to blanket the globe with high-altitude connectivity. Loon (Google’s sister company), Facebook, and satellite companies like SpaceX have announced connectivity agendas and broadband infrastructure projects, which include the development of potentially disruptive satellite and telecommunications infrastructure—namely, Internet-beaming balloons, drones, and low Earth orbit (LEO) satellite mega-constellations.¹²

Besides the promise of new markets and additional profits, it is important to inquire about the ethos underpinning the companies’ connectivity infrastructure projects. How should the impact of such projects, which purport to solve the lack of worldwide connectivity, be measured? When extending Internet access to billions in rural and remote areas of the world, the companies with connectivity agendas often assume that broadband infrastructure will lead

⁹ BRETT M. FRISCHMANN, *INFRASTRUCTURE: THE SOCIAL VALUE OF SHARED RESOURCES* 212–13 (2012) [hereinafter FRISCHMANN, *INFRASTRUCTURE*] (“Backhaul is a large-scale wireline infrastructure that functions similar to interexchange networks [for wired telephony], and even arterial highways in the transportation infrastructure. It permits interconnection between the carrier’s network and other networks (including other wireless carriers or the Internet).”); see also Marvin Ammori, *Competition and Investment in Wireline Broadband*, in . . . AND COMMUNICATIONS FOR ALL: A POLICY AGENDA FOR A NEW ADMINISTRATION 85 (Amit M. Schejter ed., 2009).

¹⁰ See JIM W. HALL ET AL., *THE FUTURE OF NATIONAL INFRASTRUCTURE: A SYSTEM-OF-SYSTEMS APPROACH* 181 (2016) [hereinafter HALL ET AL., *THE FUTURE OF NATIONAL INFRASTRUCTURE*].

¹¹ Geoffrey S. Kirkman, *Out of the Labs and Into the Developing World: Using Appropriate Technologies to Promote Truly Global Internet Diffusion*, 2 J. HUM. DEV. 194 (2001).

¹² See Section III.A, *infra*.

to a number of positive developments.¹³ Although socioeconomic development is typically the goal for most infrastructure investment, it should not be deemed a natural byproduct of broadband infrastructure deployment. As such, this Article cautions against the “build it and they will come” approach employed by many technology enthusiasts and techno-solutionists—who may enter emerging markets often unaware of the challenges of implementing technology in developing countries.¹⁴ There are limitations to technocratic approaches and solutions to long-standing economic and development problems. Instead of tackling the engineering challenges first, regulatory and market strategies second, and development outcomes last, this Article recommends incorporating Amartya Sen’s capability approach from the outset to anticipate and measure the potential development impact of newly available connectivity infrastructure. By sharpening the focus from potential macro- to micro-level considerations, the capability approach may alert connectivity companies to how technology may positively and negatively impact individuals’ lives and substantive freedom—and prompt collaboration with existing development actors in this space to mitigate negative impacts.

This Article is organized as follows. Part II explains the reasons for communications infrastructure underdevelopment historically, taking into account the myriad ways governments, usually through national universal service mechanisms, have attempted to correct the underprovision and

¹³ ERIC SCHMIDT & JARED COHEN, *THE NEW DIGITAL AGE: TRANSFORMING NATIONS, BUSINESSES, AND OUR LIVES* 5 (2014). With rising connectivity and the spread of new technology, Google executives believe the future will “help reallocate the concentration of power away from states and institutions and transfer it to individuals.” Besides political empowerment, connectivity is believed to yield social and economic empowerment. *See id.*; *see also* Part IV, *infra*.

¹⁴ Such tech-solutionism and optimism have been the cause of much consternation; Silicon Valley’s ability to “save the world” has been a matter of debate. *See, e.g.*, Jason Henry, *Hey Silicon Valley, John Kerry Wants You to Help Save the World*, WIRED (Nov. 1, 2016), <https://www.wired.com/2016/11/hey-silicon-valley-john-kerry-wants-help-save-world/> [<https://perma.cc/3P9E-Q9UZ>]; Charles Kenny & Justin Sandefur, *Can Silicon Valley Save the World?*, FOREIGN POL’Y (June 24, 2013), <https://foreignpolicy.com/2013/06/24/can-silicon-valley-save-the-world/> [<https://perma.cc/TDG6-T7NM>]; Kevin Maney, *Why the World Hates Silicon Valley*, NEWSWEEK (June 9, 2016), <https://www.newsweek.com/2016/06/17/silicon-valley-takeover-468182.html> [<https://perma.cc/GS7K-8L9J>]; Pankaj Mishra, *Can Silicon Valley Save the World?*, BLOOMBERG (Sept. 29, 2015), <https://www.bloomberg.com/opinion/articles/2015-09-30/silicon-valley-can-t-save-the-developing-world> [<https://perma.cc/L9RP-UNQZ>]; *Is Silicon Valley Saving the World or Just Making Money?*, N.Y. TIMES (July 22, 2015), <https://www.nytimes.com/roomfordebate/2015/07/22/is-silicon-valley-saving-the-world-or-just-making-money> [<https://perma.cc/R2UT-5999>]; Claire Cain Miller, *Can Technology Save the World? Experts Disagree*, N.Y. TIMES, (May 2, 2014), <https://www.nytimes.com/2014/05/03/upshot/can-technology-save-the-world-experts-disagree.html> [<https://perma.cc/NBZ2-NZLR>].

positing why this opportunity to create global broadband infrastructure has surfaced. In essence, this portion of the paper explains the last mile problem that innovative infrastructure projects purport to solve. Part III then describes the broadband infrastructure projects, the consequences of multi-jurisdictional regulatory complexities for bringing the projects to market, and the disruptive potential of the infrastructure to change the economics of broadband access and provision. Lastly, Part IV considers whether the companies are indeed solving the last mile problem beyond mere provision. Accordingly, the potential impacts of Internet access are surveyed using Amartya Sen's capability approach, which seeks to place the individual and his or her freedom at the center of development.

II. HISTORICAL SOLUTIONS TO THE LAST MILE PROBLEM

Encouraging private investment in infrastructure is no small feat when taking into account the economics of infrastructure financing. Although there is no separate field of infrastructure study or settled definitions of infrastructure,¹⁵ the term generally brings to mind large-scale physical resources, usually developed for public consumption.¹⁶ Many infrastructure resources have a similar cost structure, entailing high fixed costs of initial production but low and decreasing marginal costs for each additional use.¹⁷ Fixed costs for infrastructure services are significant, needing to be spread among a large number of consumers for the investment to be deemed profitable and worthwhile.¹⁸ Duplicating resources in some infrastructure sectors is inefficient, leading to the market being supplied by the one firm that

¹⁵ See HALL ET AL., *THE FUTURE OF NATIONAL INFRASTRUCTURE*, *supra* note 10, at 4 (“[A]nalysis of infrastructure . . . is complicated by the absence of a single comprehensive, functional and practical definition of infrastructure.”).

¹⁶ See FRISCHMANN, *INFRASTRUCTURE*, *supra* note 9, at 3.

¹⁷ See Jerry A. Hausman, *Valuing the Effect of Regulation on New Services in Telecommunications*, BROOKINGS PAPERS ON ECON. ACTIVITY, 1997, at 27 (discussing implications of prices being set at marginal cost); R. S. Khemani & D. M. Shapiro, OECD, GLOSSARY OF INDUSTRIAL ORGANISATION ECONOMICS AND COMPETITION LAW 62 (1993) [hereinafter OECD GLOSSARY], <http://www.oecd.org/regreform/sectors/2376087.pdf> (“[N]atural monopolies are characterized by steeply declining long-run average and marginal-cost curves such that there is room for only one firm to fully exploit available economies of scale and supply the market.”) [<https://perma.cc/8JAA-V8V8>].

¹⁸ See Brett M. Frischmann, *An Economic Theory of Infrastructure and Commons Management*, 89 MINN. L. REV. 946–47, n. 104 (2005) [hereinafter Frischmann, *Economic Theory of Infrastructure*].

has achieved economies of scale.¹⁹ Such natural monopolies commonly exist in industries relating to public consumption, such as energy, transportation, and telecommunications.²⁰

The above holds true for telecommunications infrastructure in both developed and developing countries. For many years, economists treated telephone networks as a token example of natural monopolies: telecom was a sector in which competition was unlikely due to large fixed costs, as well as decreasing average and marginal costs, resulting in economies of scale.²¹ Moreover, to this day, the most expensive investment for telecommunications providers is the “local loop,” or the last mile of the network that connects every home to a network provider’s local office, which is connected to the backhaul network.²² For a last mile project to be financially viable, the prospective revenues must be greater than the investment costs; in other words, the large fixed capital costs common to infrastructure resources need to be spread over many potential customers, making population density an important factor for telecom providers when considering market expansion.²³ That is why coverage gaps and insufficient infrastructure upgrades persist in low-density areas: rural markets are the least profitable.²⁴

In order to increase public access to and usage of communications infrastructure and solve the last mile problem, the regulatory response in most countries has been two-fold, sometimes pursued simultaneously and other times discretely: (1) address the *market efficiency gap* using liberalization policies to enhance competition in the sector; and (2) address the *access gap* by designing financing mechanisms in compliance with universal service mandates.²⁵ These two gaps and corresponding regulatory responses are discussed below.

¹⁹ See *id.* at 929. See also Mariana Mota Prado, *Regulatory Choices in the Privatization of Infrastructure*, in PRIVATE SECURITY, PUBLIC ORDER: THE OUTSOURCING OF PUBLIC SERVICES AND ITS LIMITS 123 (Simon Chesterman & Angelina Fisher eds., 2009) (discussing rationales for competition).

²⁰ See HALL ET AL., THE FUTURE OF NATIONAL INFRASTRUCTURE, *supra* note 10, at 5 (noting infrastructure resources show characteristics of public goods); OECD GLOSSARY, *supra* note 17, at 62.

²¹ See OECD, THE DEVELOPMENT OF FIXED BROADBAND NETWORKS, 11 (2014), <http://dx.doi.org/10.1787/5jz2m5mlb1q2-en> [hereinafter OECD, FIXED BROADBAND NETWORKS] [<https://perma.cc/Y6T7-MS3R>]. See also FRISCHMANN, INFRASTRUCTURE, *supra* note 9, at 213 (citing three reasons why “[t]he natural monopoly designation made sense for a while”).

²² Ammori, *supra* note 9, at 85; FRISCHMANN, INFRASTRUCTURE, *supra* note 9, at 212.

²³ See HALL ET AL., THE FUTURE OF NATIONAL INFRASTRUCTURE, *supra* note 10, at 181.

²⁴ See *id.* at 191.

²⁵ Luis D. Emiliani, *Universal Service and Universal Access to Telecommunications: A Review*, 36TH RES. CONF. ON COMM., INFO. & INTERNET POL’Y, 10 (Sept. 2008) (explaining the *market efficiency gap* versus the *access gap* with a chart).

Before proceeding, it is important to note that the last mile problem, and the resultant policies that attempt to solve it, historically focused on voice telephony and copper wire deployment. Nevertheless, the same principles discussed in this Part endure and are applicable to broadband infrastructure development.²⁶ This is because Internet access and data flow relies on the same mediums as voice telephony: cables (usually copper wire or optical fiber) and electromagnetic waves (for satellite, wireless, and mobile networks). Moreover, broadband development falls within the scope of the regulatory authority in each country that deals with telecommunications. Telecommunications regulatory agencies exercise authority over physical wired networks, as well as determine spectrum allocations required for wireless networks.²⁷ In other words, regulatory telecommunications agencies oversee the physical layer of communications infrastructure, which includes the Internet.

A. Addressing the Market Efficiency Gap: Liberalization in Fixed Wireline Services

During the 20th century, governments dealt with the last mile problem by requiring cross-subsidies to keep urban and rural rates equal. In 1934, instead of nationalizing the then unregulated AT&T, Congress established the Federal Communications Commission (FCC) to regulate AT&T's monopoly in exchange for universal access included in the Communications Act of 1934: "to make available, so far as possible, to all the people of the United States . . . a rapid, efficient, nationwide, and worldwide wire and radio communication service with adequate facilities at reasonable charges."²⁸ The term universal access then meant that "everyone gained access at a (somewhat) flat fee, regardless of his or her geographical location . . . AT&T was required to amortize connection charges across all customers."²⁹ In other countries, state-owned (rather than privately-owned) telecommunications companies similarly pursued cross-subsidies among rural and urban areas.³⁰ Eventually, governments realized that although monopolies (if regulated) could provide

²⁶ *Id.* at 9 (noting convergence of services over TCP/IP, expanding the definition of universal access policies to include not only voice services but broadband Internet access).

²⁷ 47 U.S.C. § 151 (2018). With the passing of the Communications Act of 1934, Congress created the FCC, an independent agency with a mandate to regulate interstate communications by radio, television, wire, satellite, and cable. The name of the act was slightly changed when it was amended decades later to the *Telecommunications Act* of 1996. *Id.*

²⁸ *Id.*

²⁹ Ted G. Lewis, *Telecommunication: Critical Infrastructure Protection*, in COMMUNICATIONS AND INFORMATION INFRASTRUCTURE SECURITY 6 (John G. Voeller ed., 2015).

³⁰ Emiliani, *supra* note 25, at 13–14.

universal access, the consequence was decreased innovation and underprovision absent the incentives provided by competition.

In the last thirty years, governments have stepped in by liberalizing the telecommunications market and enforcing pro-competitive regulations.³¹ In countries where telecom liberalization policies were pursued, sectoral authorities were also established to regulate the sector and “safeguard [the] public interest in the provision of these essential services.”³² The U.S. Congress, for instance, passed the Telecommunications Act of 1996 “to provide for a pro-competitive, de-regulatory national policy framework” in a market that was *already* occupied by the private sector.³³ In India and some countries in Africa the pro-competitive policies sanctioned the first move away from full state ownership, encouraging the entry of private sector operators and partially-privatized national monopolies. In India specifically, liberalization and private investment were recommended after the government realized it had insufficient resources to fulfill its universal service targets codified in the National Telecom Policy of 1994,³⁴ and the government pursued a strategy of “encouraging the entry of multiple Indian private sector operators with foreign partners in different regions and market segments.”³⁵ In Kenya, the government also struggled to meet its universal service obligations under the Kenya Communication Act (KCA) of 1998, which spelled out a pathway towards liberalization. Although state-run Telkom Kenya was to hold an interim monopoly for five years, it needed the support of other regional operators to fulfill the universal service obligation spelled

³¹ Emmanuel O. Arakpogun, Roseline Wanjiru & Jason Whalley, *Impediments to the Implementation of Universal Service Funds in Africa—A Cross-Country Comparative Analysis*, 41 TELECOMM. POL’Y 617, 617–20 (2017) [hereinafter *Universal Service Funds in Africa*].

³² Paolo Gerli, Marlies Van der Wee, Sofie Verbrugge & Jason Whalley, *The Involvement of Utilities in the Development of Broadband Infrastructure: A Comparison of EU Case Studies*, 42 TELECOMM. POL’Y 726, 727 (2018).

³³ H.R. REP. NO. 104-458, at 1 (1996) (Conf. Rep.) (noting that the legislation aimed “to accelerate rapidly private sector deployment of advanced telecommunications and information technologies and services to all Americans by opening all telecommunications markets to competition.”).

³⁴ Krishna Jayakar & Chun Liub, *Universal Service in China and India: Legitimizing the State?*, 38 TELECOMM. POL’Y 186, 191 (2014). (“With respect to universal service, the National Telecom Policy aimed to provide telephone on demand by 1997, connect all villages to the telephone network by 1997, provide a public call office (PCO) for every 500 persons in urban areas, and introduce value added services in India, on par with those available internationally.”).

³⁵ Stephen McDowell & Jenghoon Lee, *India’s Experiments in Mobile Licensing*, 27 TELECOMM. POL’Y 371, 373 (2003).

out in the UN's Millennium Development Goals and the KCA.³⁶

Although some initially did not support liberalization, believing that universal service would not be sustainable in a competitive market absent cross-subsidies in monopolists' prices, it was generally agreed upon that competitive policies resulted in greater access by making services cheaper.³⁷ Coupled with requirements for carriers to serve underserved populations, these policies regularly increased coverage in the U.S. and other countries.³⁸ For instance, in Kenya, competition led some telecommunications companies to develop areas previously considered uneconomical.³⁹ Today, a total of 108 World Trade Organization (WTO) members have made commitments to facilitate trade in telecommunications services, and many of them have had to give up price support regimes for telecom operators and devise other methods of support.⁴⁰

Pro-competitive policies increased coverage and brought significant changes to the once monopolized landline market, yet advances in mobile phone technologies and competition between mobile and fixed-line services

³⁶ Monica Kerretts, *ICT Regulation and Policy at a Crossroads: A Case Study of the Licensing Process in Kenya*, 5 SOUTHERN AFR. J. INFO. & COMM. 49, 52 (2005); MILLENNIUM DEVELOPMENT GOALS AND BEYOND 2015, UN, <http://www.un.org/millenniumgoals/global.shtml> ("In cooperation with the private sector, make available benefits of new technologies, especially information and communications.") [<https://perma.cc/F4NE-2S8R>].

³⁷ See Carolyn Gideon & David Gabel, *Disconnecting: Understanding Decline in Universal Service*, 35 TELECOMM. POL'Y 737, 738 (2011). In India, telecommunications workers' unions and the general public opposed privatization. The promise of universal service made "privatization . . . more palatable: the National Telecom Policy explicitly stated that the newly licensed telecommunications operators would be required to maintain 'a balance in their coverage between urban and rural areas' as well as acquiesce to tariff regulation and revenue sharing arrangements." Jayakar & Liub, *supra* note 34, at 194.

³⁸ FCC, *CONNECTING THE GLOBE: A REGULATOR'S GUIDE TO BUILDING A GLOBAL INFORMATION COMMUNITY*, at i (1999) [hereinafter *CONNECTING THE GLOBE*], <https://transition.fcc.gov/connectglobe/regguide.pdf> [<https://perma.cc/VU7C-T4G2>]; Cristina Casanueva-Reguart, *Institutions, Telecommunications Reform, and Universal Service Policy in Mexico (1990–2014)*, 9 INT'L J. COMM. 2092, 2095 (2015) (noting that competition is now viewed as essential to the development of modern telecommunications infrastructure).

³⁹ Omae Malack Oteri, Langat Philip Kibet & Ndung'u Edward N., *Mobile Subscription, Penetration and Coverage Trends in Kenya's Telecommunication Sector*, 4 INT'L J. ADVANCED RES. ARTIFICIAL INTELLIGENCE 1, 5 (2015).

⁴⁰ See Do Manh Thai & Morten Falch, *Universal Service in Vietnam: An Institutional Approach*, 42 TELECOMM. POL'Y 323, 329 (2018). For example, in the United States-Colombia and United States-Peru trade agreements, "the universal service clause (article 14.8) indicates that each party has the right to manage their USO definition in a competitively-neutral fashion, and Clause 14.4 forbids each party to engage in cross subsidization." Emiliani, *supra* note 25, at 16.

contributed even more toward achieving universal service.⁴¹ Today, mobile services account for about forty percent of the global telecommunications market, with mobile subscribers outnumbering fixed telephone line users by more than two to one.⁴² Competition among mobile operators has helped narrow the digital divide: it reduced tariff rates and increased the affordability of communication services for many.⁴³ For instance, both India and China have been ranked among the fastest growing mobile markets in the world during the last two decades as a result of competition policies.⁴⁴ Similarly, with four mobile network operators in the last two decades, Kenya now has one of the highest mobile penetration rates in Africa.⁴⁵ The dramatic growth in mobile phone usage in Kenya has increased Internet penetration, with most Kenyans accessing the Internet via mobile phone rather than fixed wireline connections.⁴⁶ International comparisons identify few countries where liberalization and advances in mobile telephony did not dramatically increase access to communication services.⁴⁷

For those countries that pursued liberalization policies and saw advances in mobile innovations, the market efficiency gap in the telecommunications market has mostly closed, but only in commercially viable locations. There are and continue to exist some remote and high cost areas where market forces alone do not guarantee satisfactory deployment of infrastructure, and universal service remains an unrealized promise.⁴⁸ These areas, which are beyond the “market efficiency frontier,” point to “access

⁴¹ *Telecommunications Services*, WORLD TRADE ORGANIZATION, https://www.wto.org/english/tratop_e/serv_e/telecom_e/telecom_e.htm [<https://perma.cc/V4HC-7QU2>].

⁴² *Id.* In some Latin American countries, the “number of mobile subscribers has surpassed the number of fixed lines installed. Due to market liberalization and the low number of restrictions to handset acquisition and activation, mobile telephony has become crucial in closing the market gap in urban and rural areas.” Emiliani, *supra* note 25, at 19.

⁴³ Oteri et al., *supra* note 39, at 6.

⁴⁴ Jayakar & Liub, *supra* note 34, at 194.

⁴⁵ In 2016, Kenya had a 90% mobile penetration rate, more than the African continent average of 76.2%. Idi Jackson Mdoe & George Kariuki Kinyanjui, *Mobile Telephony, Social Networks and Credit Access: Evidence from MSMEs in Kenya*, 6 COGENT ECON. & FIN. 1, 4 (2018).

⁴⁶ In 2011, about 17 million Kenyans used the Internet (roughly forty three percent of the population). Though more than 6 million were subscribed to the Internet, about 23,000 subscribed to fixed fiber-optic, with the rest receiving access via mobile phones. GEORGE NYABUGA & NANCY BOOKER, OPEN SOC’Y FOUNDS., MAPPING DIGITAL MEDIA: KENYA 1, 17 (2013).

⁴⁷ Casanueva-Reguart, *supra* note 38, at 2094. Teledensity in mobile services in Africa is greater than in Mexico, and Mexico’s teledensity is below the average for Latin America. *Id.* Teledensity in mobile services was 88.3% in Mexico, 104.6% in Africa, and 113.0% in Latin America. *Id.* Mexico’s teledensity is low when taking into account countries in the region with a similar level of wealth. *Id.*

⁴⁸ CONNECTING THE GLOBE, *supra* note 38, at i.

gaps.”⁴⁹ Another term to capture this phenomenon is the “territorial divide,” or a gap in infrastructure availability between rural and urban areas.⁵⁰ For example, in Mexican rural areas especially, mobile density is low and residential landline availability is even more limited.⁵¹ Additionally, many parts of Kenya also have no mobile coverage, especially the arid and semi-arid areas, pointing to a general trend in Africa, India, and elsewhere: liberalization and mobile telephony have drastically reduced coverage gaps in urban areas and decreased prices, but the same cannot be said of less densely populated areas, where people continue to be unserved or underserved.⁵² This reflects the commercial orientation of network operators, where “[m]anagement decisions made by firms regarding the pace and geographic coverage of network roll-out may be guided by attempts to reach the highest density of customers, and those with the greatest purchasing power.”⁵³ The exorbitant costs of network infrastructure continue to present a challenge for universal service by discouraging investment in high-cost areas. To fully close this access gap, a different kind of public intervention is needed and has typically come in the form of various universal service obligations and mechanisms.

B. Addressing the Access Gap: National Universal Service Mechanisms

To ensure universal service, governments have typically turned to policies that aim to reduce *both* market efficiency and access gaps. This results in pro-competitive policies complemented by universal service mechanisms whose aim is to improve access and social welfare, as well as mitigate the problem of the digital divide between commercially viable and non-viable locations.⁵⁴ Not only are there many different variations of universal service visions and definitions, the implementation and funding methods vary by country. The analysis below examines some historic and present-time

⁴⁹ Casanueva-Reguart, *supra* note 38, at 2095–97; Emiliani, *supra* note 25, at 11, 18.

⁵⁰ Emiliani, *supra* note 25, at 18.

⁵¹ Casanueva-Reguart, *supra* note 38, at 2094. In poorer states of Mexico, where 15% of the country’s population resides, residential landline availability is more limited (25.5%), and these states also have low mobile density (71.1%). *Id.*

⁵² Oteri et al., *supra* note 39, at 6 (“With a national coverage of about 77% of the population, the mobile industry invariably covers over 31 million people in the country. However, the 38% geographic coverage implies that many parts of the country are not covered, especially the arid and semi-arid areas”); *Universal Service Funds in Africa*, *supra* note 31, at 618; Jayakar & Liub, *supra* note 34, at 191 (“[A]nalysis by the Telecommunications Regulatory Authority of India [in 2004] showed that telecom licensees were far from achieving their targets. Providers often disdained to serve rural areas due to an assumption that they were not viable markets.”).

⁵³ McDowell & Lee, *supra* note 35, at 372.

⁵⁴ *Universal Service Funds in Africa*, *supra* note 31, at 622; *see also* Casanueva-Reguart, *supra* note 38, at 2095.

definitions of universal service, government priorities in terms of which services and modes of expansion to prioritize, as well as popular financing mechanisms. The Part ends with a discussion of common reasons for failure of universal service policies and programs to adequately address the access gap worldwide.

1. *Universal Service: History, Priorities, and Modes of Expansion*

The concept of universal service, or the “universal service obligation,” has existed and evolved since the 1900s.⁵⁵ The concept was initially proposed by AT&T president Theodore Newton Vail, who believed that a telecommunications network should reach all *households* of a country.⁵⁶ At that time, the concept was used as a defense for maintaining a monopoly over telephone services in the United States.⁵⁷ The principle of universal service was later defined in the U.S. Communications Act of 1934 as “making available, so far as possible, to all the *people* of the U.S., rapid, efficient, nationwide and worldwide wire and radio communication services with adequate facilities at reasonable charges.”⁵⁸ In the early 1970s, Bell System relied on a cross-subsidy method to implement its universal service obligation and subsidize the last mile.⁵⁹ After liberalization of telecom markets in the United States, the notion that the last mile should be in some way subsidized spread around the world.

The problem with analyzing universal service is the lack of a commonly accepted definition. Though the concept has been around for decades, the definition continues to change as certain services and beneficiaries are prioritized.⁶⁰ Still, the concept in both developed and developing countries has the same general meaning: wider access to telecommunications services in a manner that, according to the International

⁵⁵ See Emiliani, *supra* note 25, at 3.

⁵⁶ See *id.*

⁵⁷ See *id.*

⁵⁸ The Communications Act, 47 U.S.C. § 151 *et seq.* (1934).

⁵⁹ See Gary Madden, *Economic Welfare and Universal Service*, 34 TELECOMM. POL’Y 110, 111 (2010).

⁶⁰ Jayakar & Liub, *supra* note 34, at 187. See generally, James Alleman, Paul Rappoport & Aniruddha Banerjee, *Universal Service: A New Definition?*, 34 TELECOMM. POL’Y 86, 86–91 (2010).

Telecommunication Union (ITU), is available, accessible, and affordable.⁶¹ Today, there are differences among countries about which services should be available, accessible, and affordable.

Although some countries emphasize access to telephone service in every home or aim to provide a public telephone within a given distance, others focus on access to fast Internet.⁶² The service to be prioritized once depended on a country's wealth and stage of telecommunications development. For example, in Canada, Europe, and the United States, historically, there has been a greater emphasis on ensuring universal access to broadband.⁶³ However, recognizing that universal connectivity through broadband infrastructure is vital in today's digital era, many developing countries are now including broadband in their universal service definitions and have started adopting National Broadband Plans (NBPs) to articulate their growth plans and broadband targets.⁶⁴ The number of NBPs increased from 38 in 2008 to 151 in 2016.⁶⁵ Some NBPs focus on the design, construction, and financing of wholly new high-speed broadband networks and fiber deployment, while others articulate a more medium-term goal of improving existing telecommunications infrastructure.⁶⁶ For example, the Nigerian NBP

⁶¹ More broadly, the ITU defines universal service as making telecommunication services (of specified quality and in light of national conditions) available at affordable prices to currently unserved potential users. See Emiliani, *supra* note 25, at 7; Madden, *supra* note 59, at 111; Thai & Falch, *supra* note 40, at 324; *Universal Service Funds in Africa*, *supra* note 31, at 619; *Universal Service*, in *Telecommunications Services: Glossary*, WORLD TRADE ORGANIZATION, https://www.wto.org/english/tratop_e/serv_e/telecom_e/tel12_e.htm [<https://perma.cc/Q8YD-NS34>].

⁶² *Universal Service Funds in Africa*, *supra* note 31, at 626; *Telecommunications Services: Glossary*, *supra* note 61.

⁶³ Broadband describes “high-speed, 'always-on' access to the Internet,” enabling enhanced online shopping and banking experiences, downloading and sharing larger files, as well as videoconferencing. Nathaniel E. Uramaab & Osita Ogbuc, *Evaluating Consumer Perception and Willingness to Pay for Broadband in Nigeria*, 42 TELECOMM. POL'Y 421, 423 (2018) [hereinafter *Willingness to Pay for Broadband in Nigeria*]. In Europe, the attention has been on increasing household access to Internet connections above 100 Mbps. See *Europe 2020: A European Strategy for Smart, Sustainable and Inclusive Growth*, EUROPEAN COMM'N, 12 (Mar. 3, 2010). The EU Strategy 2020 aimed for all European households to have broadband access by 2013, access to much higher Internet speeds (30 Mbps or above) by 2020, and for fifty percent or more of European households subscribing to Internet connections above 100 Mbps. *Id.* The targets were updated in 2016 towards achieving a “European Gigabit Society” by 2025. *Id.*; Gerli et al., *supra* note 32, at 728.

⁶⁴ See generally *Benchmarking 15 National Broadband Plans*, ERICSSON (2014).

⁶⁵ See Winseck, *supra* note 3, at 256–57.

⁶⁶ See PHILLIPPA BIGGS ET AL., INT'L TELECOMM. UNION, PLANNING FOR PROGRESS: WHY NATIONAL BROADBAND PLANS MATTER 12 (2013), <http://www.broadbandcommission.org/documents/reportNBP2013.pdf> [<https://perma.cc/9GKQ-4LXM>]; *Digital Single Market: Investment Models*, EUROPEAN

targets a five-fold increase in broadband penetration in five years via numerous access initiatives and fiber network build-out plans.⁶⁷ Similarly, the Kenyan plan focuses on providing connections that are at least 5 mbps, extending national fiber optic cables by 30,000 kilometers, building neutral national data centers, as well as enhancing digital literacy.⁶⁸

One point of differentiation between developed and developing countries is the greater emphasis on mobile broadband in universal service policies of developing countries. For example, in the NBP for Kenya, Mexico, and India, there is a greater focus on building out wireless rather than fixed wireline infrastructure.⁶⁹ The belief is that mobile phones are “the best candidate for connectivity,” since mobile phones in these countries are usually substitutes, and not complements, for wireline access to the Internet.⁷⁰ For instance, mobile operators using wireless technologies are the primary providers of Internet services in Kenya, accounting for ninety-eight percent of total Internet subscriptions.⁷¹ The prominence of mobile infrastructure in policies is due to mobile networks costing about half as much as fixed-line networks.⁷² Moreover, rollout is more flexible and faster, enabling swifter reach to rural areas. Although in some areas closing the market efficiency gap with competitive policies is enough to ensure adequate supply of mobile infrastructure, developing countries still require public intervention, because their market forces are insufficient to cover even lower mobile network

COMMISSION (Oct. 3, 2017), <https://ec.europa.eu/digital-single-market/investment-models> (examples of investment models for broadband plans) [<https://perma.cc/ATW8-SB6V>].

⁶⁷ See *Willingness to Pay for Broadband in Nigeria*, *supra* note 63, at 422. The Nigerian Telecommunications Commissions initiatives include ensuring that no place is more than thirty miles away from backbone infrastructure; encouraging the private sector to build and manage wireless broadband services in many state capitals, cities and towns; financing the community centers via a universal service fund; and various rural initiatives. *Id.*

⁶⁸ GOV'T KENYA, THE NATIONAL BROADBAND STRATEGY, 6–7 (2013), http://icta.go.ke/pdf/The_National_Broadband_Strategy.pdf [<https://perma.cc/772L-J9R2>].

⁶⁹ See FERNANDO BERMEJO ET AL., OPEN SOC'Y FUNDS., MAPPING DIGITAL MEDIA: GLOBAL FINDINGS 299 (2014) (describing the history of India's Universal Service Obligation Fund, which now intends to help with expansion of mobile telecom infrastructure in rural, commercially non-viable areas); Jayakar & Liub, *supra* note 34, at 193 (noting how India's universal service fund supports infrastructure for mobile services, such as the set-up and maintenance of cellular towers in remote areas without mobile service); *Universal Service Funds in Africa*, *supra* note 31, at 622 (noting that many universal service funds in Africa provide some financial incentives for mobile network operators to deploy networks in economically unattractive areas).

⁷⁰ See Alleman et al., *supra* note 60, at 87–88. Many services like e-mail, web surfing, movies, photos are now available on smart mobile phones. Mobile phone subscribers exceed fixed connection subscriptions in 200 countries; in 166 countries the mobile penetration is over twice that of fixed connections. *See id.*

⁷¹ See THE NATIONAL BROADBAND STRATEGY, *supra* note 68, at 15.

⁷² See Casanueva-Reguart, *supra* note 38, at 2093.

infrastructure costs.

Many governments and regulators are challenged by how to effectively design and implement universal service policies and programs that will truly increase access.⁷³ Variations in the definition of universal service result in differing priorities, use of policy tools, and degrees of public sector intervention.⁷⁴ Several well-regarded individuals in the field explain that universal service policies and programs call for demographic, territorial, and/or layered modes of expansion, which directly benefit people, high cost areas, or discrete programs, respectively.⁷⁵ Some countries choose to pursue all three modes “simultaneously, or in varying sequential combinations.”⁷⁶

In the demographic expansion mode, universal service programs aim to provide access to end-users of communication services in order to gradually expand access to more *people*. The most common method is to provide subsidies or rate reductions to end-users. The Lifeline program in the United States does exactly that: it targets assistance to low-income households, making up the difference between the cost of the service and the amount consumers can pay.⁷⁷ Some credit these programs with the high telephone penetration rate in the United States these past decades.⁷⁸ Today, the program has expanded to provide subsidies for mobile and Internet service. In comparison, until recently universal service policies in India and China mostly lacked the demographic expansion mode. Some note this is because “universal household access would be too expensive, considering the population, geographical area and stage of economic development of these two countries.”⁷⁹ Others note that greater *mobile* penetration makes policymakers less compelled to add household access as a universal service goal.⁸⁰ However, to advance broadband growth in rural areas, the regulatory agency in India is considering implementing subsidized tariffs for rural subscribers.⁸¹ Overall, by implementing strategies that are directly targeted at end-users, policies

⁷³ See CONNECTING THE GLOBE, *supra* note 38, at I-7, V-2, V-10, VI-4, VII-2.

⁷⁴ Thai & Falch, *supra* note 40, at 323.

⁷⁵ See Harmeet Sawhney & Krishna Jayakar, *Universal Service: Migration of Metaphors*, in MAKING UNIVERSAL SERVICE POLICY: ENHANCING THE PROCESS THROUGH MULTIDISCIPLINARY EVALUATION 32-37 (Barbara A. Cherry et al. eds., 1999).

⁷⁶ Jayakar & Liub, *supra* note 34, at 188.

⁷⁷ Whereas Lifeline provides a discounted monthly fee for one telephone line in a residence, the Link-Up program discounts the initial installation fee. See Gideon & Gabell, *supra* note 37, at 740 (“Optimal policy provides assistance only to those who otherwise would not subscribe.”); CONNECTING THE GLOBE, *supra* note 38, at i (“These universal subsidy schemes are most effective when they are targeted, explicit and competitively neutral.”).

⁷⁸ See Gideon & Gabell, *supra* note 37, at 740.

⁷⁹ Jayakar & Liub, *supra* note 34, at 194.

⁸⁰ See *id.*

⁸¹ *Id.* at 192.

tackle factors inherent in rural and low-income areas, such as low purchasing power, low usage, and seasonal income; and through subsidies, increase potential demand for services.⁸²

In the territorial expansion mode, universal service programs seek to extend network infrastructure across geographical space. The distinction between territorial and demographic expansion can be blurry, since both aim to connect people to services as a final goal. Territorial expansion, however, is more about increasing access in a certain location rather than for a specific end-user.⁸³ In Colombia, Peru, and India, for example, universal access initiatives initially aimed to place services within a certain distance of a person or habitation.⁸⁴ The classic example of a territorial expansion policy is the U.S. Connect America Fund, which is one of four initiatives under the Universal Service Fund and provides subsidies to eligible telecommunications carriers who develop infrastructure in high-cost areas.⁸⁵

The layered expansion mode refers to the gradual increase in the number of services or programs that receive public support. In many countries, such as the United States and Mexico, this has meant expanding universal service to directory assistance and 911 emergency services, schools and libraries, and rural hospitals.⁸⁶ In developing countries, layered expansion somewhat overlaps with territorial expansion as it refers to supporting digital community centers, or “telecenters,” usually located in rural areas. In Kenya, the government is funding “digital villages” that will enable rural inhabitants to access broadband Internet,⁸⁷ and in Mexico a similar program has led to a 500 percent increase in Internet access points in public places.⁸⁸

⁸² See Arturo Muent-Kunigami & Juan Navas-Sabater, *Options to Increase Access to Telecommunications Services in Rural and Low-Income Areas* 11 (World Bank, Working Paper No. 178, 2010).

⁸³ See *id.* at 28.

⁸⁴ See Emiliani, *supra* note 25, at 19 (describing universal access initiatives in Colombia and Peru). In India, the government initially made targets for telephone access, where a public phone had to be available within a certain distance from any habitation.) The government also focused on providing access to 290,000 “uncovered villages.” Jayakar & Liub, *supra* note 34, at 191.

⁸⁵ See Sanford V. Berg et al., *Universal Service Subsidies and Cost Overstatement: Evidence from the U.S. Telecommunications Sector*, 35 TELECOMM. POL’Y 583, 590 (2011).

⁸⁶ See Telecommunications Access Policy Division, FCC, *Universal Service, Connectado Program, INT’L TELECOMM. UNION*, <https://www.fcc.gov/general/universal-service> [<https://perma.cc/PDP6-96M8>]; *Punto México*, <https://www.itu.int/net4/wsis/archive/stocktaking/Project/Details?projectId=1487175478> [<https://perma.cc/E33S-YQDW>].

⁸⁷ NYABUGA & BOOKER, *supra* note 46, at 6.

⁸⁸ Casanueva-Reguart, *supra* note 38, at 2111.

2. *Universal Service Financing Mechanisms*

Besides the variation of which services and beneficiaries' countries prioritize, countries also use different approaches to finance their universal service agendas. Historically, the most common mechanisms were cross-subsidies among long-distance and local rates, as well as carrier mandates for development in rural areas. First, monopolies or state-owned companies used internal cross-subsidies where they charged high prices for international calls to subsidize local calls.⁸⁹ However, this was soon not feasible in competitive and liberalized environments.⁹⁰ Another common mechanism national regulators used was to mandate carriers to rollout certain services in underserved areas and to underserved populations. Such a rollout obligation in India, for example, required operators who received a spectrum license to extend services within a specific time period in their territories as a condition of the license.⁹¹ At one point, Indian operators were required to have at least ten percent coverage in rural areas under their licensing agreement.⁹² These rollout obligations have been largely ineffective. In Mexico, Telmex's designated network operator status at times led to evasion of obligations to rural communities.⁹³

Today, the preferred mechanism to finance universal service is through a publicly-administered Universal Service Fund (USF). Studies show that USFs have been adopted by numerous countries in Africa and Latin America.⁹⁴ Nations usually finance USFs through fees charged during the licensing process or, more commonly, levies on telecommunications

⁸⁹ Madden, *supra* note 59, at 111 ("In former monopoly markets universal service obligations are mostly provided through cross-subsidies on profitable market segments."); *see generally* Munte-Kunigami & Navas-Sabater, *supra* note 82, at 30.

⁹⁰ Emiliani, *supra* note 25, at 12 (noting five strategies in a liberalized market: imposing the universal service obligation on the incumbent; imposing the obligation on all firms; procuring universal service from one firm; procuring universal service from several firms; and offering universal service subsidies).

⁹¹ VIBODH PARTHASARATHI ET AL., OPEN SOC'Y FOUNDS, MAPPING DIGITAL MEDIA: INDIA 94–95 (2012); Jayakar & Liub, *supra* note 34, at 191.

⁹² Jayakar & Liub, *supra* note 34, at 191.

⁹³ Telmex was the designated network operator in Mexico and after negotiations with regulators, it was freed from its obligation to serve communities with fewer than 500 people—at one point, that was 47.2% of rural inhabitants in Mexico. *See* Casanueva-Reguart, *supra* note 38, at 2102.

⁹⁴ U.N. Conference on Trade and Development, *Financing Mechanisms for Information and Communication Technology for Development*, 14, UNCTAD/DTL/STICT/2009/5 (2010) (noting the popularity of USFs); *Universal Service Funds in Africa*, *supra* note 31, at 618 (noting studies showing that universal service funds are the most adopted universal service strategy across Africa); Emiliani, *supra* note 25, at 14 (citing data from 2006 that indicates that twelve Latin American countries use special funds; thirteen rely on contributions from international agencies and NGOs; and three continue to use cross-subsidies).

operators. To finance USFs, telecommunications operators pay (1) a fixed fee in exchange for a spectrum and operating license, and/or (2) a variable fee calculated as a percentage of the company's revenues.⁹⁵ In India, the regulatory authority initially set the levy at five percent of adjusted gross revenue,⁹⁶ and in various countries in Africa, the levies have ranged from 0.2 to 5 percent.⁹⁷ In some countries such as Burkina Faso, Ghana, Tanzania, and Uganda, the government and international donors like the World Bank also contribute to USFs.⁹⁸ Telecom companies in the U.S. usually recover their USF levies by passing the cost directly to their customers, which some argue is essentially a tax.⁹⁹ Once funded, the USF is then administered by a regulatory or independent agency, whose mandate typically entails the allocation of funds "to provide financial support to rural access and other ICT development projects, often through a competitive bidding process, in which the same operators that contribute to the fund are invited to bid for the subsidies and the mandate to deliver rural access and services."¹⁰⁰

In addition to financing commercially nonviable projects with levies, some countries fund universal service projects via direct grants. Some argue universal service obligations should in fact be financed in accordance with democratic principles, meaning that the government should fund the program through general tax revenues.¹⁰¹ In the United States, the Obama Administration allocated \$7.2 billion to broadband infrastructure deployment in accordance with the National Broadband Plan.¹⁰² The turn to direct grants

⁹⁵ Mark A. Jamison et al., *Competition in Wireless: Spectrum, Service and Technology Wars*, 27 TELECOMM. POL'Y 319, 321 (2003). Some mobile operators defend high service charges because of high licensing fees paid to the government as a result of a licensing process that leads to high bids. See Kerretts, *supra* note 36, at 57.

⁹⁶ Jayakar & Liub, *supra* note 34, at 192.

⁹⁷ *Universal Service Funds in Africa*, *supra* note 31, at 622 (reporting that thirty-four countries' universal service funds were surveyed in Africa and most of those funds were financed by levies contributed by mobile network operators. Only Morocco and Togo financed with a "pay or play" strategy. The levies ranged from 5% (Tunisia), to 0.5% (Mauritius and Kenya), as well as 0.2% (South Africa)).

⁹⁸ *Id.*

⁹⁹ Although the FCC does not require this, most companies do it anyway. See Berg et al., *supra* note 85, at 583; Rob Frieden, *Killing with Kindness: Fatal Flaws in the \$5.7 Billion Universal Service Funding Mission and What Should Be Done to Narrow the Digital Divide*, 24 CARDOZO ARTS & ENT. L. J. 448, 456 (2006).

¹⁰⁰ U.N. Conference on Trade and Development, *supra* note 94, at 14; see Jayakar & Liub, *supra* note 34, at 191 (explaining how the universal access levy in India was collected from licensed telecom operators as a percentage of their revenues; the funds were then used to reimburse service providers who provide universal service in remote areas); see also Berg et al., *supra* note 85, at 585 (noting that the Universal Service Administrative Company (USAC) was created in 1997 to administer the universal fund for the FCC).

¹⁰¹ Alleman et al., *supra* note 60, at 88.

¹⁰² *Id.* at 90; Gideon & Gabell, *supra* note 37, at 737.

in the United States, however, was in addition to USF levies and was intended to serve as a financial stimulus package after the economic downturn.¹⁰³ While in the United States this was a one-off stimulus package, in Chile universal access projects are funded directly by the national treasury rather than levies on telecom companies.¹⁰⁴

After years of relying on market mechanisms and universal service levies to fund infrastructure development, governments are adopting a new wisdom regarding the advantages of increased public funding for faster development of broadband networks.¹⁰⁵ In fact, many national broadband plans tackle backbone infrastructure challenges by increasing public funding for such projects and partnering with private actors via public-private partnerships (PPPs) to collaborate in the design, construction, operation, and financing of the infrastructure.¹⁰⁶ Developed and developing countries are both pursuing the PPP model.¹⁰⁷ Some countries like Australia and Nigeria are

¹⁰³ Alleman et al., *supra* note 60, at 88.

¹⁰⁴ Emiliani, *supra* note 25, at 23; Muenta-Kunigami & Navas-Sabater, *supra* note 82, at 31.

¹⁰⁵ See generally Bronwyn Howell & Bert Sadowski, *Anatomy of a Public-Private Partnership: Hold-up and Regulatory Commitment in Ultrafast Broadband*, 42 TELECOMM. POL'Y 552 (2018).

¹⁰⁶ See *id.* at 553; see generally WORLD BANK, PUBLIC-PRIVATE PARTNERSHIP IN TELECOMMUNICATIONS INFRASTRUCTURE PROJECTS: CASE OF THE REPUBLIC OF CONGO 5, <https://openknowledge.worldbank.org/bitstream/handle/10986/12540/687020ESW0P1220cover0PO1223950Congo.pdf> [<https://perma.cc/2ZNB-PDL6>].

¹⁰⁷ For example, in New Zealand, the government has partnered with four firms to fund a fiber-to-the-home (FTTH) network aimed to reach eighty-five percent of the population by 2024. See Howell & Sadowski, *supra* note 105, at 553. In the United States, rural towns are joining together using consortium and PPP models to extend fiber cables to the home, pushing “for much higher speed connectivity than is likely in the near term to be provided by existing service providers.” *Case Studies on PPP Arrangements for Telecommunications*, PUBLIC-PRIVATE-PARTNERSHIP LEGAL RESOURCE CTR. (Sept. 12, 2019), <https://ppp.worldbank.org/public-private-partnership/sector/telecom/telecom-laws/case-studies-telecommunications> [<https://perma.cc/57E2-RNTC>].

establishing infrastructure corporations,¹⁰⁸ turning to a wholesale open access model where they share the national backbone infrastructure with retail service providers.¹⁰⁹ Because most projects are still in the early stages, reports on investment partnerships are limited.¹¹⁰ Further research is needed regarding the extent of risk sharing among partners and the role of international investors and institutions in infrastructure development including, for example, the implications of Chinese engagement in many African countries' backbone telecommunications development.¹¹¹

3. *Failure of Universal Service Funds*

Overall, in liberalized telecommunications markets, universal service policies have played a critical role in mitigating persistent access gaps in commercially non-viable locations. Scholars note that the universality of voice telephone service in the United States two decades ago was a result of a

¹⁰⁸ Australia's National Broadband Network project resulted in a fully funded government corporation, which designs, builds and operates Australia's wholesale broadband access network. The project started with the aim of FTTH network deployment to most residences, "supplemented by satellite and wireless connections to the remainder." Howell & Sadowski, *supra* note 105, at 553; see also Matthew L. James, *National Broadband Network*, PARLIAMENT AUSTL., https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/BudgetReview201314/NBN [<https://perma.cc/GX66-KHNR>]. In Nigeria, several InfraCos were licensed for geographic zones of the country and are utilizing the Open Access Model to offer fiber penetration and backbone infrastructure available on a non-discriminatory basis to telecommunications operators. See Paul Adepoju, *60 Bidders Compete for Fibre Deployment Licenses*, ITWEB AFRICA (June 20, 2017), <http://www.itwebafrica.com/networks/270-nigeria/238034-60-bidders-compete-for-fibre-deployment-license> [<https://perma.cc/838U-L78W>]. This ensures reliable and fast broadband services for the country. *Id.*

¹⁰⁹ See OECD, *FIXED BROADBAND NETWORKS*, *supra* note 21, at 26 ("In nearly all countries in the OECD area, incumbent telecommunication providers that operate copper-based broadband networks are required to offer unbundled wholesale access to competitors at regulated rates. Such open access mandates ensure that even if competition among physical network infrastructure is infeasible, there is the opportunity for competition at the retail level to drive consumer benefits and innovation.")

¹¹⁰ Howell & Sadowski, *supra* note 105, at 553 (noting that information is scant since most projects are in the early stages, and it is hard to compare projects since scopes vary by country).

¹¹¹ See generally VIVIEN FOSTER ET AL., WORLD BANK, *BUILDING BRIDGES: CHINA'S GROWING ROLE AS INFRASTRUCTURE FINANCIER FOR SUB-SAHARAN AFRICA* xiii (2008). Kenya's National Optic Fibre Backbone Infrastructure (NOFBI) project takes on a multi-stakeholder approach towards governance and is being financed by many actors, including China's Export-Import Bank. See Ewan Sutherland, *China and Africa: Alternative Telecommunication Policies and Practices*, 17 AFR. J. INFO. & COMM 120, 183–84 (2016) (table showing Chinese support of backhaul networks).

liberalized market and universal service funding strategies.¹¹² However, the recent turn to national broadband plans and increased public funding supports scholars' conclusions that universal fund mechanisms are not well positioned to bridge the digital divide and close the broadband access gap. This is consistent with some scholars' findings that USFs are failing around the world, citing the cause as partially ill-conceived policies.¹¹³

One reason for USF's failure to close the access gap is government corruption and lack of accountability. Many funds in developing countries lack detailed public information—specifics which are necessary to keep the government and operators accountable—calling the transparency and accountability of funds into question.¹¹⁴ Concerns about corruption are especially persistent in countries where service expansion by telecom operators is mandated. The lack of performance monitoring and enforcement of mandated universal service commitments in African countries and other developing countries have often been noted.¹¹⁵ More commonly, sometimes operators under-prioritize truly unserved markets, upgrading broadband infrastructure for existing slower networks instead.¹¹⁶ Predictably, the potential success of universal service policies is limited if operators are not held accountable to expand coverage in rural areas.

Another reason for the failure of USFs to bridge the access gap is unsatisfactory operations and implementation of USFs by governments, marked by operator cost overstatements and inefficient outcomes. Typically, after levied funds have been collected (1) the cost of universal service is

¹¹² Penetration of telephone service in the U.S. increased from 91.4% in 1983 to 95.5% in March 2003. Gideon & Gabell, *supra* note 37, at 738.

¹¹³ *Universal Service Funds in Africa*, *supra* note 31, at 625.

¹¹⁴ *Id.* at 624 (noting that a majority of universal service funds in Africa lacked more specific public information about the usage of funds and that financial records were “somewhat patchy and outdated.”); Jayakar & Liub, *supra* note 34, at 192 (noting that the universal service fund disbursements in India were problematic because the collected levies were credited to the national treasury as general revenue, and thus available for other budgetary needs).

¹¹⁵ *Universal Service Funds in Africa*, *supra* note 31, at 625–26 (observing lack of performance monitoring of universal service funds in African countries); Casanueva-Reguart, *supra* note 38, at 2113 (noting that “Telmex’s dominance made it difficult for government authorities to enforce the universal service commitments” in the company’s operator license and contract under the Social Coverage Fund program); PARTHASARATHI ET AL., *supra* note 91, at 95 (remarking that the government has full discretionary powers to withdraw, modify, or distort certain commitments and initiatives); NYABUGA & BOOKER, *supra* note 46, at 84 (mentioning cases of corruption in spectrum allocation, the lack of enforcement of the use-it-or-lose-it principle with spectrum, and government bias towards Telkom Kenya given the government’s ownership stake).

¹¹⁶ See DOUG BRAKE, INFO. TECH. & INNOVATION FOUND., A POLICYMAKER’S GUIDE TO RURAL BROADBAND INFRASTRUCTURE 12 (2017).

assessed, and (2) the provider is compensated for the cost.¹¹⁷ Considering the information asymmetries regarding infrastructure build-out, the procedure can lead to waste as it entices universal service providers to overstate costs to receive greater compensation.¹¹⁸ In response, some governments have implemented least-cost auctions, also called reverse auctions, whereby USF administrators award projects on a least quoted subsidy basis, lowering the costs of universal service programs.¹¹⁹ Though the auctions have improved the disbursement process—resulting in more structure, transparency, and competition among bidders—in many countries USFs remain underutilized, sometimes because of poor implementation.¹²⁰ Although not many comprehensive studies on the effectiveness of USFs exist, one study showed that at one point \$400 million of funds lay idle in over twenty African countries.¹²¹

Overall, studies show that USFs have failed to address the limited coverage of telecom services in many countries due to “poor policy formulation, inadequate stakeholder engagement, lack of accountability, inaccurate data, undue political influence and the narrow scope of universal services.”¹²² As a result, the digital divide persists, especially in the last mile of the globe. Moreover, there has been a shift in solutions, with new rhetoric focusing on innovation and global partnerships to close persisting market access gaps. The UN has notably called for this in the SDGs. Seeing the last

¹¹⁷ See generally Axel Gautier & Xavier Wauthy, *Competitively Neutral Universal Service Obligations*, 24 INFO. ECON. & POL’Y 254 (2012).

¹¹⁸ Berg et al., *supra* note 85, at 589.

¹¹⁹ Jayakar & Liub, *supra* note 34, at 192 (describing a study of the use of a multi-layered bidding process on a least quoted subsidy basis in India, which scholars have credited with lowering of universal service program costs); Alleman et al., *supra* note 60, at 90 (finding that least-cost auctions noticeably reduced subsidies in many countries); *id.* at 90 (explaining that the U.S. FCC is reconsidering auctions in the USF context because of alleged waste and fraud in funds distribution); Emiliani, *supra* note 25, at 15 (detailing the process of a “reverse” auction and a “beauty contest”).

¹²⁰ See generally Jayakar & Liub, *supra* note 34, at 192; Emiliani, *supra* note 25, at 29-30 (explaining that in Latin American countries USFs are not disbursing resources efficiently, there is a lack of a clear “collect-disburse mechanism,” and too much “bureaucracy involved in the approval of the disbursements.”); Lilian Ochieng, *United Nations Report Lauds Kenya’s Broadband Plan*, DAILY NATION (Oct. 5, 2015), <https://www.nation.co.ke/business/United-Nations-report-lauds-Kenya-s-internet-plan/996-2899190-rcf3k7/index.html> (noting USF implementation challenges in Kenya) [<https://perma.cc/Y8P7-2FZ4>].

¹²¹ FERNANDO BERMEJO ET AL., OPEN SOC’Y FOUNDS. DIGITAL JOURNALISM: MAKING NEWS, BREAKING NEWS 299 (Marius Dragomir et al. eds., 2014) (citing a watchdog foundation which noted that the Indian government failed to utilize raised resources to develop telecom infrastructure); *Universal Service Funds in Africa*, *supra* note 31, at 622 (observing that implementation of USFs across Africa is inefficient and ineffective, and most funds lack regular financial reporting).

¹²² *Universal Service Funds in Africa*, *supra* note 31, at 626.

mile problem and access gaps as a business opportunity, technology and satellite companies have shifted toward developing innovative global infrastructure solutions such as Internet-beaming balloons, drones, and LEO satellite mega-constellations.

III. MOVING BEYOND LIBERALIZATION AND UNIVERSAL SERVICE FUNDS: INNOVATIVE GLOBAL BROADBAND INFRASTRUCTURE PROJECTS

Even though liberalization policies and universal service mechanisms boosted the deployment of communications infrastructure worldwide, the broadband access gap persists in many countries. On a global scale, infrastructure development for rural communities in the developing world represents the ultimate last mile. Seeing that many countries failed to bridge the Internet access gap nationally, the UN reiterated its Internet access targets in the 2030 Agenda for Development, otherwise known as the SDGs. The SDGs discuss ICT infrastructure development and Internet access, identifying ubiquitous connectivity and the digital divide as a global challenge. Ultimately, the challenge is one of global governance.

Two SDGs discuss ICT infrastructure development and universal Internet access. Primarily focused on infrastructure, industrialization, and innovation, SDG 9 notes the paucity of communications infrastructure in developing countries, remarking that sixteen percent of the global population does not have access to mobile broadband networks.¹²³ One target for SDG 9 seeks to increase access to ICTs significantly, and to provide universal and affordable access to the Internet in least developed countries by 2020. SDG 9 recognizes that the way to close the digital divide is via regional and transborder infrastructure that is affordable and provides equitable access for all. If SDG 9 spells out the *ends* in terms of more infrastructure and innovation, then SDG 17 encapsulates the *means* of implementation to achieve the SDGs: by strengthening global partnerships. SDG 17 envisions a revitalization of multi-stakeholder and public-private partnerships comprising governments, the private sector, and civil society at the global, regional, national, and local levels, as the implementation mechanism for the development agenda. SDG 17's aim is to unlock private resources and foreign direct investment to

¹²³ The goal of infrastructure development, articulated in SDG 9, is a notable addition to the global development agenda, as it was not explicitly stated in the Millennium Development Goals (MDGs). One target for SDG 9 communicates the value of infrastructure development and investment—namely, the enabling effects of infrastructure investment in enhancing economic development and human well-being, especially in least developed countries where basic infrastructures are lacking. The indicator used to measure progress towards this SDG is improvement in the “proportion of population covered by a mobile network.” *Sustainable Development Goal 9*, UNITED NATIONS, <https://sustainabledevelopment.un.org/sdg9> (follow “Targets & Indicators” hyperlink) [<https://perma.cc/63U6-YKPH>].

achieve development objectives, especially in critical sectors like ICT infrastructure. In fact, SDG 17 directly references the digital divide by citing ICT-related facts, including that more than four billion people do not use the Internet and ninety percent of them are from the developing world. The UN's vision is that global partnerships will enable achievements for the technology-related targets of SDG 17, with progress measured by indicators.¹²⁴

Several companies have recently articulated global connectivity agendas and broadband infrastructure projects which happen to coincide with the SDGs. The companies are responding to the challenge of the last mile for infrastructure provision and are motivated to bridge the global digital divide not through national universal service mechanisms, but with innovative global infrastructure. Believing that everyone should have Internet access, and that provision could be at high altitudes rather than solely terrestrial, companies like Alphabet, Facebook, SpaceX, among others, have begun developing Internet-beaming balloons, drones, and LEO satellite mega-constellations.¹²⁵ Whereas universal service mechanisms and funds were an incremental response to the need for more communications infrastructure development, these non-traditional “telecommunications” companies are introducing potentially disruptive infrastructure.¹²⁶ Their missions have left many wondering whether the companies' innovations will complement or supplement current communications infrastructure and help developing countries leapfrog certain stages of infrastructure development.

Though these companies are ostensibly working to bridge the global digital divide, the SDGs likely did not motivate the endeavors; the decision to turn to innovation was due to an appetite to reach under-connected markets

¹²⁴ *Sustainable Development Goal 17*, UNITED NATIONS, <https://sustainabledevelopment.un.org/sdg17> (the indicators are “fixed Internet broadband subscriptions per 100 inhabitants, by speed” and “proportion of individuals using the Internet.”) [<https://perma.cc/2PXP-ACRK>].

¹²⁵ See Part III.A and III.B, *infra*.

¹²⁶ The U.S. FCC initially determined that broadband services were “information” and not “telecommunications” services. Recently, the FCC reclassified broadband Internet access services as telecommunications services, but it went through “pains to emphasize that it was reclassifying . . . with a ‘light touch’”—solely, for net neutrality purposes. Matthew L. Gibson, *Evolution of the FCC's Open Internet*, 54 INFRASTRUCTURE 9 (2014-15); see also *id.* at 9, citing FCC Open Internet Order, ¶ 456–542 (2015) (“The FCC has also decided not to impose the following common carrier obligations on broadband Internet access service: Any form of rate regulation . . . ; Any requirement that providers of broadband Internet access service contribute to the universal service fund; or any other form of FCC-imposed tax or fees.”); see also Frieden, *supra* note 99, at 461.

and customers.¹²⁷ Still, most of the companies use the SDGs as valuable rhetoric to navigate complex multi-jurisdictional regulatory issues and develop viable market strategies. Moreover, these companies are inadvertently following the UN's development agenda with regards to utilizing multi-stakeholder partnerships. However, as shown in the Section below, the partnerships and support corralled by these companies embarking on connectivity projects are not being pursued at the behest of the UN or the SDGs. The partnerships are a necessary byproduct of the current regulatory landscape concerning global resources, such as spectrum, civil airspace, and outer space, as well as the transnational nature of the Internet.

To succeed, the global broadband infrastructure these companies envision must accord with various regulatory frameworks. Ultimately these companies may initiate private-led standards, most evidently with regard to satellite mega-constellations. Although SDG 17 articulates governance concerns regarding partnerships and clarifies that the public sector's role is to set clear directions for such partnerships through regulation, to date, regulations have not caught up to the tech companies' innovations. Moreover, by navigating multi-jurisdictional issues and various regulatory frameworks to test viable market strategies, the non-traditional "telecommunications" companies could create regulatory effects. Such effects may be created as these companies informally influence national and global policies regarding global communications and contribute to the development of global and long-lasting infrastructures that have the power to shape lives for decades. Although projects of this nature—which rely on international and national regulatory frameworks—are not *per se* new, the transnational nature of high-altitude network infrastructure coupled with the global development agenda calls for more debate about the proposed last mile solutions and the adequacy of current regulatory frameworks.

The first Section below describes the global infrastructure solutions envisioned by Loon, Facebook, and next-generation satellite companies to bridge Internet access gaps globally, as well as the corresponding regulatory frameworks. The second Section then compares and contrasts the different

¹²⁷ In terms of the revenue potential, one Loon engineer stated it could be a financially viable project if people paid a small portion of their income for access. See Ben Popper, *Inside Project Loon: Google's Internet in the Sky is Almost Open for Business*, VERGE (Mar. 2, 2015), <https://www.theverge.com/2015/3/2/8129543/google-x-Internet-balloon-project-loon-interview> ("Think about it—with 4.5 billion people without Internet access, take 5 percent; you're talking 250 million people . . . If those people pay just a small portion of their monthly income, say \$5 a piece, you're going to be in a billion dollars a month in revenue, tens of billions a year in revenue. So it's good business, too.") [<https://perma.cc/TNZ5-3H3E>]; see generally Peter H. Diamandis, *4 Billion New Minds Online: The Coming Era of Connectivity*, SINGULARITYHUB (July 27, 2018), <https://singularityhub.com/2018/07/27/4-billion-new-minds-online-the-coming-era-of-connectivity/> [<https://perma.cc/U5MN-Q4T5>].

broadband infrastructure projects to evaluate their disruptive potential. The last Section assesses the projects' prospective contribution towards telecommunications development, asking whether the innovations will enable some countries to leapfrog certain stages of infrastructure development.

A. High-Altitude Internet Infrastructure Projects and Applicable Regulatory Frameworks

Several companies are seeking to provide Internet access to people in rural and remote areas worldwide from high altitudes: Loon with its Internet-beaming balloons, Facebook with its solar-powered drones, and next-generation satellite companies with their LEO mega-constellations. Floating over various countries at altitudes above 60,000 feet, Loon balloons and Facebook drones trigger the application of mainly national civil aviation laws (and some international civil aviation standards typically implemented by national legislation), given that states have "complete and exclusive sovereignty over the airspace above [their] territory."¹²⁸ By floating in LEO, around 100 to 1250 miles high above Earth and sovereign airspace, the satellite mega-constellations must comply with the international and national regulations governing the global commons that is outer space. The projects and the corresponding regulatory frameworks are described below.

1. *Airspace Connectivity: Internet-beaming Balloons and Drones*

Through a network of balloons floating at high altitudes, Loon aims to extend Internet connectivity to help fill coverage gaps and bring people back online after disasters. The mission was conceived in 2011 and was under the supervision of employees at "X," Google's famous experimental division and "moonshot factory."¹²⁹ Loon is now among many Alphabet projects to become its own standalone business. Now Loon LLC, the global Internet connectivity project, is Alphabet's fully owned subsidiary and a sister company to Google.¹³⁰ Loon balloons are designed to fly roughly twice as high as

¹²⁸ Convention on International Civil Aviation art. 1, Dec. 7, 1944, https://www.icao.int/publications/Documents/7300_orig.pdf [<https://perma.cc/TRK3-7AQU>].

¹²⁹ Astro Teller, *The Unexpected Benefit of Celebrating Failure*, TED2016 (Feb. 15, 2016), https://www.ted.com/talks/astro_teller_the_unexpected_benefit_of_celebrating_failure/transcript [<https://perma.cc/NG55-NQVF>].

¹³⁰ Alphabet refers to these projects as 'Other Bets.' The businesses include Waymo (self-driving cars), Nest ("smart" thermostats), Verily (Alphabet's life sciences brand), Google Fiber, Project Loon (Internet connectivity with balloons) and Project Wing (drones), among others. Nick Statt, *Alphabet's Experimental Investments in the Future Continue to Cost it a Fortune*, VERGE (Jul. 23, 2018),

commercial air traffic and weather in the stratosphere. Loon engineers use wind forecast data sets from the U.S. National Oceanic and Atmospheric Administration and decision-making algorithms to determine which stratosphere current a balloon should ride when seeking to fly over certain areas.¹³¹ Custom-built autolaunchers launch a new balloon into the network every thirty minutes. The company notes that with its software powered by machine learning, it can send small squads of balloons to where people need service.¹³² Media coverage of Loon's test flights reveals that the balloons fly primarily over the southern hemisphere, where "services are needed most."¹³³

Loon's balloons serve as "floating cell towers" working to extend wireless broadband infrastructure and coverage to a prescribed area. Each balloon contains a box of solar-powered electronics (transmitter, receiver, etc.) that allows it to make a radio link to a telecommunications network on the ground and beam down high-speed cellular Internet coverage to smartphones and other LTE-enabled devices.¹³⁴ From already established ground stations, antennae transmit connectivity across a balloon mesh network and back down to a user's LTE phone. For wireless infrastructure, the fundamental constraint of connectivity has been the proximity of users to ground stations and cell towers. Loon's balloons address this constraint via laser optic technology that enables transmitting connectivity across longer distances between balloons. One ground-based connection point can be leveraged across several nodes, allowing a single terrestrial access point to activate a web of connectivity via multiple balloons. Recent news articles note that Loon has successfully beamed a test connection 1,000 kilometers across seven balloons.¹³⁵

Facebook has also joined the race to extend global connectivity with its own high-altitude Internet infrastructure project, housed under its Connectivity Lab.¹³⁶ With the help of an acquired company, Facebook

google-other-bets-waymo-nest-future-investments-costs-fortune [https://perma.cc/G3FY-SCPH].

¹³¹ See Popper, *supra* note 127.

¹³² Loon LLC, *Loon-Improving Navigation*, YOUTUBE (Feb. 16, 2017), https://www.youtube.com/watch?time_continue=53&v=eHCKL-fCmk8 [https://perma.cc/7D6Y-PTHG].

¹³³ The southern hemisphere is less densely populated and full of "remote areas where broadband Internet is less likely to reach." Popper, *supra* note 127.

¹³⁴ See Tom Simonite, *Project Loon*, MIT TECH. REV. (Feb. 18, 2015), <https://www.technologyreview.com/s/534986/project-loon/> [https://perma.cc/44JD-PHTG]. The balloons are made from polyethylene plastic and when fully inflated, are about fifteen meters wide. *Id.*

¹³⁵ See Kenn Abuya, *Telkom-Loon Partnership Makes Notable Milestone as Loon Sends a Single 1000km Connection Across 7 Balloons*, TECHWEEZ (Sept. 12, 2018), <https://techweez.com/2018/09/12/loon-7-balloons-1000-km/> [https://perma.cc/T8UH-2DB9].

¹³⁶ Mark Zuckerberg, *The Technology Behind Aquila*, FACEBOOK (July 21, 2016),

designed and developed a solar-powered drone with the wingspan of a Boeing 737, called Aquila, to provide Internet access in remote and rural areas worldwide.¹³⁷ Compared to Loon's balloons, drones theoretically could beam broadband coverage to a greater geographic area, provide greater Internet speed, and remain in flight for longer periods of time—some say as long as five years.¹³⁸ In addition to Facebook's Internet.org initiative and Free Basics platform,¹³⁹ Aquila was one of many ways that Facebook is working to bring people in the developing world and remote areas online.

The future for both projects includes transforming the experimental ideas into viable commercial operations. No outstanding templates exist for bringing radical technologies to market. To be truly transnational and global infrastructure, the projects must follow international and national regulations concerning airspace, outer space, and spectrum. One such regulatory framework pertaining to Loon's agenda and emerging partnerships is discussed below, given that Facebook recently stated that it is no longer building its own Internet drones.¹⁴⁰ Notwithstanding, Facebook intends to keep working with partners on high-altitude Internet systems and to influence international and national policies concerning spectrum and aviation.¹⁴¹

<https://www.facebook.com/notes/mark-zuckerberg/the-technology-behind-aquila/10153916136506634/> [https://perma.cc/YWM2-R7AE].

¹³⁷ Sean Gallagher, *Facebook's Fleet of Solar-Powered Internet Drones Grounded Forever*, ARSTECHNICA (June 27, 2018), <https://arstechnica.com/information-technology/2018/06/facebook-drops-solar-powered-internet-drone-business-cans-aquila/> [https://perma.cc/H3AC-TZYY].

¹³⁸ Soujanya Katikala, *Google Project Loon*, INSIGHT: 10 RIVIER ACAD. J. 5 (2014), https://www2.rivier.edu/journal/ROAJ-Fall-2014/J855-Katikala_Project-Loon.pdf [https://perma.cc/7CPH-FAMH].

¹³⁹ See *Free Basics Platform*, FACEBOOK FOR DEVELOPERS, <https://developers.facebook.com/docs/Internet-org/> [https://perma.cc/54VK-VEAW].

¹⁴⁰ The Aquila project conducted two public test flights of a prototype in 2016, and one of the flights resulted in serious damage to the aircraft during landing. Martin Luis Gomez & Andrew Cox, *Flying Aquila: Early Lessons from the First Full-Scale Test Flight and the Path Ahead*, FACEBOOK ENGINEERING (July 21, 2016), <https://code.fb.com/connectivity/flying-aquila-early-lessons-from-the-first-full-scale-test-flight-and-the-path-ahead/> [https://perma.cc/2H9H-PXNS]; Nick Statt, *Facebook Abandons Quest to Build Its Own Internet Drones*, VERGE (June 26, 2018), <https://www.theverge.com/2018/6/26/17507826/facebook-aquila-Internet-drone-project-shut-down> [https://perma.cc/F2RZ-6P6Z].

¹⁴¹ In the past, Facebook has lobbied to open up the global regulatory environment with regards to aviation and spectrum policy to high altitude solutions. One engineer has noted that Facebook still plans to "actively participat[e] in a number of aviation advisory boards and rule-making committees in the U.S. and internationally." Statt, *supra* note 140.

a. Loon's Connectivity Project and the Regulatory Framework

Loon's envisioned broadband infrastructure primarily concerns two regulatory spheres, overflight and spectrum. To obtain overflight permissions, Loon has approached international regulatory agencies such as the International Civil Aviation Organization, Civil Air Navigation Service Organization, among others. Loon has also approached international and national public agencies to obtain spectrum licenses. But given the high cost of spectrum licenses, Loon's efforts have focused on partnering with existing licensed telecom companies. Both regulatory spheres, in the context of Loon's already emerging partnerships, are discussed below.

Floating at an altitude of 60,000 feet, Loon's balloons require permission from several countries in the balloons' flight path to fly over their territories and reach the desired coverage area. A basic tenet of international law is that every state has complete and exclusive sovereignty over the airspace above its territory.¹⁴² The Paris Convention of 1919 sided with the "sovereignty of the air" school (as opposed to the "freedom of the air" school), emphasizing that each nation has absolute sovereignty over the airspace covering its territories and waters.¹⁴³ The Convention on International Civil Aviation of 1944, also known as the Chicago Convention, with 191 signatories, re-codified this principle of sovereignty over airspace.¹⁴⁴ Article 6 of the Chicago Convention further states that "[n]o scheduled international air service may be operated over or into the territory of a contracting State, except with the special permission or other authorization of that State."¹⁴⁵ To that end, the International Air Services Transit Agreement of 1944 (with 130 signatories) grants certain freedoms of the air, such as flying across signatory states' territory, for *scheduled* international air services.¹⁴⁶ Article 5 of the Convention, concerning *non-scheduled* air services, still requires compliance

¹⁴² See William W. Bishop, Jr., INTERNATIONAL LAW: CASE AND MATERIALS 422–23 (Little Brown & Co. Law & Business eds., 1953); John Cobb Cooper, *The Chicago Convention After—Twenty Years*, 19 U. MIAMI L. REV. 333, 334–35 (1965).

¹⁴³ Jae Woon Lee, *Revisiting Freedom of Overflight in International Air Law: Minimum Multilateralism in International Air Transport*, 38 AIR & SPACE L. 351 (2013) (the Paris Convention (1919) embraced sovereignty of the air principle, while the Chicago Convention (1944) embraced an extreme version of the principle).

¹⁴⁴ *Id.* at 360–61. The limits of the sovereignty of the air principle and the definition of "airspace" will be discussed in the context of satellite mega-constellations. See *infra* Section III.1.A.2.a.

¹⁴⁵ See Convention on International Civil Aviation, *supra* note 128, at art. 6.

¹⁴⁶ ICAO, *Information Paper on Revisiting the International Air Services Transit Agreement of 1944* (ICAO Information Paper No. LC/36-IP/2), (Nov. 25, 2015), <https://www.icao.int/Meetings/LC36/Working%20Papers/LC%2036%20-%20IP%202.en.pdf> [<https://perma.cc/333V-V7UP>].

with state regulations, otherwise the state may require landing as a condition of flying over its territory.¹⁴⁷ As a non-scheduled international flight under this framework, Loon must work with individual state aviation agencies to obtain and maintain permissions to fly over and land in certain countries.

To facilitate overflight permissions, Loon has already begun meeting with international public aviation agencies and industry groups. Loon has approached the International Civil Aviation Organization (ICAO), a UN specialized agency and global aviation forum established by the Chicago Convention. Many of ICAO's standards are the basis for member states' aviation regulations, allowing for standardization of operating procedures among ICAO countries.¹⁴⁸ ICAO provides a forum for states to discuss global aviation topics, including overland flight agreements. Ultimately, this forum facilitates and accelerates the process to obtain necessary permits from national civil aviation authorities.¹⁴⁹ In 2016, Loon went before ICAO for its triennial assembly to ask the 191 member states for airspace access. Loon specifically asked for standardized overflight agreements to allow it to expand its global and regional testing.¹⁵⁰ The ICAO General Assembly endorsed Loon, noting the balloons comply with ICAO standards and go beyond the safety requirements.¹⁵¹ When presenting to ICAO, Loon noted its intent to support SDGs 9 and 17 with its proposed infrastructure.¹⁵² Various ICAO

¹⁴⁷ See Lee, *supra* note 142, at 360; Cooper, *supra* note 142, at 339–40 (explaining ambiguities in Articles 5 and 6 of the Chicago Convention).

¹⁴⁸ States that signed the Chicago Convention are responsible for implementation of Standards and Recommended Practices (SARPs) through civil aviation authorities. See Convention on International Civil Aviation, *supra* note 128, at art. 37–38.

¹⁴⁹ See OPSGroup, *World Permit Map*, FLIGHT SERV. BEUREAU, <http://www.fsbureau.org/permitmap> [<https://perma.cc/W4BK-TF84>].

¹⁵⁰ *Google Asks for Airspace Access for Internet Balloons*, PHYS.ORG (Sept. 30, 2016), <https://phys.org/news/2016-09-google-airspace-access-Internet-balloons.html#jCp> [<https://perma.cc/5CPQ-PHLR>]; *Seventh Meeting of the North American, Central American and Caribbean Directors of Civil Aviation 3* (ICAO Working Paper No. NACC/DCA/07–WP/16, July 31, 2017) [hereinafter *Seventh Meeting*], <https://www.icao.int/NACC/Documents/Meetings/2017/NACCDCA7/NACCDCA7WP16.pdf> [<https://perma.cc/LH4B-T4X3>].

¹⁵¹ *Seventh Meeting*, *supra* note 150, at 2.

¹⁵² Loon maintained that the Internet-carrying balloons directly support the UN's SDGs, particularly goals 9 and 17, targets 9.2, 17.6, and 17.8, and indicators 17.6.2 (“Fixed Internet broadband subscriptions, by speed”); 17.8.1 (“Proportion of individuals using the Internet”); and 9.c.3 (“Percentage of population covered by a mobile network, by technology”). See ICAO, REPORT OF THE TWENTY-FIRST MEETING OF THE AFRICA-INDIAN OCEAN PLANNING AND IMPLEMENTATION REGIONAL GROUP (APIRG/21) 33 (Feb. 9, 2018), <https://www.icao.int/ESAF/Documents/APIRG/APIRG%2021/Final%20Report%20and%20Appendices/Final%20Report/APIRG%2021%20Report%20-%20FINAL.pdf> [<https://perma.cc/D2NP-F3XT>].

committees have acknowledged Loon's intent to support the SDGs.¹⁵³ In a letter to member states, ICAO's Secretary-General recommended that states finalize operational letters of agreement with Loon "to provide trans-global Internet access."¹⁵⁴ Some countries signed overflight agreements during the ICAO Assembly.¹⁵⁵ Desiring more multilateral agreements to secure necessary authorizations and partnerships for ease of operations and direct "float" paths, Loon held a briefing workshop for the Eastern African Community (EAC) and its member states.¹⁵⁶ To date, Loon has not announced the total amount of procured overflight agreements formed.

Besides working with governmental civil aviation agencies, Loon has also joined the Civil Air Navigation Service Organization (CANSO) as a way to work in partnership with Air Navigation Service Providers (ANSPs).¹⁵⁷ CANSO is a representative body of companies that provide air traffic control services and represents the interests of ANSPs. Air traffic control issues are integral to Loon's operations. Loon must coordinate directly with local air traffic control when balloons are launched, throughout their flight, and upon descent. Each balloon is equipped with a transponder that constantly transmits its position and altitude to air traffic control.

In addition to partnering with various aviation organizations and agencies, Loon balloons must comply with telecommunications regulations. To extend connectivity, the Loon balloon network relies on radio waves, which are fundamental to many telecommunications services and technologies. Rules and guidelines ensure the orderly allocation and use of radio frequencies, or "electromagnetic spectrum," at the international and

¹⁵³ "The Committee noted that [Loon] directly supports SDGs 9 and 17 and encouraged the Assembly to endorse the spirit of paper extending its coverage to all aviation solutions that are compliant with SARPs that assist in the achievement of the SDGs related to bringing the Internet to underserved parts of the world." CANSO, *Project Loon—Floating Cell Phone Towers in the Sky 2* (ICAO DG Conference Africa) (document undated), [https://www.icao.int/WACAF/Documents/DGCA/DGCA-6/WP%2010.2%20-%20Africa%20DG%20Conf%20-%20Project%20Loon%20\(CANSO\).pdf](https://www.icao.int/WACAF/Documents/DGCA/DGCA-6/WP%2010.2%20-%20Africa%20DG%20Conf%20-%20Project%20Loon%20(CANSO).pdf) [https://perma.cc/V86G-GY76].

¹⁵⁴ *Seventh Meeting*, *supra* note 150, at 4 (referencing ICAO State Letter (Ref. AN13/22.1-16/42, June 17, 2016, *High Altitude Operations of Unmanned Free Balloon*).

¹⁵⁵ Kenya and Nigeria signed agreements for overflight during the ICAO Assembly. See *Project Loon – Floating Cell Phone Towers in The Sky*, *supra* note 153, at 2.

¹⁵⁶ ICAO, *Seventh Meeting of the Directors General of Civil Aviation Administration of the AFI Region*, (AFI-DGCA/7–WP/18, July 20, 2018); Civil Aviation Safety and Security Oversight Agency (CASSOA), *Project Loon – Aviation Powered Internet CASSOA.ORG* (June 26, 2018), <http://www.cassoa.org/cassoa/?p=1713> [https://perma.cc/6KFY-RYYK].

¹⁵⁷ *CANSO Members: Project Loon*, CIVIL AIR NAVIGATION SERVS. ORG., <https://www.canso.org/sites/default/files/Article%20-%20Project%20Loon.pdf> [https://perma.cc/Z72Q-R26Q]; *Project Loon: Managing Balloon Technology in Airspace*, 4 AIRSPACE 26–27 (2016).

national levels of governance. At the international level, the ITU has long regulated the global community's use of wired and wireless communications mediums.¹⁵⁸ The ITU specifically regulates radio frequencies, working to prevent interference across borders and ensure efficient use of the finite natural resource.¹⁵⁹ With a long history of international coordination, the UN specialized agency brings together 193 member states, more than 700 private entities, and 150 academic institutions to weigh in on a complex frequency allocation plan.¹⁶⁰ The ITU regularly holds World Radio Conferences to establish regulations and allocation plans for the global use of radio spectrum. At the conferences, the latest versions of the Radio Regulations, the main body of ITU laws which have treaty status, are adopted.¹⁶¹ Each country has its own frequency allocation plan based on the ITU plan.¹⁶² With the right to exclude certain frequency allocations within national borders and with three different allocation plans divided by regions, countries' frequency allocation charts differ. Even so, there is much overlap for most allocations.¹⁶³ Beyond complying with regional and international allocations, spectrum licenses ultimately depend on national regulations and local availability.¹⁶⁴

The nature of this regulatory framework has led Loon to begin partnering with various telecommunications companies. Loon employees have done outreach with the ITU, but no significant lobbying efforts exist at the

¹⁵⁸ See generally INT'L TELECOMM. UNION, THE INTERNATIONAL TELECOMMUNICATION UNION: AN OVERVIEW (2002), <https://www.itu.int/itudoc/gs/promo/gs/81150.pdf> [<https://perma.cc/9GDT-GMX3>]; see also SCOTT MADRY ET AL., INNOVATIVE DESIGN, MANUFACTURING AND TESTING OF SMALL SATELLITES 92 (2018) ("The ITU addresses and agrees on global transmission standards for all types of media and transmission services whether via wire, coaxial cable, optical transmission systems, radio frequency and infrared transmission, or wireless mobile telecommunications systems of all types including cellular telephone, radio communications services (including specialized commercial, medical, and emergency services), satellite services of all types, and even links to UAVs and High-Altitude Platforms.").

¹⁵⁹ Veronique Wavre, *Universal Service Obligation (USO) and Spectrum Management*, in POLICY DIFFUSION AND TELECOMMUNICATIONS REGULATION 7374 (2018).

¹⁶⁰ OECD, INTERNATIONAL TELECOMMUNICATION UNION (ITU) PROFILE, 1 (2016), <https://www.oecd.org/gov/regulatory-policy/ITU%20profile.pdf> [<https://perma.cc/7XJV-WWPT>]; MADRY ET AL., *supra* note 158, at 92.

¹⁶¹ See NAT'L AERONAUTICS AND SPACE ADMIN., SPECTRUM 101: AN INTRODUCTION TO NATIONAL AERONAUTICS AND SPACE ADMINISTRATION SPECTRUM MANAGEMENT, at iv, 18 (2016), https://www.nasa.gov/sites/default/files/atoms/files/spectrum_101.pdf [<https://perma.cc/3Y2Q-T6XG>]; François Rancy, *ITU Radio Regulations – 110 Years of Success*, ITU NEWS MAGAZINE, 15 (May 2016) https://www.itu.int/en/itunews/Documents/2016-05/2016_ITUNews05-en.pdf [<https://perma.cc/4Y3M-FC4E>].

¹⁶² MADRY ET AL., *supra* note 158, at 93.

¹⁶³ *Id.*

¹⁶⁴ KERRETT, *supra* note 36, at 53.

international level with regards to spectrum given the role of national telecommunications agencies in regulating spectrum use.¹⁶⁵ Moreover, instead of procuring spectrum licenses or utilizing the already crowded unlicensed spectrum, Loon's market strategy is to partner with existing telecom providers.¹⁶⁶ Initially, Loon leaders wanted to lease spectrum, but another vision prevailed—lease the balloons to wireless carriers to expand their network coverage. This saves Loon spectrum license expenses and ensures that telecom companies are partners, not competitors.¹⁶⁷ In that regard, Loon is not an Internet service provider; individuals can only access the balloon network through their mobile network operator. Therefore, Loon now must not only gain support from national and occasionally international regulators, but also private actors, adding an additional layer of complexity.

In fact, Loon has worked with national regulatory agencies and existing telecom providers to conduct tests in various countries, signaling that the company is ready to enter into commercial deals with network operators around the globe. Prior to 2017, the Loon team completed beta tests in New Zealand, Brazil, Nevada, and Central California—typically using unlicensed spectrum.¹⁶⁸ In 2017, Loon partnered with the Peruvian government and Telefonica to deliver Internet access to flooded areas around Lima.¹⁶⁹ That same year, the FCC awarded Loon an experimental license to help connect

¹⁶⁵ Olivia Hatalsky et al., PowerPoint Presentation, ITU Outreach-Project Loon, (Sept. 14, 2017), https://www.itu.int/en/ITU-R/seminars/rrs/RRS-17-Americas/Documents/Forum/3_Google%20Olivia%20Hatalsky.pdf [<https://perma.cc/GNA7-P556>].

¹⁶⁶ The radio waves used in Loon's New Zealand launch operated on unlicensed spectrum in the 2.4GHz and 5.8GHz bands used in wi-fi. See Liam Tung, *Google Trials LTE in Project Loon's Balloons over Brazil*, ZDNET (June 17, 2014), <https://www.zdnet.com/article/google-trials-lte-in-project-loons-balloons-over-brazil/> [<https://perma.cc/Q82M-AJGD>].

¹⁶⁷ See Simonite, *supra* note 134.

¹⁶⁸ The Loon team has conducted tests with a few telecom companies that have spectrum and infrastructure in place to market the service to customers: Vodafone in New Zealand, Telstra in Australia, and Telefonica in Latin America. The Loon team has conducted research flights in California's Central Valley to test the strength of the balloon-powered Internet connection and small-scale field trials in rural areas of New Zealand. Loon successfully connected a local school in an isolated area of northeastern Brazil to the Internet for the first time. In 2016, the Indian government was approached about a pilot project with Loon, and in the summer of 2017, Kenya. See Simonite, *supra* note 134. See also FCC, Experimental Radio Station Construction Permit and License, File No. 0251-EX-CR-2017, <https://apps.fcc.gov/els/GetAtt.html?id=192339&x=> [<https://perma.cc/5SVZ-Y6ZJ>].

¹⁶⁹ After working with O3B networks, Level 3 and Ecologistica Peru to set up ground stations (which connect the balloons to the backbone of the Internet) and integrating the balloon-powered Internet into Telefonica's network, the balloon network transmitted more than 160 GB worth of data over the span of seven weeks, or about 2 million emails worth of data. Alastair Westgarth, *Helping out in Peru*, MEDIUM (May 17, 2017), <https://medium.com/loon-for-all/helping-out-in-peru-9e5a84839fd2> [<https://perma.cc/AT2V-Q4VN>].

people in Puerto Rico after Hurricane Maria.¹⁷⁰ To this day, Loon maintains its relationship with FCC, for example by recently encouraging the agency to “clarify that [USF] resources can be used by mobile carriers to support Loon and other new solutions that expand the preparedness and resilience of the communications networks.”¹⁷¹ Alphabet recently announced Loon’s first commercial contract with Telkom Kenya, a partially state-owned enterprise, to provide coverage to remote areas of Kenya in 2019.¹⁷²

Ultimately, the company is navigating multi-jurisdictional regulatory issues in light of its transnational infrastructure solution. In doing so, it is following the UN’s advice regarding multi-stakeholder partnerships for infrastructure development—a consequence of the current regulatory landscape concerning public goods such as spectrum and airspace.

2. *Outer Space Connectivity: The Global Broadband Space Race*

Alphabet is not alone in its mission to connect the globe; over the past several years, there has been renewed interest in satellite Internet. Remarkably, this revival has emerged despite satellite Internet companies’ troubled history,

¹⁷⁰ Loon obtained consent agreements to use land mobile radio spectrum in the 900 MHz band from existing carriers operating within Puerto Rico. By collaborating with the FCC, the FAA, FEMA, AT&T, T-Mobile, and many others, Loon was able to provide connectivity to 200,000 Puerto Ricans after Hurricane Maria. See Press Release, FCC, *FCC Grants Experimental License for Project Loon to Operate in Puerto Rico* (Oct. 7, 2017), <https://www.fcc.gov/document/fcc-grants-experimental-license-project-loon-puerto-rico/>, [https://perma.cc/29ZD-KXHY].

¹⁷¹ In the aftermath of Hurricane Marina, eligible telecommunications companies leveraged innovative technologies to reconnect their subscriber base, ranging from cells-on-wheels, to cells-on-UAVs, to Loon balloons. Loon called for project costs to be reimbursable for eligible telecommunications carriers from the High Cost Fund. See *In the Matter of The Uniendo a Puerto Rico Fund and the Connect USVI Fund*, Comments of Loon LLC, FCC, (July 26, 2018), [https://ecfsapi.fcc.gov/file/107261684001853/2018-07-26%20Loon%20Comments%20\(WC%2018-143\).pdf](https://ecfsapi.fcc.gov/file/107261684001853/2018-07-26%20Loon%20Comments%20(WC%2018-143).pdf) [https://perma.cc/8U5D-SBEW]; Monica Avellen, *Loon Highlights Value Of Mobile Carrier Partnerships in Puerto Rico*, FIERCEWIRELESS (July 30, 2018), <https://www.fiercewireless.com/wireless/loon-highlights-value-mobile-carrier-partnerships-puerto-rico/> [https://perma.cc/6T8Y-8EDF].

¹⁷² Safaricom Kenya Limited controls 71.2% of the total subscription followed by Airtel Kenya with 17.6% with Telkom Kenya and Finserve East Africa (Equitel) controlling a market share of 7.4% and 3.8% respectively. See Alfred Kipyegon Bett et al., *Analysis of Information Systems Capabilities and Performance of Firms in Telecommunications Industry, Kenya*, 6 INT’L J. SCI. RES. & MGMT. 319, 320 (2018); Jamal Carnette, *Is Alphabet’s Craziest Moonshot Starting to Pay Off?*, MOTLEY FOOL (July 26, 2018), <https://www.fool.com/investing/2018/07/26/is-alphabets-craziest-moonshot-starting-to-pay-off.aspx> [https://perma.cc/YNW6-ZNKL].

marked by bankruptcies in the 1990s.¹⁷³ A handful of satellite companies, including OneWeb, SpaceX, and O3b (or “Other 3 Billion”), are developing fleets of smaller satellites to circumscribe Earth and provide truly global Internet service. The mega-constellations that they propose are on a mission to bridge the digital divide. In 2017, executives of these companies spoke to the United States Senate Committee on Commerce, Science, and Transportation about the future of the commercial satellite industry and discussed regulations that could aid in their efforts to bridge the divide and enhance social and economic development.¹⁷⁴ So far, none of the companies have deployed their full constellations to space, setting off a global broadband space race to see who will obtain first mover advantage.

The innovators at SpaceX, OneWeb, and O3b are working to change the cost structure and quality of satellite service. Satellite Internet has been an option for people living in rural areas (especially in developed countries) for decades—albeit an expensive alternative, often providing patchy, low-quality service. By launching thousands of smaller satellites in lower orbits, next-generation satellite service providers expect to see less latency. Traditionally, satellite communications systems “hover” in geosynchronous orbit, about 22,000 miles above the Earth, whereas satellites in LEO float roughly 100 to 1,250 miles above Earth and are not aligned with Earth’s rotational period.¹⁷⁵

¹⁷³ The trend to launch commercial small satellite constellations began in the early 1990s. As a result of insufficient market demand and technical issues, the companies Iridium and Globalstar went through bankruptcy proceedings. Iridium in fact went bankrupt one year after deployment. See RAM S. JAKHU & JOSEPH N. PELTON, GLOBAL SPACE GOVERNANCE: AN INTERNATIONAL STUDY 358-61, 370 (2017). Since the 1990s, projected costs of satellite operations have decreased by two orders of magnitude. See BHAVYA LAL ET AL., INST. FOR DEF. ANALYSES, GLOBAL TRENDS IN SMALL SATELLITES 284 (2017), <https://www.ida.org/-/media/feature/publications/g/gl/global-trends-in-small-satellites/p-8638.ashx> [<https://perma.cc/WS8R-L8WE>].

¹⁷⁴ The Commercial Satellite Industry: What’s Up and What’s on the Horizon: Hearing Before the S. Comm. on Commerce, Sci., and Transp., 115th Cong. (2017). Recent scholarship indeed argues that companies specializing in satellite and space technology can positively contribute to the global development agenda. See JAKHU & PELTON, *supra* note 173, at 520–23, 529–36; Small Satellites and the U.N. Sustainable Development Goals, in MADRY ET AL., *supra* note 158, at 65; European Space Agency, *Sustainable Development With A Little Help From Space*, ESA.INT, (May 31, 2016), http://www.esa.int/Our_Activities/Preparing_for_the_Future/Space_for_Earth/Sustainable_development_with_a_little_help_from_space/ [<https://perma.cc/X4ZP-BS5L>]; EUROCONSULT EC, PROSPECTS FOR THE SMALL SATELLITE MARKET, (2018), <http://www.euroconsult-ec.com/research/smallsats-2018-brochure.pdf> [<https://perma.cc/B8VL-5P7N>].

¹⁷⁵ See Klint Finley, *Can These Small Satellites Solve The Riddle Of Internet From Space?*, WIRED (March 1, 2018), <https://www.wired.com/story/can-these-small-satellites-solve-the-riddle-of-internet-from-space/> [<https://perma.cc/R7PB-Z4KV>]. In geostationary orbit, a

In higher orbits, connections can lag quite a bit, making real-time connections impracticable. Yet geostationary satellites provide more geographic coverage on Earth with fewer satellites, whereas a LEO network requires thousands of smaller satellites to achieve the same amount of geographic coverage.¹⁷⁶ For instance, SpaceX's constellation will have nearly 12,000 satellites.¹⁷⁷ Although smaller LEO satellites are a fraction of the cost of traditional ones, launching thousands of them could still result in significant total project costs.¹⁷⁸ Consequently, many companies are working to reduce the costs of launching satellites into space.¹⁷⁹ Some companies are also considering smaller fleets that will operate in middle Earth orbits.¹⁸⁰

Like Loon, satellite companies are also on mission to address the problem of the global last mile. For example, OneWeb's mission is to bridge the global digital divide by 2027, and the company frequently uses the term "global" to market its ambitions: OneWeb is "building a new *global* knowledge infrastructure accessible to all" and "can address the most

satellite "hovers" over one spot because the satellite's orbital period aligns with Earth's rotational period.

¹⁷⁶ See Lawrence D. Roberts, *A Lost Connection: Geostationary Satellite Networks and the International Telecommunication Union*, 15 BERKELEY TECH. L.J. 1095, 1099–100 (2000); Stewart Sanders, *The New Space Race is All About Satellites: Pros and Cons of Each Orbit*, THE NEXT WEB (Nov. 3, 2018), <https://thenextweb.com/contributors/2018/11/03/the-new-space-race-is-all-about-satellites-pros-and-cons-of-each-orbit/> [<https://perma.cc/E3QX-EKEP>].

¹⁷⁷ See Finley, *supra* note 175. Recently, OneWeb received FCC approval to serve U.S. customers with a constellation of 720 satellites, far from the thousands envisioned by some LEO networks. FCC FACT SHEET, *OneWeb Market Access Grant*, Order and Declaratory Ruling - IBFS File No. SAT-LOI-20160428-00041, https://transition.fcc.gov/Daily_Releases/Daily_Business/2017/db0601/DOC-345159A1.pdf [<https://perma.cc/X8D4-XZEM>].

¹⁷⁸ Small satellites are less massive, which substantially reduces launch costs. Also, the designing, manufacturing and launching stages can be completed in less than two years, lowering mission costs. See JAKHU & PELTON, *supra* note 173, at 366. In 2015, OneWeb's goal was to build satellites at \$500,000 per satellite or less. The total estimated cost of building, launching and operating the constellation is \$3.5 billion. See Caleb Henry & Brian Berger, *Amid Concerns, OneWeb Gets Vague About Constellation's Cost*, SPACENEWS (Sept. 12, 2018), <https://spacenews.com/amid-concerns-oneweb-gets-vague-about-constellations-cost/> [<https://perma.cc/TCF7-YBZK>]. Recently, OneWeb refused to affirm satellite unit costs or total program costs, disclosing only that each satellite cost less than \$1 million to produce. Caleb Henry, *OneWeb Vouches for High Reliability of Its Deorbit System*, SPACENEWS (July 10, 2017), <https://spacenews.com/oneweb-vouches-for-high-reliability-of-its-deorbit-system/> [<https://perma.cc/2PYC-X9GT>].

¹⁷⁹ See Finley, *supra* note 175.

¹⁸⁰ Satellites in middle Earth orbit still provide lower latency connections than satellites in geostationary orbit, but the total number of satellites needed to provide complete coverage would also decrease. See Finley, *supra* note 175.

demanding *global* connectivity challenges. . . .”¹⁸¹ In many respects, its mission is similar to Loon’s. For instance, the goal is not only to provide rural areas with Internet access, but to “assure *global* communications” by bringing areas back online after natural disasters. Aspiring to be a global infrastructure solution, companies like OneWeb and O3b are also creating go-to-market strategies that opt for Loon’s approach, planning to partner with other operators who have already leased spectrum to extend current networks. In other words, some companies plan on selling bandwidth to Internet service providers or mobile providers, rather than directly to end-users.¹⁸²

Next-generation satellite companies’ next steps include planning viable commercial deployments and gaining regulatory approvals. To successfully bring global infrastructure to market, companies will have to work with international and national public agencies regulating space and spectrum to gain necessary approvals. However, companies planning large-scale satellite deployments could face legal and regulatory issues because current international and national space regulations do not fully address large-scale satellite constellations.¹⁸³ Companies will have to work with an outdated and complex regulatory framework, potentially undermining the ability of satellite broadband infrastructure to bridge the global digital divide.

a. Connectivity via Mega-Constellations and the Regulatory Framework

Plans to launch next-generation satellites will trigger the application of national and international air and space laws. As explained above, states have “complete and exclusive sovereignty over the airspace above [their] territory” in accordance with Article 1 of the Chicago Convention.¹⁸⁴ However, the concept of complete and exclusive sovereignty over the airspace above a state’s territory has evolved in the last century. The concept of *usque ad coelum*, whereby sovereignty extended “to the heavens,” was modified by the Outer Space Treaty of 1967. The treaty “set vertical limits to a [s]tate’s sovereignty over airspace.”¹⁸⁵ While the boundary between airspace and outer space is not clear, states admit that their sovereignty over airspace is no longer

¹⁸¹ ONEWEB, <http://www.oneweb.world/> [<https://perma.cc/4S9G-3CH7>]. Even the web address has a “.world” domain name.

¹⁸² *See id.*; *Seamlessly Scaling our O3b Fleet to Meet Exponential Demand for Connectivity*, SES (Feb. 26, 2018), <https://www.ses.com/newsroom/seamlessly-scaling-our-o3b-fleet-meet-exponential-demand-connectivity> [<https://perma.cc/32RC-L9AU>]; Finley, *supra* note 175.

¹⁸³ *The Commercial Satellite Industry: What’s Up and What’s on the Horizon: Hearing Before the S. Comm. on Commerce, Sci., and Transp.*, 115th Cong. (2017).

¹⁸⁴ Convention on International Civil Aviation, *supra* note 128, at art. 1.

¹⁸⁵ Legal Committee, ICAO, *Revisiting the International Air Services Revisiting the International Air Services Transit Agreement Of 1944*, at 2 (No. LC/36-IP/2 2015).

infinite and ends where outer space begins.¹⁸⁶ Launched satellites transiting through a state's airspace must therefore accord international and national civil aviation regimes. Upon reaching the undefined vertical limits of national airspace and entering outer space, however, a different regulatory framework applies.¹⁸⁷

The legal framework for outer space activities consists of five international treaties which were adopted between 1967 and 1979, including the Outer Space Treaty, the Liability Convention, and the Registration Convention. Many terms in the treaties have been adopted by national space legislation. With 107 signatories, the Outer Space Treaty—which contains basic principles for space activities—is considered to contain principles of customary international law.¹⁸⁸ One such principle is that outer space and celestial bodies are global commons, not subject to national sovereignty and jurisdiction.¹⁸⁹ According to the Outer Space Treaty and the Liability Convention, a launching state (i.e., a state that launches or procures the launching of a space object or from whose territory or facility an object is launched) is internationally liable for damage caused by its space object.¹⁹⁰ Determining the launching state can be complex with regard to multinational launches because more than one state may be implicated, and the international

¹⁸⁶ There is no international agreement that defines the altitude that is considered outer space, nor the demarcation between airspace and outer space. *See* JAKHU & PELTON, *supra* note 173, at 653–54.

¹⁸⁷ *See id.* at 311.

¹⁸⁸ The Outer Space Treaty is considered to contain principles of customary international law, binding signatories and non-signatories alike. The customary principles can be found in Articles I–IV, VI, VII, VIII and arguably also Art. IX. *See* Rada Popova & Volker Schaus, *The Legal Framework for Space Debris Remediation as a Tool for Sustainability in Outer Space*, 5 *AEROSPACE* 1, 4 (2018), <https://www.mdpi.com/2226-4310/5/2/55> [<https://perma.cc/GYV3-YST9>].

¹⁸⁹ This is stated in the first paragraph of Article 1 in the Outer Space Treaty, “according to which the use and exploration and use [sic.] of outer space should be regarded as the ‘province of all mankind.’” Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies, art. 1, Jan. 27, 1967 (hereinafter “Outer Space Treaty”); *see also* Popova & Schaus, *supra* note 189, at 4.

¹⁹⁰ Convention on International Liability for Damage Caused by Space Objects, art. I(c), II, III, Nov. 29, 1971; Outer Space Treaty, *supra* note 189, at art. VII. *See also* CHRISTOPHER D. JOHNSON, SECURE WORLD FOUND., LEGAL AND REGULATORY CONSIDERATIONS OF SMALL SATELLITE PROJECTS 8, https://swfound.org/media/188605/small_satellite_program_guide_-_chapter_5_-_legal_and_regulatory_considerations_by_chris_johnson.pdf (noting four potential categories of launching states: (1) the state that launches; (2) the state that procures the launch; (3) the state from whose territory an object is launched; and (4) the state from whose facility an object is launched) [<https://perma.cc/R2QH-5MM4>].

rights and obligations of the launching states may differ.¹⁹¹ Depending on treaty obligations, each launching state may be internationally responsible and potentially liable under international space law—prompting governments to regulate, supervise, license and oversee satellite projects.¹⁹² This authorization requirement extends to small satellite constellations, which are legally considered “space objects,” independent of their size.¹⁹³ Moreover, if a state is a signatory to the Registration Convention, it is required to register launched space objects nationally and internationally with the UN.¹⁹⁴ Even if there is more than one launching state, there is only one state of international registry for each satellite.¹⁹⁵

The launching of communications satellites also triggers the application of national and international spectrum laws. Each country has its own frequency allocation plan, which mostly overlaps with the frequency allocation plan put forth by ITU members for satellite services and communications.¹⁹⁶ Electromagnetic spectrum for radio communications between satellites and ground stations is limited, requiring responsible use and sharing of the finite natural resource.¹⁹⁷ The process of obtaining a license to operate a satellite in a certain market normally involves filing an application for the intended frequency with the national telecommunications agency in the country where the satellite will provide connectivity.¹⁹⁸ Under the ITU Radio

¹⁹¹ See JOHNSON, *supra* note 190, at 10 (“Once the launching states question is answered, it is wise to investigate what international space treaties these states are party to. . . . Determining the launching states will show which ones have what international responsibilities under international space law, including registration with the UN.”).

¹⁹² States bear international responsibility for national activities in outer space, no matter if the activities are performed by governmental or commercial organizations. Thus, under the various treaties states are required to authorize and license national space activities and ensure continuous supervision and control. See Outer Space Treaty, *supra* note 189, at art. VI; JOHNSON, *supra* note 190, at 10. Some satellite constellations require additional authorizations and permits from government authorities. See JAKHU & PELTON, *supra* note 173, at 371.

¹⁹³ See JAKHU & PELTON, *supra* note 173, at 371.

¹⁹⁴ International space law encourages and sometimes mandates states to register their space objects in the international registry in order to notify other states of their space activities. The United Nations Office for Outer Space Affairs (UNOOSA) keeps the space object registry. Some of the information is voluntarily supplied to UNOOSA by states, whereas registration is mandatory for states party to the 1975 Registration Convention. See JAKHU & PELTON, *supra* note 173, at 371; JOHNSON, *supra* note 190, at 10.

¹⁹⁵ JOHNSON, *supra* note 190, at 14.

¹⁹⁶ MADRY ET AL., *supra* note 158, at 93.

¹⁹⁷ See JOHNSON, *supra* note 190, at 3.

¹⁹⁸ Entities may file a petition for a declaratory ruling to access the U.S. market using a non-U.S.-licensed space station. See 47 C.F.R. § 25.137 (2019); In the Matter of Streamlining Licensing Procedures for Small Satellites, FCC 18-44, April 17, 2018, para. 23 https://transition.fcc.gov/Daily_Releases/Daily_Business/2018/db0417/FCC-18-44A1.pdf [<https://perma.cc/UGL4-G8JN>]. If a satellite company wants to reach a certain national

Regulations, satellites may not be operated without a license “by or on behalf of the government of the country to which the station in question is subject.”¹⁹⁹ For example, under the U.S. Communications Act of 1934, satellite operators must be issued a license allowing communications to and from the United States or from any U.S. satellite.²⁰⁰ Moreover, the same Act grants the FCC authority to implement the ITU’s Radio Regulations.²⁰¹ Upon meeting the filing requirements of the national licensing agency, which may take years, the national administration, as a member of the ITU, then notifies the ITU and enters the satellite’s frequency on the master chart of internationally used frequencies.²⁰² The filings with the ITU must detail the “specific frequency bands that are to be utilized as well as the specific orbits and orbital patterns to be used by the intended system.”²⁰³ Filing initiates the publication stage of the ITU process, whereby the ITU checks the submitted information for completeness and publishes the filing for members to review and comment.²⁰⁴ This stage does not confer rights or any priority on the filing administration. The next stage of the process, called “coordination,” entails negotiations between affected countries and the notifying administration.²⁰⁵ The last stage

market, it must apply for a spectrum license with the relevant telecommunications agency. For example, recently Canada-based Telesat and Space Norway, with constellations authorized by other administrations, applied for U.S. market access with the FCC. The FCC declined Telesat’s request that its constellation should have spectrum priority based on its filing date with the ITU facilitated by the Canadian administration. *See* In the Matter of Telesat Canada, FCC 17-147 (Nov. 3, 2017) <https://www.fcc.gov/document/telesat-ngso-market-access-grant> [<https://perma.cc/6YJM-RKGGX>]; Caleb Henry, *FCC Grants Telesat LEO Market Access Despite Viasat Protests*, SPACE NEWS (Nov. 6, 2017), <https://spacenews.com/fcc-grants-telesat-leo-market-access-despite-viasat-protests/> [<https://perma.cc/D9DJ-W532>].

¹⁹⁹ ITU Radio Regulations, No. 18.1 (2016), <http://search.itu.int/history/HistoryDigitalCollectionDocLibrary/1.43.48.en.101.pdf> [<https://perma.cc/V6PS-UBZX>].

²⁰⁰ 47 U.S.C. § 301(d), (f) (2018).

²⁰¹ 47 U.S.C. § 303(r) (2018).

²⁰² *See* JOHNSON, *supra* note 190, at 4.

²⁰³ States that are party to the Outer Space Treaty and Registration Convention enact national laws embodying international standards, which require detailed information to issue a license. The information to be filed typically includes, “the number and size of the satellites and many details of their technical characteristics, plus business plans for the services to be provided, the builder of the satellite and related contractual details, the financial details as to financing, contractors to build and launch the satellite, specific details as to mitigation procedures to lessen the possibility of creating orbital debris. . . .” MADRY ET AL., *supra* note 158, at 94-95.

²⁰⁴ Int’l Telecomm. Union, *How Satellites are Brought into Service: A Brief Account of the Regulatory Steps for Satellites Using Frequency Bands Falling Under the “Coordination Procedures,”* WORLD RADIO CONFERENCE 2000 <https://www.itu.int/newsarchive/wrc2000/presskit/how-sat.html> [<https://perma.cc/QDB4-GY4D>].

²⁰⁵ *Id.*

is notification.

With LEO mega-constellations several steps ahead of regulation, scholars point to several issues requiring international cooperation and intersystem coordination: namely, the ITU's "first come, first served" principle, inadequate processes for orbital congestion and spectrum interference, and ineffective space debris mitigation requirements.

First, although the ITU's "first come, first served" principle aims to provide equitable access to spectrum resources, some have noted the tendency of countries to abuse the system by filing "paper satellites."²⁰⁶ Today, the worry is that companies launching in LEO might go "regulation shopping" to a "government of convenience," which might require fulfillment of fewer national requirements and deployment deadlines before filing applications to the ITU for large constellations.²⁰⁷ Some question whether the ITU's approval process for frequencies and orbital allocation for mega-LEO constellations "represents a reasonable economic, regulatory, and safety assessment model to follow."²⁰⁸

Second, scholars note the inadequate processes to mitigate orbital congestion and spectrum interference. Present-day national and international regulatory procedures do not address the question of "how many new mega-LEO systems can be plausibly deployed. . . ."²⁰⁹ With so many companies competing in the global broadband space race, absent international coordination, future orbital congestion and harmful spectrum inference are a given.²¹⁰ To tackle this issue, international cooperation would need to address which frequency plans and constellation orbital deployment locations are "reasonable in terms of approving these systems for launch."²¹¹ Currently, regulation on orbital allocation only specifies that geostationary satellites have

²⁰⁶ See MADRY ET AL., *supra* note 158, at 95 ("Some countries have in the past accelerated (and abused) the national review process to file so-called "paper satellites" with certain technical characteristics with the ITU simply to take advantage of the ITU's "first come, first served" principle. Needless to say, such practices undermine the principle of equitable access to spectrum resources. The ITU now has created charges for satellite filing and other milestone procedures to limit such 'paper filings.'").

²⁰⁷ JAKHU & PELTON, *supra* note 173, at 360. The FCC imposes deployment deadlines to prevent companies from "warehousing" spectrum, laying claim to frequencies and barring them from use by other companies. See Caleb Henry, *Oneweb Asks FCC to Authorize 1,200 More Satellites*, SPACENEWS (Mar. 20, 2018), <https://spacenews.com/oneweb-asks-fcc-to-authorize-1200-more-satellites/> [<https://perma.cc/MK6W-QUCY>].

²⁰⁸ JAKHU & PELTON, *supra* note 173, at 360.

²⁰⁹ MADRY ET AL., *supra* note 158, at 96.

²¹⁰ JAKHU & PELTON, *supra* note 173, at 5, 373. Scholars have commented on possible interference by LEO satellites with geosynchronous satellites. See MADRY ET AL., *supra* note 158, at 89.

²¹¹ MADRY ET AL., *supra* note 158, at 96.

protected status against non-geostationary satellites.²¹² No procedures or guidelines at the ITU or national level exist that detail the priority of orbital allocations for LEO constellations.²¹³ As more companies are applying for licenses for their constellations, some propose instituting a moratorium on deploying LEO satellites, at least until a reasonable global decision-making process is established to equitably prioritize mega-LEO deployments internationally.²¹⁴

Third, large-scale LEO satellite deployments could potentially lead to the aggravation of orbital space debris. Since the 1950s, when space missions commenced, it is estimated that more than 621,000 human-made objects that are greater than one centimeter in diameter have come to reside in orbit.²¹⁵ With the addition of potentially 15,000 small communications satellites in LEO, ten times the amount currently in orbit, some worry that, absent properly allocated altitudes for constellations, there will be more collisions of space objects, leading to more space debris.²¹⁶ The ever-increasing number of satellites and human-made objects colliding in space could quickly reach the tipping point for runaway debris, also called the Kessler syndrome, jeopardizing sustainable use of space.²¹⁷ Although there are some international guidelines in place,²¹⁸ voluntary orbital cleanup projects underway,²¹⁹ and

²¹² *See id.* at 95, 100 (“The filing process is different in the case of [geostationary] satellite networks. This is because it is necessary to seek specific orbital locations in the GEO belt and to identify slots that might be available that are not occupied by existing satellite networks.”).

²¹³ *Id.* at 96.

²¹⁴ *Id.*

²¹⁵ There are currently in orbit about 21,000 man-made objects that are more than 10 centimeters; about 600,000 objects measuring between 1 and 10 centimeters; and hundreds of millions of objects less than 1 centimeter. *See* Ram S. Jakhu et al., *Regulatory Framework and Organization for Space Debris Removal and On Orbit Servicing of Satellites*, 4 J. SPACE SAFETY ENG’G 129 (2017), <https://www.sciencedirect.com/science/article/pii/S2468896717300836> [<https://perma.cc/3GFH-RSDF>].

²¹⁶ MADRY ET AL., *supra* note 158, at 82.

²¹⁷ *Id.* at 96.

²¹⁸ The Inter-Agency Space Debris Committee has published Space Debris Mitigation guidelines, urging removal of satellites from orbit within twenty-five years of when their mission ends. *See id.* at 81; Popova & Schaus, *supra* note 188, at 3, 10–11. Some companies have committed to deorbiting out-of-use satellites within five years of the mission’s end date. *See* Caleb Henry, *OneWeb Vouches for High Reliability of Its Deorbit System*, SPACENEWS (July 10, 2017), <https://spacenews.com/oneweb-vouches-for-high-reliability-of-its-deorbit-system/> [<https://perma.cc/H7U9-JFQT>].

²¹⁹ Some cleanup efforts are currently underway, but there is no international agreement on the obligations of countries or companies to clean up orbital debris. Many actors are developing projects to capture and de-orbit small satellites, including Swiss-based CleanSpace One, the U.S. Defense Advanced Research Projects Agency, and the German space agency. *See* MADRY ET AL., *supra* note 158, at 86. Cooperation between industry, government, and

national mechanisms requiring companies to detail space debris mitigation plans,²²⁰ most national and international legal frameworks for space activities do not impose legal obligations for debris removal.²²¹ This deficiency has caused some scholars to advocate for modernized global space governance to address the risk of runaway debris,²²² such as establishing an international regulatory framework and intergovernmental organization with an active debris removal (ADR) space debris mandate,²²³ and amending the Liability Convention to include ADR obligations.²²⁴

In sum, like Loon, satellite companies will also have to navigate multi-jurisdictional regulatory issues in light of the envisioned global infrastructure. In doing so, they will have to rely on various partnerships to not only operate under existing spectrum licenses, but, more importantly, address regulatory gaps concerning international space activities and mega-constellations.

B. Disruptive Potential of the Innovations: Changing the Economics of Broadband Access?

Internet service relies on telecommunications infrastructure as the medium through which data flows, including cables (copper wires or optical fiber) and electromagnetic waves (for satellite, wireless, mobile networks). The projects identified in the previous Section aim to globally transmit Internet

academia is also common. See Tereza Pultarova, *This Space Junk Removal Experiment Will Harpoon & Net Debris in Orbit*, SPACE (Apr. 6, 2018), <https://www.space.com/40221-space-junk-debris-sweeper-experiment.html> [perma.cc link unavailable].

²²⁰ National regulations usually require companies to at least specify their orbital debris mitigation procedures. In France, national laws allow for enforcement of space debris removal guidelines, and companies are penalized if they have not timely deorbited. MADRY ET AL., *supra* note 158, at 82, 95. In contrast, the FCC and other U.S. federal agencies require companies to submit plans during the licensing process that ensure compliance with the space debris mitigation guidelines—specifically, they must show that the spacecraft will deorbit within twenty-five years of the mission’s end. See TED WACKLER, WHITE HOUSE OFFICE OF SCI. AND TECH. POL’Y, ORBITAL DEBRIS REPORT 3 (2017), <https://www.whitehouse.gov/wp-content/uploads/2017/12/08-14-17-OSTP-Orbital-Debris-Report.pdf> [https://perma.cc/M246-38RP].

²²¹ See Popova & Schaus, *supra* note 188, at 1.

²²² MADRY ET AL., *supra* note 158, at 90.

²²³ Jakhu et al., *supra* note 215, at 129 (“[M]itigation efforts must be accompanied by active debris removal (ADR) of *existing* pieces of debris from space in order to effectively protect the space environment . . . [which] cannot be effectively undertaken due to the existing complex legal problems, primarily at international level.”).

²²⁴ See MADRY ET AL., *supra* note 158, at 86 (noting that launching states are already liable for collisions in space). See also Popova & Schaus, *supra* note 188, at 10 (noting “no agreement on whether space debris should be considered to be space objects, as per the definition of ‘space object’ of Art. I . . . of the Liability Convention”); Jakhu et al., *supra* note 215, at 131 (arguing a piece of space debris is a space object).

connectivity with non-terrestrial infrastructure—circumventing the need to dig trenches, install cables, and determine property rights. Loon’s balloon deployments, the new LEO mega-constellations, and Facebook’s drones are possible *wireless* solutions to the last mile problem, relying on overcrowded radio waves. This Section explores how these projects could potentially do more than bridge the access gap—including disrupting the telecommunications industry and changing the economics of broadband access.

The analysis below contemplates the disruptive potential of Loon’s balloons and mega-constellations, the two projects that are furthest along. Upon briefly examining the theory of disruptive innovation, the Section proceeds to compare the projects to other wireless and wireline solutions, as well as to each other. First, the Section considers how both projects directly compete with incumbent wireless providers in their respective sectors: Internet balloons via non-terrestrial “floating” cell towers could potentially challenge mobile operators, while LEO satellite mega-constellations may challenge the dominance of geostationary satellites. Second, the Section considers whether the two infrastructure projects might also disrupt the customer base of wireline Internet service providers. The Section ends by considering whether, between the two wireless infrastructure solutions, one project has more disruptive potential.

Relying on the theory of disruptive innovation to assess the two wireless projects’ disruptive potential, it is unclear whether either project will prove to be disruptive. Not all innovations are considered disruptive, in the sense that they alter or transform established markets and incumbent players. Scholars argue that an innovative technology is disruptive when it alters the status quo and creates a new market, rapidly.²²⁵ One theory of disruptive innovation, articulated by Clayton Christensen in the *Innovator’s Dilemma*, explains how new technologies may cause incumbents to fail.²²⁶ First, the disruption begins when a smaller company with fewer resources gains a foothold in a market that an incumbent has overlooked. This process usually means creating a market where none previously existed by offering a lower cost, inferior product—in essence, converting non-consumers into consumers.²²⁷ Typically, incumbents’ customers initially ignore the new entrant because they believe the offered service is inferior. The new entrant begins to improve the quality of its service over time—enough so that

²²⁵ JAKHU & PELTON, *supra* note 173, at 363–64.

²²⁶ See generally CLAYTON CHRISTENSEN, *THE INNOVATOR’S DILEMMA: WHEN NEW TECHNOLOGIES CAUSE GREAT FIRMS TO FAIL* (1997).

²²⁷ JAKHU & PELTON, *supra* note 173, at 366–67; Clayton M. Christensen et al., *What is Disruptive Innovation?*, HARV. BUS. REV. (Dec. 2015) at 5, <http://pedrotrillo.com/wp-content/uploads/2016/01/Whatisdisruptiveinnovation.pdf> [<https://perma.cc/M4QC-LTG8>].

incumbents' customers begin to notice. Only when incumbents' customers switch to the new entrant's product has disruption occurred. To summarize, the three key characteristics of disruptive innovation include (1) adoption by an underserved market, (2) initially inferior performance at a lower cost, and (3) greater market capture as quality improves.²²⁸

Applying the first factor to mega-constellations and Loon's balloons, there are signs that disruption could occur as both infrastructure projects target underserved markets. First, by seeking to connect the unconnected, both projects attempt to turn non-consumers into consumers. Positioning itself to be a wireless Internet provider to unconnected markets, Loon's balloons could directly compete with existing providers, such as mobile network operators, although the company has chosen not to.²²⁹ Likewise, LEO satellite companies may directly compete with incumbent geostationary satellite telecommunications operators for underserved markets.

The second characteristic of disruptive innovation, inferior performance at lower costs, is pertinent to analyzing the disruptive potential of LEO satellites. The rationale driving mega-constellations is the use of lower orbits to improve latency. Absent more satellite deployments and testing at LEO, it is difficult to ascertain the possible service quality. A number of factors could negatively impact service quality, including weather, availability of spectrum, and potential interference with other satellites' signals. But the potential unreliability has also been dealt with by design through redundancy in the form of larger constellations, a potentially key driver affecting the economics of access.²³⁰ Even with the uncertainty regarding quality, some argue that LEO satellites will nevertheless be disruptive vis-à-vis geostationary satellites—offering at least equal or superior performance at

²²⁸ JAKHU & PELTON, *supra* note 173, at 364.

²²⁹ From the outset Loon has chosen to partner with, instead of compete with, incumbent wireless carriers. In considering Loon balloons' disruptive potential vis-à-vis mobile providers, the decision to supply network capacity to local mobile network operators renders the disruptive innovation analysis inapplicable. See Tung, *supra* note 166 and accompanying text; SIMONITE, *supra* note 134; *supra* text accompanying note 167.

²³⁰ See JAKHU & PELTON, *supra* note 173, at 367; Caleb Henry, *LEO and MEO Broadband Constellations Mega Source of Consternation*, SPACE NEWS (Mar. 13, 2018), <https://spacenews.com/divining-what-the-stars-hold-in-store-for-broadband-megaconstellations/> (“[The industry is] recognizing that the megaconstellation approach to capacity expansion represents a sea-change in the economics of the satellite industry. . . We see this recognition across the traditional value chain, from manufacturing and launch through operators and service providers, as well as customers . . . Manufacturers are angling for constellation construction contracts by promoting new smallsat platforms. Launch providers are designing adapters and deployers for constellations, or building new rockets specifically sized for dedicated smallsat missions. And operators of ground-based satellite gateways are installing new antennas around the world to provide turn-key solutions for constellation operators.”) [<https://perma.cc/Q39F-PXFU>].

lower costs.²³¹ Over time, this may result in incumbent satellite providers' customers switching service providers, signaling the arrival of a disruptive innovation.

It is not enough to consider whether either project is capable of effectively competing with companies in their respective sectors. For example, many have wondered if satellite Internet service providers can effectively compete with traditional broadband, such as cable Internet.²³² Whether either project will disrupt the business models of traditional wireline service providers greatly depends on the targeted location for service provision. The two wireless solutions aim to serve unconnected markets, where there is limited availability of wireline services, or as identified in Part II, *supra*, the last mile portions of infrastructure that incumbent wireline providers find difficult and uneconomical to develop. In areas where wireline development is uneconomical, the infrastructure projects cannot "disrupt" certain markets because there is already limited or nonexistent competition.

One of the two wireless solutions may prove disruptive in urban markets in developing countries. Even if these areas receive more attention by incumbent wireline providers, the speed of mobile broadband or satellite Internet could be greater than what is currently available with wireline services or roughly equal at lower costs of access. Notwithstanding the increase in high-speed fixed-broadband subscriptions globally, the lack of high-speed wireline connections in developing countries is notable: the penetration rate is 6 percent (1.6 percent excluding China), compared with 24 percent in developed countries.²³³ In developing countries, mobile broadband is more affordable than fixed broadband, resulting in Internet access being increasingly mobile.²³⁴

In contrast, it is unlikely that either of the two wireless solutions will prove disruptive in urban markets in developed countries, where fixed broadband connections are more common. Given that mobile and satellite broadband have not yet reached the speeds of wireline services, it is unlikely that the inferior performance of Loon or LEO satellite Internet will result in great uptake by incumbents' customers in developed countries. Although satellite Internet connection speeds are sky-rocketing, latency remains an issue. Both mobile broadband and satellite broadband speeds are still

²³¹ The next-generation satellites have been coined "NewSpace" ventures and are predicted to challenge the economic models for existing satellites in a "disruptive way" with low cost technology and improved designs. Some wonder if mega-constellations will "lead to the obsolescence of existing operators' economic models." JAKHU & PELTON, *supra* note 173, at 360, 362.

²³² See Finley, *supra* note 175.

²³³ See INT'L TELECOMM. UNION, *supra* note 5, at 6.

²³⁴ *Id.* at 4–5.

considered slow compared to speeds available through a wireline connection. Connection reliability and quality are important to users in developed countries, who expect less latency to conduct real-time activities such as videoconferencing. Most customers of incumbent carriers would consider the quality of wireless services to be inferior to wireline services, and they would not subscribe to a low-cost wireless service unless quality significantly improves. However, switchover to wireless-based Internet providers might be conceivable in the future with greater advancements in 5G and alleviation of spectrum congestion by governments.²³⁵

Nevertheless, in last mile areas where limited competition currently exists, Loon and LEO companies eventually may have to compete with wireline providers. At low price points, former noncustomers may begin subscribing to lower performance Internet service, such as mobile broadband. In these areas, individuals may initially find slow or unreliable Internet adequate to suit their needs, but over time the increased adoption and market demand may spark the need for and interest in higher reliability connections and, thus, wireline development.²³⁶ Such was the case in India, which moved from the bottom of the list for “mobile broadband penetration to the world’s largest mobile data-consuming nation.”²³⁷ As a result of increased adoption and greater demand, telecom companies in India are responding by installing more wireline infrastructure in the form of optical fiber.

Given that the two infrastructure solutions posed by Loon balloons and LEO satellites are both in the wireless space and are high altitude projects, it is worth asking whether one solution is more likely to be disruptive than the other in developing countries. Applying the key characteristics of disruptive innovation, determining which project has more disruptive potential requires predicting how much demand each wireless solution will generate. Demand will depend on the cost of access, an important driver for adoption in developing countries,²³⁸ and the regulatory landscape.

If the mission is to connect the unconnected, Loon may have greater success because of the relatively low cost of its proposed infrastructure. First, Loon balloons were designed from the outset to be inexpensive, requiring limited amounts of expensive hardware. The balloons utilize solar energy, an

²³⁵ Recent wireless broadband and mobile communications technology has increased the demand for radio frequencies and requires governments to find a solution to optimize spectrum use. See JOVAN KURBALIJA, AN INTRODUCTION TO INTERNET GOVERNANCE, 39 (2016).

²³⁶ According to survey data, many Nigerians perceive broadband as expensive relative to its quality, and are willing to pay, on average, an extra 166% per month for more reliable and faster broadband. *Willingness to Pay for Broadband in Nigeria*, *supra* note 63, at 428.

²³⁷ Luke Beirne, *How Long Are We Going to Need that City?*, TRADE ARABIA (Sept. 12, 2018), http://www.tradearabia.com/news/REAL_345017.html [<https://perma.cc/CZD8-4U9N>].

²³⁸ See *infra* notes 329–331 and accompanying text.

inexpensive energy source when compared to the cost of fuel required to move satellites to their orbital positions in space. Moreover, the decision to complement the capabilities of mobile operators translates into savings on spectrum licenses for Loon. With the ability to provide between twenty to thirty times broader coverage than cell towers, savings for mobile operators from balloon “cell towers” may result in lower costs of access in some markets.²³⁹ Additionally, launching Loon balloons, although far from a simple task, is rather mechanical, does not necessarily rely on third parties, and is far less expensive than rocket launches. In contrast, satellite companies must rely on the launch industry, such that the availability and costs of launch vehicles could potentially “undermine the whole economic advantage” of small satellites.²⁴⁰ In sum, Internet access via Loon balloons is arguably more cost-effective. As satellite companies experiment with variables affecting cost — such as the orbit height, satellite size, and launch vehicle availability—they will inevitably have to consider whether their business model offers access at rates that subscribers can afford.

A less onerous regulatory framework may also reduce Loon’s costs. Regulatory burdens tend to translate into higher project costs that could affect a project’s success.²⁴¹ Loon balloons fly in national airspace, for which the transnational regulatory framework is more stable because it is similar to the current regulatory framework for international civil aviation.²⁴² In contrast, given how outdated and underdeveloped international outer space law is, the regulatory framework for LEO satellites may be less stable in the future as

²³⁹ The Loon System includes the “most essential components of a cell tower and redesigned them to be light and durable enough to be carried by a balloon twenty km up, on the edge of space.” *The Loon System*, LOON, <https://loon.co/technology> [<https://perma.cc/RHR4-WFRA>]. In theory, the service should “reach more people for less money than it would take to install base stations or fiber optic cables in those areas.” Amy Nordrum, *How Project Loon Built the Navigation System That Kept Its Balloons Over Puerto Rico*, IEEE SPECTRUM (Mar. 8, 2018), <https://spectrum.ieee.org/tech-talk/telecom/Internet/how-project-loon-built-the-navigation-system-that-kept-its-balloons-over-puerto-rico> [[perma.cc link unavailable](#)].

²⁴⁰ JAKHU & PELTON, *supra* note 173, at 368.

²⁴¹ One author aptly notes “a problematic and persistent behavior in the space industry is thinking of engineering as the first step, and potential commercial markets and economic and regulatory consequences as a second step.” JAKHU & PELTON, *supra* note 173, at 370. This applies to space industry projects and tech solutions.

²⁴² Although the regulatory framework is more stable, this does not render the commercialization of Loon balloons straightforward. Loon will have to obtain permission from many states to fly over their territories—at least many more states than would be required to register a satellite.

greater international cooperation is required to address space innovation.²⁴³ The lack of international cooperation to develop an updated regulatory framework may translate to greater costs for satellite companies. Some other points of differentiation include difficulties securing spectrum allocated to satellites,²⁴⁴ as well as greater state involvement due to state responsibility for liabilities caused by space objects under international law.²⁴⁵

Ultimately the global last mile problem may require many solution providers and may not be such a “space race” after all. In fact, news reports signal that Loon and satellite companies are even collaborating with regards to the same challenge: the routing of data packets through constantly moving balloons and satellites.²⁴⁶ Moreover, although it remains unclear how much disruptive potential each project has at this stage, it *is* clear that the cost of access will be a key driver in generating demand in low-end markets. An important factor for adoption in developing countries will be whether Internet service is affordable, which partly depends on the sufficiency of backhaul infrastructure.

C. Leapfrogging Stages of Development?

Connecting users to broadband without having to build new terrestrial infrastructure would, potentially, overcome a significant obstacle to bringing Internet access to under-served markets worldwide. As such, many wonder whether wireless broadband infrastructure solutions such as balloons and mega-constellations will enable developing countries to “leapfrog” expensive

²⁴³ Many new regulations are needed at the national and international level to address challenges, such as the burst in ITU filings for spectrum priority; issues with orbital space debris, orbital congestion, and limited orbit control capabilities; shortage of radio frequencies; and potential interference to satellite systems. *See* JAKHU & PELTON, *supra* note 173, at 362–72. Cooperation in international forums is especially crucial with regards to these challenges as “the existing international space governance system is insufficient, inadequate, and inappropriate for facilitating the rapid introduction of small satellites as well as regulating their negative implications.” *Id.* at 371.

²⁴⁴ *See* JAKHU & PELTON, *supra* note 173, at 149 (discussing accommodating radio frequency spectrum allocations as a governance challenge, especially as “demand for terrestrial wireless broadband continues to expand sharply (i.e., nearly 40% per annum).”).

²⁴⁵ Because states bear international responsibility for activities in outer space, no matter if the activities are performed by governmental or commercial organizations, launching or registering states may ultimately be liable for any damage caused by their space objects under the various treaties. *See* Johnson, *supra* note 190, at 7–8, 10. As a result, greater costs from increased state involvement in the satellite industry are foreseeable given the potential for greater harm stemming from collisions, and thus liabilities.

²⁴⁶ *See* Salvatore Candido, *The Connectivity Brain Behind Loon’s Network*, MEDIUM (Jan. 31, 2019), <https://medium.com/loon-for-all/the-connectivity-brain-behind-loons-network-f26c2b0b4288> [<https://perma.cc/JRV2-CBAF>].

underground cables, just as fiber enabled them to leapfrog over cable and DSL.²⁴⁷ As many consider the wireless infrastructure projects' potential to accelerate telecommunications development in developing countries, the importance of backhaul network infrastructure has played a central role in the debate. This Section views skeptically the claim that wireless telecommunications innovations will help countries leapfrog stages of infrastructure development. Instead, it argues that wireless solutions complement, rather than compete with, wireline broadband and, in fact, rely on wired backhaul networks to decrease the cost of access and increase reliability.

Wireless solutions depend on spectrum, which is already quite limited and overcrowded, and these "spectrum limitations have a significant impact on the broadband delivery capabilities of a wireless service."²⁴⁸ Wireless technology is not considered a comprehensive solution but an "intermediary" one because it cannot serve large areas given the physical limits of radio spectrum; this limits the number of devices that can be connected at any given time.²⁴⁹ One key advantage of wireline broadband is that because the service does not rely on spectrum, greater speed and capacity are available. Moreover, fewer factors could weaken the signal, such as proximity to cell towers and weather.

Experts mostly agree on the advantages of wireline broadband, but still highlight that both wireless and wireline broadband services are needed and that one will not displace the other.²⁵⁰ They note that wireless services will remain complementary because wireless service depends on the speed and quality of wireline connections. For wireless networks, data traffic travels over the air for only a short distance and then requires a high-capacity wired connection.²⁵¹ For instance, Loon balloons rely on backhaul networks to

²⁴⁷ See INT'L TELECOMM. UNION, *supra* note 5, at 6.

²⁴⁸ See Larry Thompson et al., *Comparing Wired and Wireless Broadband*, BROADBAND COMMUNITIES, 86 (2015) (noting that spectrum scarcity "significantly constrains the amount of broadband that can be provided").

²⁴⁹ KURBALIJA, *supra* note 235, at 179.

²⁵⁰ See Thompson et al., *supra* note 248, at 92 (noting that because wireline technologies are not plagued by issues of scarce spectrum, they provide better service compared to wireless technologies in terms of speed, latency, capacity, and reliability). Whereas wireless services are mainly used to meeting customers' mobile needs, high-quality wireline services are necessary for activities such as videoconferencing and streaming. *See id.*

²⁵¹ *See id.* ("Wireless service depends on the speed and quality of wireline connections. Wireless towers require high-capacity connections, typically using Ethernet delivered over a landline carrier's fiber network."); OECD, FIXED BROADBAND NETWORKS, *supra* note 21, at 5 (noting that the growth of Wi-Fi and other mechanisms for offloading mobile traffic is expected to place greater demands on wired networks).

extend connectivity.²⁵² Without a nearby ground station with a backhaul connection, some might be unable to access Loon Internet. Until more advances are made in laser communications systems to address the shortcomings of radio waves,²⁵³ it is prudent to continue with both terrestrial and wireless Internet broadband networks.

Although leapfrogging stages of development could lower access prices, so would ample provision of backbone infrastructure. Internet access in developing countries is as much a last mile issue as it is a backbone issue. Backhaul networks—whether international connectivity via submarine cables or satellites, regional backbone, or domestic backbone—are key for carrying

²⁵² When data is sent wirelessly to a balloon, the floating cell tower then transmits the data wirelessly to a ground station that is connected to the wired backbone Internet. *See* Salvatore Candido, *1 Connection, 7 Balloons, 1,000 Kilometers*, MEDIUM (Sept. 11, 2018), <https://medium.com/loon-for-all/1-connection-7-balloons-1-000-kilometers-74da60b9e283> (“[A] backhaul connection must pass from a ground access point to a balloon—a big jump that ultimately has some constraints.”) [<https://perma.cc/GR62-8KSM>]; Taylor Hatmaker, *These Google X Moonshots Will Radically Change the World*, KERNEL MAG (Mar. 8, 2015), <https://kernelmag.dailydot.com/issue-sections/features-issue-sections/12083/google-x-project-loon-titan-makani/> (data “travel[ing] through the balloon network is ultimately relayed to [Google’s] local telecommunications partners’ ground stations, where it connects to pre-existing Internet infrastructure.”) [<https://perma.cc/E9C2-LQXP>]. In contrast, fixed cell towers transmit the data via fixed connectivity to the main network. *See* HALL ET AL., THE FUTURE OF NATIONAL INFRASTRUCTURE, *supra* note 10, at 185. Although Loon balloons are designed to expand coverage areas, there are still “people who live outside the reach of one of [the] balloons operating adjacent to a backhaul connection on the ground.” Candido, *supra* note 246.

²⁵³ Radio waves are a weak form of light and have many limitations. Optical communications systems seek to address the limitations of radio frequency communications. Although laser is already used in cables, experts are experimenting with using laser through air and space. NASA and other space agencies are studying laser communications systems, which can transfer data ten to one hundred times better than those of radio systems and maintain better signal strength across long distances. Many note that the future, of space communications at least, is optical. *See* Rebecca Boyle, *Space Communications Are Stuck In The Dial-Up Age. Which Means It’s Time For More Lasers.*, FIVETHIRTYEIGHT (Apr. 18, 2018, 9:41 AM), <https://fivethirtyeight.com/features/space-communications-are-stuck-in-the-dial-up-age-which-means-its-time-for-more-lasers/> [<https://perma.cc/9GQV-HAGH>]; Nat’l Aeronautics and Space Admin., *Benefits of Optical Communications*, NASA.GOV (May 6, 2014), https://www.nasa.gov/directorates/heo/scan/engineering/technology/txt_opticalcomm_benefits.html [<https://perma.cc/58VK-GDHL>]; Adam Hadhazy, *How It Works: NASA’s Experimental Laser Communication System*, POPULAR MECHANICS (Sept. 6, 2011), <https://www.popularmechanics.com/space/a7194/how-it-works-nasas-experimental-laser-communication-system/> [<https://perma.cc/WU6L-LRFD>].

traffic.²⁵⁴ One estimate is that fiber backhaul of a 4G LTE network constitutes seventy to eighty percent of a network's total cost.²⁵⁵ For the cost of access to wireless services to decrease and reliability to increase, more regional backhaul infrastructure is required.²⁵⁶ Regional connectivity depends on Internet Exchange Points (IXPs), physical infrastructure that allows local Internet Service Providers (ISPs) to exchange local traffic. Called the core of the Internet, IXPs contribute to better performance and lower costs by keeping Internet traffic local.²⁵⁷ The lack of IXPs in developing countries has meant that data is not routed locally, but as long-distance international traffic via the backbone networks of developed countries. This phenomenon is called boomerang or indirect routing, it increases costs, and decreases speed due to the multiple network hops required to route traffic.²⁵⁸ For example, intra-African data from one African Internet user to another is often passed outside of the region and is exchanged by ISPs in Europe or North America.²⁵⁹ The high costs of international traffic routing of intra-regional data result in higher costs of access. In countries where there has been increased regional backbone infrastructure development, Internet services became more affordable and reliable.²⁶⁰

Additionally, developing countries aiming to leapfrog stages of telecom development should consider whether they would be leapfrogging into greater dependence on foreign companies. Following the era of

²⁵⁴ See MARK D. J. WILLIAMS, REBECCA MAYER & MICHAEL MINGES, WORLD BANK, AFRICA'S ICT INFRASTRUCTURE: BUILDING ON THE MOBILE REVOLUTION, 84 (2011); KURBALIJA, *supra* note 235, at 178. More than ninety percent of all global Internet traffic flows through submarine cables. The Digital Silk Road is planning terrestrial cable investment, shifting traffic from the seabed to land. *Id.* at 38.

²⁵⁵ See OECD, FIXED BROADBAND NETWORKS, *supra* note 21, at 12.

²⁵⁶ See Mike Jensen, *Promoting the Use of Internet Exchange Points: A Guide to Policy, Management, and Technical Issues*, INTERNET SOCIETY REPORTS 5 (2009) (“[E]stablishing a local IXP decreases Internet access prices to the end user and provides faster response times from local web sites and other local interactive services”).

²⁵⁷ Carolyn D. Marsan, *Experts Say Economics and Politics Hamper Efficient Routing of Internet Data*, IETF JOURNAL (Nov. 1, 2014), <https://www.ietfjournal.org/experts-say-economics-and-politics-hamper-efficient-routing-of-internet-data/> [<https://perma.cc/JRV2-CBAF>].

²⁵⁸ KURBALIJA, *supra* note 235, at 178; Jensen, *supra* note 256, at 4 (noting telecommunication and management costs and impacts on speed depending on the number of network hops required to reach another local network).

²⁵⁹ WILLIAMS, MAYER, & MINGES, *supra* note 254, at 64.

²⁶⁰ Improvements in Internet performance and economics from IXP installations have been noted in Argentina, Brazil, Ecuador, and Kenya. In Ecuador, local traffic became \$1 per megabit per second, much cheaper than when international transit was \$100 per megabit per second. In Kenya, latency reductions went from 200 to 600 milliseconds down to a range of 2 to 10 milliseconds, and local mobile operators saved \$1.5 million per year on international transit. See Marsan, *supra* note 257.

liberalization of the telecommunications sector, many countries have generally welcomed foreign investment. However, some countries have continued to maintain foreign ownership restrictions in sensitive sectors like telecom, seeing the value in retaining a grip on infrastructure that is considered a critical input for economic development.²⁶¹ One related issue for developing countries is that dependence on foreign companies for technological solutions does not always translate into the development of tech expertise and stimulation of the local tech industry. This issue is only exacerbated in the satellite context: only about a dozen states possess launch capabilities, and the issue of foreign dependence is especially prominent for non-spacefaring nations who rely on satellite services.²⁶² In the context of Loon's first commercial deal with Kenya, local experts expressed concerns about "building a reliance on commercial, foreign technology for something as critical as connectivity" and "the potential for dependency on a single foreign company."²⁶³ Some propose additional backbone infrastructure in the form of IXPs to help countries become more self-sufficient with regards to technological development. Developing countries are already heavily burdened when financing access to backbones based in developed countries, and from a public policy standpoint, more IXPs would decrease foreign dependence.²⁶⁴ Although there are barriers to overcome, such as lack of tech expertise,²⁶⁵ IXPs would ensure large capital outflows are not paid to foreign ISPs.²⁶⁶

In sum, though the envisioned broadband infrastructure solutions

²⁶¹ See Margit Molnar, *Different Regulations, Different Impacts: What Regulations Affect Trade in Telecommunications Services?*, OECD Experts Meeting on Telecommunication Services 10 (2008), <https://fddocuments.us/document/different-regulations-different-impacts-what-.html> [<https://perma.cc/X3GJ-C7EQ>].

²⁶² See JAKHU & PELTON, *supra* note 173, at 366–67.

²⁶³ *Google's Loon Brings Internet-by-Balloon to Kenya*, BBC NEWS (July 19, 2018), <https://www.bbc.com/news/technology-44886803> ("Once these networks are in place, and dependency has reached a critical level, users are at the mercy of changes in business strategy, pricing, terms and conditions and so on.") [<https://perma.cc/SSK9-P229>]; Nathan Mattise, *Project Loon Signs its First Deal for Internet-delivering Balloons—in Kenya*, ARS TECHNICA (July 29, 2018), <https://arstechnica.com/gadgets/2018/07/project-loon-signs-its-first-deal-for-Internet-delivering-balloons-in-kenya/> [<https://perma.cc/YM9T-FPE4>]. See also Alex Davies, *Inside X, The Moonshot Factory Racing to Build the Next Google*, WIRED (July 11, 2018), <https://www.wired.com/story/alphabet-google-x-innovation-loon-wing-graduation/> ("Critics already call Google a monopoly. Now imagine its dominion extending . . . into how we connect to the Internet at all.") [<https://perma.cc/9BPU-9D5H>].

²⁶⁴ See KURBALIJA, *supra* note 235, at 174; Jensen, *supra* note 256, at 2.

²⁶⁵ See Jensen, *supra* note 256, at 3 ("The barriers to establishing IXPs in countries where they do not yet exist are largely non-financial. . . [L]imited technical skills and a lack of open competitive markets in telecommunication and Internet services make it more difficult to establish an IXP.")

²⁶⁶ See Jensen, *supra* note 256, at 2.

appear to be convenient alternatives to investing in wireline infrastructure, they do not address the issue of adequate access to backbone infrastructure and are at best intermediary or complementary. However, the popularity of national broadband plans and universal service obligations across countries points to a genuine demand for any infrastructure that may bridge the last mile by reaching poor and rural communities. These imperfect solutions may help some countries to at least provide provisional Internet access to rural areas that are decades away from being profitable for wireline investment. The technology might, after all, be disruptive in reaching unconnected communities years ahead of schedule, prompting demand for services and future wireline infrastructure development.

Nevertheless, the financial and social constraints of the developing world will be important factors driving adoption in underserved markets. Skeptics criticize companies such as Loon for assuming that demand for broadband Internet exists in developing countries.²⁶⁷ They explain that the people these companies want to provide with Internet service might not actually want it, emphasizing obstacles such as lack of disposable income and unmet basic needs. Questioning whether many of the world's unconnected people "want all this Internet," skeptics urge tech companies interested in connectivity to fully consider the needs of people in the developing world.²⁶⁸ As such, the financial and social constraints of the developing world are explored next, in the context of a theoretical framework that aims to assess the potential development impacts of broadband infrastructure: via the capability approach.

IV. EVALUATING THE POTENTIAL IMPACT OF BROADBAND INFRASTRUCTURE USING THE CAPABILITY APPROACH

Parts II and III above addressed some of the regulatory frameworks that companies interested in nontraditional broadband infrastructure must pay attention to—universal service policies and regulations concerning airspace, outer space, and spectrum—and evaluated the disruptive potential of high-altitude broadband infrastructure. However, a greater question is posed by the infrastructure projects described above: besides the promise of new markets and additional profits, what is the ethos or purpose behind the infrastructure provision? Moreover, if the connectivity projects purport to solve a need—

²⁶⁷ See Finley, *supra* note 175 ("We hear many times from satellite operators—especially those launching massive constellations—the pitch of 'connecting the other half of the population. . . . But the truth is that, of the total global unconnected population, two-thirds is not connected because they choose not to be connected.")

²⁶⁸ See *id.* (noting that what a remote village may want from the Internet is going to differ from what a person in Menlo Park wants).

namely the lack of worldwide connectivity and loss of associated benefits—how should the impact of these projects be measured?

As will be discussed below, socioeconomic development is typically the goal for most infrastructure provision. Infrastructure resources are not built for their own sake, but for a purpose; usually they are seen as development inputs that help generate productive downstream activities. This view has long been held by the World Bank, well versed in traditional infrastructure assets, and is articulated in the UN SDGs, which aim to strengthen social, economic, and environmental development.²⁶⁹ The executives at Alphabet and some satellite companies have also articulated a similar vision whereby they expect nontraditional connectivity infrastructure to benefit recipients socioeconomically. By extending Internet access to billions in rural and remotes areas of the world, those affiliated with Loon believe connectivity will lead to a number of positive developments, including access to online learning, better access to medical information and doctors, enhanced employment opportunities, and increased political participation.²⁷⁰ The goal of socioeconomic development also aligns with views articulated by the satellite company O3b, whose founders believe that ubiquitous connectivity will be an enabler “of industry productivity, economic growth, and social opportunity.”²⁷¹

Instead of assuming that development will be a natural byproduct of broadband infrastructure deployment globally, this article cautions against a “build it and they will come approach,” often employed naively by technology enthusiasts. Entering into emerging markets, private actors have not always engaged with development studies literature and may be unaware of the

²⁶⁹ UN, *Social Development for Sustainable Development*, <https://www.un.org/development/desa/dspd/2030agenda-sdgs.html> [<https://perma.cc/F29A-26G7>].

²⁷⁰ See SCHMIDT & COHEN, *supra* note 13, at 13 (where the chairman of Google and a former U.S. State Department official (who is now the president of Jigsaw (previously Google Ideas)) predict that five billion people will join the virtual world and that the “boom in digital connectivity will bring gains in productivity, health, education, quality of life and myriad other avenues in the physical world.”).

²⁷¹ Simon Gatty Saunt, *Accelerating a Pan-Africa Broadband Revolution*, SES, (Nov. 22, 2018), <https://www.ses.com/blog/accelerating-pan-africa-broadband-revolution> [<https://perma.cc/WNX3-RWJ9>]. O3B is now owned by “SES.”

“unique challenges of implementing technology” in developing countries.²⁷² There are limitations to technocratic approaches and solutions to development. Moreover, this article recommends employing Amartya Sen’s capability approach to anticipate and measure the potential development impact of newly available connectivity infrastructure. Sen’s capability approach sharpens the focus from macro- to micro-level considerations when implementing infrastructure solutions to anticipate how newly available technology may positively and negatively impact individuals’ lives and substantive freedom.

The first Section will evaluate the alleged direct relationship between infrastructure resources and socioeconomic development. The second Section briefly summarizes and applies the capability approach as an alternative to evaluating project outcomes in the aggregate and broadly in terms of “socioeconomic development.” The Section also addresses macro- and micro-level barriers that prevent the development of capabilities and examines the potential impact technology might have once introduced on existing and newly acquired capabilities. The second Section ends with an appraisal of community-led approaches to development to ensure that access to technology does not become harmfully disruptive—rather, it should respect the autonomy of individuals and communities.

A. Exploring the Relationship Between Infrastructure and Socioeconomic Development

Infrastructure resources are not built for their own sake, but for a purpose—usually, to enhance development. Definitions of infrastructure typically highlight the role such resources play in enabling development. One definition notes that infrastructure resources are “*shared means to many ends*.”²⁷³ As such, the value of the infrastructure resource is not generated primarily from consuming the resource for its own sake, but from

²⁷² See Jean-Yves Hamel, UNDP, *ICT4D and the Human Development and Capabilities Approach: The Potentials of Information and Communication Technology*, 57 (Human Development Research Paper, no. 37, 2010) (“A fundamental problem in the application of ICTs is apparently the domination of the field by technologists approaching the implementation of the tools and techniques in purely technological terms, with insufficient attention to local capacities and the diversity encountered in the field, which can make it or break it in developing countries.”); Mark Thompson, *ICT and Development Studies: Towards Development 2.0*, 3 (Working Paper Series 27, 2007).

²⁷³ FRISCHMANN, INFRASTRUCTURE, *supra* note 9, at 4; OECD, 2 INFRASTRUCTURE TO 2030: MAPPING POLICY FOR ELECTRICITY, WATER AND TRANSPORT, 20 (2007) (“Infrastructures are not an end in themselves. Rather, they are a means for ensuring the delivery of goods and services that promote prosperity and growth and contribute to quality of life, including the social well-being, health and safety of citizens, and the quality of their environments.”).

“downstream productive activity that requires the resource as an input.”²⁷⁴ Even if there is a direct, immediate benefit from consuming an infrastructure resource, the value for most lies in their “intermediate production capacity.”²⁷⁵

More specifically, telecommunications infrastructure is widely recognized as a vital development input.²⁷⁶ Scholars note that telecommunications benefits go beyond connecting people—calling the infrastructure a “link in the chain of the development process itself” and as important as water and electricity.²⁷⁷ Besides the term “telecommunications” infrastructure, scholars use interchangeably terms such as broadband or ICT infrastructure, the definition of which is commonly believed to encompass the Internet.²⁷⁸ Many scholars use words like essential, basic, or fundamental when describing the Internet as an infrastructure resource,²⁷⁹ and underscore that it is a “public and social infrastructure that is transforming our society.”²⁸⁰ Aside from the commercial activities it enables, the Internet is believed to be socially valuable because of the downstream activities it facilitates, for example via platforms which enable the exchanging of ideas and goods and social interactions.²⁸¹

As “shared means to many ends,” it is important to ask what the end goals of broadband infrastructure projects are. Scholars regularly emphasize the economic and social benefits of infrastructure and its role in economic and

²⁷⁴ See Frischmann, *Economic Theory of Infrastructure*, *supra* note 18, at 956–58 (explaining why a road system may provide directly realizable consumptive benefits, but the social benefits accrue from “activities it facilitates at the ends, including, for example, commerce, labor, communications, and recreation”).

²⁷⁵ See HALL ET AL., THE FUTURE OF NATIONAL INFRASTRUCTURE, *supra* note 10, at 32.

²⁷⁶ Jayakar & Liub, *supra* note 34, at 186.

²⁷⁷ Frieden, *supra* note 99, at 453–54.

²⁷⁸ See Hamel, *supra* note 272, at 1 (The World Bank defines ICTs as “tools or techniques that allow recording, storing, using, diffusing and accessing electronic information,” and others say ICTs “facilitate communication and the processing and transmission of information and the sharing of knowledge by electronic means”).

²⁷⁹ WILLIAM H. LEHR & LORENZO MARIA PUPILLO, INTERNET POLICY AND ECONOMICS: CHALLENGES AND PERSPECTIVES 3 (2d ed. 2009) (“The Internet is now widely regarded as *essential* infrastructure for our global economy and society.”); *id.* at 6 (“As *basic* infrastructure, the Internet supports the production and consumption of both market and nonmarket goods.”); *see also* OECD, *Internet Governance*, <http://www.oecd.org/Internet/Internet-governance.htm> (“The Internet is a *fundamental* infrastructure with a still largely untapped potential to address a wide array of economic and social challenges.”) [<https://perma.cc/L683-P645>].

²⁸⁰ PRESIDENT’S INFO. TECH. ADVISORY COMM., INFORMATION TECHNOLOGY RESEARCH: INVESTING IN OUR FUTURE 11–20 (1999), http://www.itrd.gov/pitac/report/pitac_report.pdf [<https://perma.cc/558L-ML6U>].

²⁸¹ FRISCHMANN, INFRASTRUCTURE, *supra* note 9, at 217. *See* Frischmann, *Economic Theory of Infrastructure*, *supra* note 18, at 957 n.43 (explaining how platforms are enabling technologies that other firms use to build more innovative products).

social development.²⁸² Along that vein, scholars increasingly perceive broadband as an essential service to foster economic growth and social development.²⁸³ This view is shared by traditional development actors, such as the World Bank, the OECD, and the UN, but also by Alphabet executives, who believe in the power of connectivity to generate socioeconomic progress.²⁸⁴ However, even though the term is invoked by many, socioeconomic development is not clearly defined, and the concept lacks a general framework, much less an agreed-upon methodology for measuring impact.²⁸⁵ The term development usually alludes to macro-level progress, change, or growth, and the modifier “socioeconomic” relates development to social and economic factors.²⁸⁶ To better understand the aims of broadband infrastructure provision, the two factors are briefly examined separately.

One goal of infrastructure investment is economic development,

²⁸² See FRISCHMANN, *INFRASTRUCTURE*, *supra* note 9, at 11 (infrastructure resources are essentially “intermediate capital resources that serve as critical foundations for productive behavior within *economic and social systems*.”); OECD, *INFRASTRUCTURE TO 2030*, *supra* note 273, at 20 (“In the past, infrastructures have provided significant social and economic benefits. Looking to the future, they will continue to play a vital role in *economic and social development*. . .”); John M. Anderies, et al., *Institutions and the Performance Of Coupled Infrastructure Systems*, 10 INT’L J. OF THE COMMONS 502 (2016) (studying how infrastructure “functions in structuring *social and economic* processes”).

²⁸³ See HOLZNAGEL ET AL., *supra* note 1, at 15 (discussing “socioeconomic impacts of broadband”); LEHR & PUPILLO, *supra* note 279, at 3 (noting the Internet is “essential infrastructure for our global *economy and society*”); Gerli et al., *supra* note 32, at 726 (“Superfast broadband is increasingly perceived as an essential service to foster *economic growth and social development*”); Harvard Berkman Center for Internet and Society, *Next Generation Connectivity: A Review Of Broadband Internet Transitions And Policy From Around The World*, 129 (2010) (“Broadband access is necessary to participate in the 21st-century economy. It’s also good policy as well: Broadband boosts *social* opportunity and *economic growth*.”); *id.* at 242 (“The wide *social and economic* impact of broadband has given all levels of government an interest in the quality and price of services.”).

²⁸⁴ See SCHMIDT & COHEN, *supra* note 13, note 270; WORLD BANK, *SOCIO-ECONOMIC ASSESSMENT OF BROADBAND DEVELOPMENT IN EGYPT*, (2010), <https://openknowledge.worldbank.org/handle/10986/12690> [<https://perma.cc/49BZ-XE5F>]; OECD, *BROADBAND POLICIES FOR LATIN AMERICA AND THE CARIBBEAN: A DIGITAL ECONOMY TOOLKIT*, <http://www.oecd.org/Internet/broadband/lac-digital-toolkit/Home/toolkit-text-chapter1.htm> [<https://perma.cc/2ZXB-TPQ3>].

²⁸⁵ Narcyz Roztocki & H. Roland Weistroffer, *Conceptualizing and Researching the Adoption of ICT and the Impact on Socioeconomic Development*, 22 INFO. TECH. FOR DEV., 541–42 (2016).

²⁸⁶ *Id.*

considered tantamount to economic growth and productivity.²⁸⁷ On a macro-level, growth related to infrastructure is measured by increases in national income or GDP, and on an individual level, by increases in personal income.²⁸⁸ Because infrastructure resources are often inputs in productive downstream activity, many believe infrastructure and economic development are at least positively correlated.²⁸⁹ The same holds for broadband infrastructure: the recent move toward national broadband plans is because policymakers and economists judge that access to broadband provides economic growth opportunities.²⁹⁰ The World Bank, with many years of experience in traditional infrastructure, deems broadband infrastructure as vital to producing economic development and has found that increased broadband penetration in low and middle-income countries results in economic growth.²⁹¹ One recent World Bank report noted that “a 10 percent increase in broadband penetration yields an additional 1.38 percent increase in GDP growth for low to middle-income countries.”²⁹² Despite the apparent strength of the claim, the causal link between infrastructure development and economic growth remains

²⁸⁷ See FRISCHMANN, *INFRASTRUCTURE*, *supra* note 9, at 19–20 (“Whether it is the Internet or freeways, infrastructure improves the functioning of an economy. Road building and improvements in telecommunications infrastructure have both been found to have a significant impact on productivity and growth for a wide selection of OECD countries. . . .”); OECD, *INFRASTRUCTURE TO 2030*, *supra* note 273, at 20 (“[Infrastructures] are a means for ensuring the delivery of goods and services that promote *prosperity and growth*. . .”).

²⁸⁸ See Roztocki & Weistroffer, *supra* note 285, at 542.

²⁸⁹ See HALL ET AL., *THE FUTURE OF NATIONAL INFRASTRUCTURE*, *supra* note 10, at 7 (the relationship is symbiotic).

²⁹⁰ See OECD, *National Broadband Plans*, 6 (OECD Digital Economy Papers No. 181 June 2011). See also *id.* at 10 (“Some economists have found that investment in broadband Internet access directly correlates to growth in GDP and gains in productivity . . . a ‘consensus’ view was that a 10% increase in household penetration of broadband boosted GDP by 0.1% to 1.3%.”); HOLZNAGEL ET AL., *supra* note 1, at 21 (high-speed broadband networks are correlated with GDP, higher employment rates, and increased productivity).

²⁹¹ See Christine Zhen-Wei Qiang et al., *Economic Impacts of Broadband*, in *INFORMATION AND COMMUNICATION FOR DEVELOPMENT: EXTENDING REACH AND INCREASING IMPACT* 45 (World Bank ed., 2009), https://siteresources.worldbank.org/EXTIC4D/Resources/IC4D_Broadband_35_50.pdf [<https://perma.cc/79HW-KP6D>].

²⁹² *Willingness to Pay for Broadband in Nigeria*, *supra* note 63, at 421; Qiang et al., *supra* note 291, at 45. This statistic was cited by a Loon engineer, who also explained that Internet access and associated GDP growth could double the standard of living in many countries around the world. Verge, *Inside Google's Wildly Ambitious Internet Balloon Project*, YOUTUBE, <https://www.youtube.com/watch?v=OFGW2sZsUiQ> [<https://perma.cc/WZD5-7QMQ>].

subject to much debate.²⁹³

Another goal of infrastructure investment is social development, but unlike economic development, it is difficult to quantify or measure. To grasp the social value of infrastructure, some suggest examining downstream uses and interactions to see if they benefit society as a whole.²⁹⁴ For instance, telecommunications infrastructure benefits society and individuals by “facilitate[ing] all sorts of personal, social, and business interactions as people talk, exchange ideas, make plans, and socialize; develop[ing] business and personal relationships; participat[ing] in political discourse; and so on.”²⁹⁵ Moreover, some downstream uses and interactions which telecom infrastructure facilitates are also known to generate economic growth.²⁹⁶ Nevertheless, though many policy documents and academic articles use terms like improved *well-being*, *quality of life*, and *social welfare* to measure the social impacts of infrastructure, such terms do not always answer the question of why certain activities are “socially valuable,” nor do they address the unintended social consequences of newly available infrastructure.²⁹⁷

²⁹³ See HALL ET AL., THE FUTURE OF NATIONAL INFRASTRUCTURE, *supra* note 10, at 8, 31 (citing many scholars for the idea that the link between infrastructure availability, economic growth and productivity is uncertain because the “relationships between infrastructure and the economy are multiple and complex”); BJÖRN-SÖREN GIGLER, WORLD BANK GROUP, DEVELOPMENT AS FREEDOM IN A DIGITAL AGE: EXPERIENCES OF THE RURAL POOR IN BOLIVIA 4–6 (2015) (noting that proponents of the ICT for Development (ICT4D) agenda assume a direct relationship between ICTs and economic growth and social development, but that the claim is not sufficiently supported empirically). Some concede that an output of infrastructure investment—namely, economic growth—is not as clear with telecommunications as with roads, for example, but highlight that there are still notable increases in productivity, positive effects on business activities and transfer of knowledge, as well as indirect effects on empowering the workforce. See Emiliani, *supra* note 25, at 1–2.

²⁹⁴ See Frischmann, *Economic Theory of Infrastructure*, *supra* note 18, at 1016–17 (listing social activities the Internet enables: “[individuals] engage in innovation and creation; they speak about anything and everything; they maintain family connections and friendships; they debate, comment, and engage in political and nonpolitical discourse; they meet new people; they search, research, learn, and educate; and they build and sustain communities. These are the types of productive activities that generate *substantial social value*. . .”); *id.* at 1017–18 (“Public participation in such activities results in external benefits that accrue to society as a whole (online and offline) . . .”).

²⁹⁵ FRISCHMANN, INFRASTRUCTURE, *supra* note 9, at 217.

²⁹⁶ For instance, economic theory posits that technological advancements and social networks have economic outcomes as they “enable creation of trust relationships that in turn facilitate business between various agents” beyond the family. Mdoe & Kinyanjui, *supra* note 45, at 3–4.

²⁹⁷ See FRISCHMANN, INFRASTRUCTURE, *supra* note 9, at 23 (measures of social welfare include “quality of life or living standards”); OECD, INFRASTRUCTURE TO 2030, *supra* note 273, at 20 (Infrastructure “contribute[s] to quality of life, including the social well-being, health and safety of citizens, and the quality of their environments.”).

Instead, what constitutes development broadly, and what is socially valuable specifically, should be assessed using Amartya Sen's capability approach to evaluate whether an infrastructure resource enables opportunities that an individual or a community has reason to value.

B. Applying the Capability Approach to Broadband Infrastructure Projects

The Section above explored one commonly recognized end goal for infrastructure investment—socioeconomic development—and reminded that companies such as Loon and O3B show interest in enhancing socioeconomic development with global connectivity projects. Given that public and private actors believe that ICT infrastructure and connectivity will lead to socioeconomic development for the billions without Internet access, the next step is to consider how such connectivity might actually contribute to development. Aspiring to reorient approaches to development away from focusing solely on economic growth and measuring well-being by resources or utility, Sen's capability approach instead offers a different aspiration for development: to improve the "ability of persons to lead a life that they have reason to value."²⁹⁸ The capability approach (sometimes called the human development approach²⁹⁹) is typically applied to development projects in developing nations, yet its use does not have to be restricted in scope.³⁰⁰ In essence, the capability approach views development as freedom: the expansion of freedom is viewed as both (1) the primary end and (2) the principal means of development.³⁰¹ Social benefits derive from a person's

²⁹⁸ See Alexandre Apsan Frediani et al., *Approaching Development Projects from a Human Development and Capability Perspective*, 15 J. HUM. DEV. AND CAPABILITIES, 1 (2014) (The capability approach undertakes to define well-being as going "beyond economic conditions."); Sabina Alkire, *The Capability Approach and Well-Being Measurement for Public Policy*, 1 (OPHI Working Paper No. 94, 2015), <https://www.ophi.org.uk/wp-content/uploads/OPHIWP094.pdf> [<https://perma.cc/9BNJ-7JEH>].

²⁹⁹ Another term used interchangeably with the capability approach is "Human Development," historically associated with an office of the United Nations Development Program (UNDP) that produces annual Human Development Reports. See MARTHA C. NUSSBAUM, *CREATING CAPABILITIES: THE HUMAN DEVELOPMENT APPROACH* 17 (2013). The reports have introduced a people-centered approach for advancing human well-being, centered on "expanding the richness of human life, rather than simply the richness of the economy in which human beings live." *About Human Development*, UNITED NATIONS DEVELOPMENT PROGRAMME, <http://hdr.undp.org/en/humandev> [<https://perma.cc/FEJ6-YJEB>]. The Human Development Reports continue to recognize that ICTs, such as mobile phones and Internet use, are tools for expanding people's freedom. See Hamel, *supra* note 272, at 1.

³⁰⁰ See Mark Coeckelbergh, *Human Development or Human Enhancement? A Methodological Reflection on Capabilities and the Evaluation of Information Technologies*, 13 ETHICS INFO. TECH. 82 (2011).

³⁰¹ AMARTYA SEN, *DEVELOPMENT AS FREEDOM* 36 (1999).

freedom to act and ability to pursue what he or she believes is valuable.³⁰² Though the aim of infrastructure investment is typically socioeconomic development, this Section proposes an alternative vision: the end goal of broadband infrastructure development should be to enhance individual and collective “capabilities,” defined below.

First, the content below provides an explanation of the capability approach, examining key concepts like “capabilities” encompassed by this alternative view of development. Next, the Section examines how ICTs, such as mobile phones and the Internet, are used to improve development outcomes and deliver services in global development projects—a field known as ICT for Development, or ICT4D.³⁰³ In light of the ICT4D studies’ assessments of the actual impacts of ICT infrastructure on human development, the Section casts doubt on the ability of mere infrastructure availability to meaningfully expand capabilities and argues for a comprehensive meaning of access. Examining how ICTs impact the development of capabilities, the Section also briefly assesses how ICTs can work for and against capabilities. The Section ends with an appraisal of community-led approaches to development to ensure that access to technology does not become harmfully disruptive, but rather respects the autonomy of individuals and communities.

1. *Key Concepts*

Sen defines development as “a process of expanding the real freedoms that people enjoy” and emphasizes the need for the “expansion of ‘capabilities’ of persons to lead the kinds of lives they value—and have reason to value.”³⁰⁴ A person’s “capability” refers to the alternative combinations or set of functionings that are feasible for her or him to achieve.³⁰⁵ Functionings reveal the various things a person may value doing or being.³⁰⁶ For example, elementary functionings might include “being safe, well-nourished, and literate,” while more complex functionings might include campaigning for political office.³⁰⁷ Realized functionings are a way of describing what a person

³⁰² See Travis Godwin Good et al., *Investigating Capabilities Associated with ICT Access and Use in Latino Micro-enterprises*, 500 AMCIS PROC. (2010).

³⁰³ ICTs are pervasively utilized in global development projects to enhance outcomes. ICT4D initiatives are in response to the growing digital divide and focus on providing those living on less than \$2 a day with access to current technologies. See ITU, *HARNESSING THE INTERNET OF THINGS FOR GLOBAL DEVELOPMENT* 7 (2016); Henry Tinashe Manara, *Factors Affecting Sustainability of ICT4D: A Case Study of Mobile-Cinemas in Rural South Africa* 10 (2015) (unpublished MCom dissertation, University of Pretoria) (on file with author).

³⁰⁴ SEN, *supra* note 301, at 3, 18.

³⁰⁵ *Id.* at 75.

³⁰⁶ *Id.*

³⁰⁷ See Alkire, *supra* note 298, at 3.

actually achieves when deliberately using his or her capabilities, and potential functionings (synonymous with capabilities) represent possible choices available to individuals.³⁰⁸ These choices are circumscribed by personal, social, and environmental conversion factors (see Figure 1 below).³⁰⁹

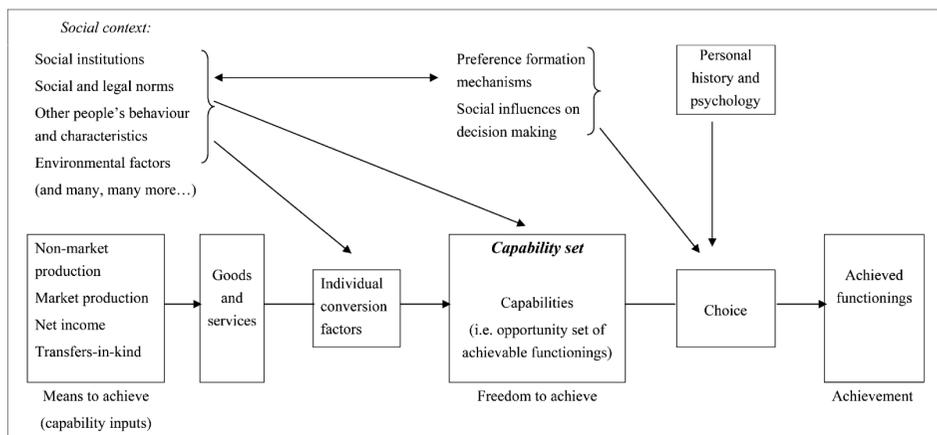


Figure 1: *A person's capability set and his/her social and personal context*³¹⁰

Emphasizing capabilities expansion, the capability approach is about creating real opportunities for people to access and realize functionings that they have reason to value.³¹¹ A person converts capabilities into functionings through agency and choice,³¹² and thus the capability approach emphasizes access to valuable functionings without insisting on the achievement of certain functionings.³¹³ In essence, the approach respects individuals' use of agency so that people are empowered to help themselves and become "actors of their own development."³¹⁴ Capability is akin to freedom; it is "the substantive

³⁰⁸ Good et al., *supra* note 302, at 3.

³⁰⁹ See *id.* at 3; Ingrid Robeyns, *The Capability Approach: A Theoretical Survey*, 6 J. HUM. DEV. 98–99 (2005); GIGLER, *supra* note 293, at 15 (noting valued functionings depend on capabilities and also on existing livelihood resources or assets, and the key is to analyze what people are capable of being or doing with available resources).

³¹⁰ See Robeyns, *supra* note 309, at 98.

³¹¹ See Alkire, *supra* note 298, at 4. Some capability expansion examples for an electrification project include promotion of connections with other people and the possibilities of studying in the evening. See Frediani et al., *supra* note 298, at 2.

³¹² See Helena Grunfeld et al., *Challenges in Operationalising the Capability Approach for Evaluating the Contribution of the Cambodian ICT4D Project, iREACH, to Capabilities, Empowerment and Sustainability*, HUMAN DEV. & CAPABILITY ASSN., 3 (2010).

³¹³ See Frediani et al., *supra* note 298, at 4.

³¹⁴ JOSEPH E. STIGLITZ, AMARTYA SEN & JEAN-PAUL FITOUSSI, REPORT BY THE COMMISSION ON THE MEASUREMENT OF ECONOMIC PERFORMANCE AND SOCIAL PROGRESS 151 (2009).

freedom to achieve combinations of alternative potential functionings.”³¹⁵ The evaluation of well-being is not achieved functionings, or visible states of being. An illustrative example is starving versus fasting: people in both situations do not exercise the functioning of nutrition. However, in the fasting scenario, one has chosen the valued state and still retains the capability of nutrition, while the person who is starving lacks agency.³¹⁶ The better evaluative space is thus capabilities, since operationalizing the approach is quite difficult without an agreed upon list of basic capabilities.³¹⁷

2. *From Mere Availability to Genuine Access*

By bringing to market Internet access in countries lacking adequate broadband infrastructure, tech companies are seemingly expanding capability sets for individuals. Indeed, the inadequate broadband infrastructure in developing countries is viewed as “an external macro-level barrier impeding the development of capabilities and their conversion into functionings at the micro-level.”³¹⁸ Some definitions of infrastructure capture the inherent expansion of opportunities: “[a]n infrastructure service is the *provision of an option for an activity* by operating physical facilities and accompanying human systems to convert, store and transmit resources (physical and virtual).”³¹⁹ However, infrastructure provision on its own will not translate into the expansion of capabilities for all; individuals with high capabilities in the first place might quickly capitalize on advantages of newly available technology and infrastructure. Therefore, interventions may be necessary to enhance the capabilities of individuals with already limited capability sets.

Many ICT4D studies find that access to broadband infrastructure or ICTs does not ensure expansion of individual capabilities.³²⁰ By itself, access

³¹⁵ SEN, *supra* note 301, at 75.

³¹⁶ See Nicholas Garnham, *Amartya Sen's Capabilities Approach to the Evaluation of Welfare: Its Application to Communications*, 4 JAVNOST - THE PUBLIC, at 29 (1997).

³¹⁷ See Grunfeld et al., *supra* note 312, at 4 (“[A] main difficulty of applying [the capability approach] is its lack of operationalisation . . . Listing basic capabilities could be one way of operationalising the CA, thereby making it more useful to development policy and Nussbaum’s . . . tentative list of basic capabilities is one example of this.”).

³¹⁸ See Grunfeld et al., *supra* note 312, at 10.

³¹⁹ See HALL ET AL., *THE FUTURE OF NATIONAL INFRASTRUCTURE*, *supra* note 10, at 6.

³²⁰ Many ICT4D initiatives reporting on actual impacts of ICT infrastructure on human development and “lessons learned” consistently convey a clear message: mere access to ICT infrastructure, whether it is computers, mobile phones, and/or the Internet, is not enough to close the digital divide and ensure development outcomes. See Grunfeld et al., *supra* note 312, at 4 (“The concept of access must include capabilities, e.g. (computer) literacy, to use the infrastructure, similar to the term ‘effective use’ . . . to reflect ‘the capacity and opportunity to successfully integrate ICTs into the accomplishment of self or collaboratively identified goals.’”).

to infrastructure is “not enough to enjoy, for instance, exercising one’s capability for affiliation. Instead, what matters is if the person can actually and effectively use the technology.”³²¹ From a capabilities perspective, physical access is a necessary but insufficient condition to expand capabilities; it is only meaningful if it translates into “effective use.”³²² This finding is consistent with the message of universal service scholars, who advocate for the reconceptualization of access to include five components: availability, affordability, accessibility, a needs assessment, and awareness.³²³ The five components should be more or less present to achieve the goals of universal service, which depend on actual adoption of ICT-based services. Under this view, the digital divide is not a gap between those who have Internet access and those who do not, but rather those who have the skills to use the Internet effectively and those who do not.³²⁴

Capabilities expansion after the introduction of new infrastructure depends on a number of factors. In addition to physical access to ICTs, effective use of ICTs depends on factors like cost, reliability, cultural acceptability, ease of use, and the relevance of available content.³²⁵ Some capability approach scholars have addressed the issue using different vocabulary, namely the term “conversion factors.” They note that resources are converted into functionings differently; differences in personal and socioeconomic factors, as well as the social, institutional and environmental context, affect one’s capabilities and functionings.³²⁶ Other scholars use the

³²¹ See Coeckelbergh, *supra* note 300, at 84.

³²² See Grunfeld et al., *supra* note 312, at 4; see also GIGLER, *supra* note 293, at 211 (“[T]he findings point to the absence of any statistically significant relationship between these two variables, suggesting that the provision of ICT infrastructure services in rural areas does not have its intended effect and de facto fails to enable people to make meaningful use of the Internet.”). Infrastructure provision barely checks the readiness and availability boxes of the ICT development cycle; the next phases interpreting “effective use” are uptake and impact. *Readiness* concerns policies and infrastructure to make ICT availability possible; *Availability* asks how to roll out the program, *Uptake* is about making ICTs useful, and *Impact* looks to maximize ICTs for the greatest development impact. See Richard Heeks, *The ICT4D 2.0 Manifesto: Where Next for ICTs and International Development?*, 28–29 (Manchester Centre for Dev. Informatics Working Paper 42, 2009).

³²³ *Universal Service Funds in Africa*, *supra* note 31, at 619-20 (“It is important to first ascertain the ‘need’ of a given community in terms of whether availability, accessibility, affordability, the possession of the right skills to use the available technology or a combination of these elements are the problem . . . [Moreover,] lack of awareness may undermine the rate of adoption.”).

³²⁴ See Coeckelbergh, *supra* note 300, at 85.

³²⁵ See Hamel, *supra* note 272, at 35 (citing TIM UNWIN, *ICT4D: INFORMATION AND COMMUNICATION TECHNOLOGY FOR DEVELOPMENT* (2009)).

³²⁶ See Frediani et al., *supra* note 298, at 4; Alkire, *supra* note 298, at 6 (“physical, social and cultural contexts affect [one’s] ability to convert resources into capabilities.”). Nussbaum describes the phenomenon in terms of combined and internal capabilities. A person’s *internal*

term “barriers” to describe certain social, institutional, and environmental contexts that may impede capabilities development and the conversion of capabilities into functionings.³²⁷ Some of these micro- and macro-level barriers are examined below.

Several micro-level factors may impede the development of capabilities or their conversion into functionings in light of broadband infrastructure availability. An ICT4D study found that Internet non-users provided several reasons for non-use despite having access to public telecenter hubs with satellite Internet, including “lack of awareness, lack of time due to . . . work, home duties, and schoolwork, lack of interest as ICT not considered important, too old, fear of damaging computers, insufficient literacy levels, and living too far away from a hub.”³²⁸ Two common micro-level factors, lack of financial resources and digital literacy, are addressed below.

An individual who lacks financial resources may be prevented from learning how to use newly available ICT infrastructure, and in the case she develops the capability, may be prevented from actual use by prohibitive cost. Studies show that in developing countries, pricing ICT services is a very important aspect of demand analysis and is the “strongest determinant of a broadband subscription.”³²⁹ For many, access to a mobile device (with Internet access) may already constitute a large portion of one’s total income, making consistent Internet usage a luxury. In some developing countries, mobile broadband services cost about 18.8 percent of monthly gross national income per capita for a gigabyte of data and fixed services cost 30.1 percent.³³⁰ This

capabilities refer to dynamic personal characteristics, while her *combined* capabilities refer to “the totality of the opportunities she has for choice and action in her specific political, social, and economic situation.” See NUSSBAUM, *supra* note 299, at 21 (Internal capabilities include both innate, as well as trained/developed traits. For example, “personality traits, intellectual and emotional capacities, states of bodily fitness and health, internalized learning, skills of perception and movement”).

³²⁷ See Grunfeld et al., *supra* note 312, at 10.

³²⁸ See *id.* at 12.

³²⁹ *Willingness to Pay for Broadband in Nigeria*, *supra* note 63, at 423. See also Hamel, *supra* note 272, at 49 (noting “higher telecommunications costs inhibits Internet use. . . Sadly, the relationship between costs and usage is most apparent in the poorest countries, where costs are exorbitant and usage rates are lowest.”); Grunfeld et al., *supra* note 312, at 11 (“The strong competition in the mobile and ISP markets, with some 37 ISPs, 10 of which were major . . . had not resulted in affordable prices by international standards.”); *Willingness to Pay for Broadband in Nigeria*, *supra* note 63, at 422, 430 (noting that the reason for weak demand for both fixed and mobile broadband in Nigeria is due to socioeconomic and demographic characteristics affecting people’s willingness to pay for broadband, such as income and level of education).

³³⁰ *Willingness to Pay for Broadband in Nigeria*, *supra* note 63, at 424. In Africa, 1GB of data currently costs an average of eighteen percent of monthly income. See generally ALLIANCE FOR AFFORDABLE INTERNET, AFRICA REGIONAL SNAPSHOT: 2017 AFFORDABILITY REPORT (2017).

is exorbitant in light of the UN Broadband Commission's target, which says that entry-level broadband should cost no more than two percent of gross national income per capita for a gigabyte of data.³³¹ Capability sets may develop or functionings be realized more quickly if access is affordable.

Another micro-level factor that may impede widespread Internet adoption is lack of digital literacy. In developing countries, computer and Internet illiteracy is frequently cited as a factor that contributes to poor Internet penetration and uptake.³³² One study of a telecenter in Vietnam highlights that many rural users were not aware of the benefits of Internet access.³³³ Nevertheless, improvement in digital literacy can significantly affect broadband demand and uptake.³³⁴ In response, many governments are combating low network utilization by moving beyond infrastructure provision and instead are focusing on improving citizens' digital literacy skills.³³⁵ In countries such as Kenya and South Africa, universal service policies include the promotion of digital literacy as a target.³³⁶ Government interventions include initiatives to increase digital literacy in schools as well as community sensitization programs to generate awareness.³³⁷ Beyond possessing digital skills to use ICTs, marginalized communities may not use an unfamiliar Internet which "does not speak their local language."³³⁸

Macro-level social, institutional, and environmental barriers may also affect the development of capabilities and their conversion into functionings in the case of broadband infrastructure provision. Resistance to the Internet

³³¹ Press Release, Int'l Telecomm. Union, UN Broadband Commission Sets Global Broadband Targets to Bring Online the World's 3.8 Billion Not Connected to the Internet (Jan. 23, 2018), <https://www.itu.int/en/mediacentre/Pages/2018-PR01.aspx> [<https://perma.cc/7VA3-KTGY>].

³³² See NYABUGA & BOOKER, *supra* note 45, at 17; Dorothea Kleine, *The Capability Approach and the "Medium of Choice": Steps Towards Conceptualising Information and Communication Technologies for Development*, 13 ETHICS INFO. TECH. 126 (2011) ("[A] certain amount of educational resources (i.e. literacy, IT skills) is needed, as well as health and psychological resources, to make use of the Internet."); Grunfeld et al., *supra* note 312, at 3 ("For example, having access to and knowing how to use ICT represent capabilities, and converting these capabilities, to send an e-mail would be a functioning."); *id.* at 4 ("The concept of access must include capabilities, e.g. (computer) literacy, to use the infrastructure. . . .").

³³³ Many users would visit the center to play games. See Thai & Falch, *supra* note 40, at 328. In Mexico, the Connected program has focused on creating community Internet access points, but little information is available whether the personnel are adequately equipped to help develop digital skills to foster adoption of ICTs. See Casanueva-Reguart, *supra* note 38, at 2111–12.

³³⁴ *Willingness to Pay for Broadband in Nigeria*, *supra* note 63, at 423–24.

³³⁵ See Thai & Falch, *supra* note 40, at 324; NYABUGA & BOOKER, *supra* note 45, at 17.

³³⁶ *Universal Service Funds in Africa*, *supra* note 31, at 627.

³³⁷ *Willingness to Pay for Broadband in Nigeria*, *supra* note 63, at 431.

³³⁸ See GIGLER, *supra* note 293, at xxxii (noting how ICTs are mostly designed for affluent people).

and smartphones in some rural areas indicates a social and cultural environment that is hostile to ICT uptake, as technology is perceived as a threat to the traditional ordering of society and beliefs.³³⁹ Such resistance disproportionately affects women. Studies show that in certain areas of the globe, there is a significant gender gap in ICT ownership and usage. In India, men on average are more than twice as likely to own a phone, thirty-six percent more likely to use a bare-bones mobile phone, and sixty-two percent more likely to use the Internet.³⁴⁰ A study reports that nearly twenty-five percent fewer women than men had access to the Internet across developing countries, and in sub-Saharan Africa, the gender gap was forty-three percent.³⁴¹ The same study revealed twenty percent of the Indian women interviewed believed that Internet use was inappropriate for them.³⁴²

Another macro-level barrier is inadequate access to basic infrastructure services. A study in Bolivia found that low access to basic infrastructure services, like water and electricity, was related to low ICT usage and capabilities.³⁴³ Many have even called the digital divide an electricity divide.³⁴⁴ About eighty-five percent of the estimated 1.5 billion people without access to electricity live in rural areas of developing countries.³⁴⁵ For instance, according to Kenya's Rural Electrification Authority, ninety percent of rural Kenya does not have electricity, and only twenty percent of Kenyans generally have access to electricity.³⁴⁶ This impacts capabilities because individuals in rural areas cannot utilize ICTs absent adequate power sources. Lack of ICT uptake persists due to inadequate electricity despite universal service

³³⁹ See Ellen Barry, *No, Google's Not a Bird: Bringing the Internet to Rural India*, N.Y. TIMES (May 21, 2017), <https://www.nytimes.com/2017/05/21/world/asia/internet-in-india.html> [<https://perma.cc/S8UA-7H3T>].

³⁴⁰ See *id.*; see also GSMA, *Women and Mobile in India: Realising the Opportunity*, GSMA.COM (Oct. 24, 2016), <https://www.gsma.com/mobilefordevelopment/programme/connected-women/women-and-mobile-in-india-realising-the-opportunity> [<https://perma.cc/9X4F-YD6H>].

³⁴¹ See Sarah Boettiger, *Tempered Enthusiasm for Digitally Enabled Networks in International Development*, 9 INNOVATIONS: TECH., GOVERNANCE, GLOBALIZATION 132 (2014).

³⁴² *Id.*

³⁴³ See also GIGLER, *supra* note 293, at 234 (“[E]xternal socioeconomic variables are crucial for indigenous peoples’ ICT capabilities. . . a structural problem exists—that is, people in the highlands continue to be excluded from many basic infrastructure services, such as electricity, water and sanitation, and ICTs—that can explain the digital inequalities that persist within Bolivia. The geographic region with the highest levels of extreme poverty, highest illiteracy rates, and lowest access to basic infrastructure services also has the lowest ICT use and capabilities.”).

³⁴⁴ See GIGLER, *supra* note 293, at 199–200.

³⁴⁵ See Hamel, *supra* note 272, at 11.

³⁴⁶ NYABUGA & BOOKER, *supra* note 45, at 14. In sub-Saharan Africa, only fifteen percent of rural households have access to electricity. See Hamel, *supra* note 272, at 11.

programs to fund telecenters.³⁴⁷ Even though engineers are attempting to create ICT devices that consume less power, ICT use is still limited in areas lacking adequate power supply.³⁴⁸

Finding that the social, institutional and environmental context is not ripe for ICT uptake should not deter broadband infrastructure development in a developing country. Research has shown that the support of the government to ensure basic infrastructure development, along with quality legal frameworks, can lead to improved technology utilization.³⁴⁹

3. *How ICTs Impact Capabilities*

Believing there is a relationship between capabilities and technology, many conceive of the Internet and ICTs generally as a means to achieve certain ends and thus inherently a tool for increasing capabilities.³⁵⁰ Sen has also stated the importance of ICTs, recognizing that “access to the web and the freedom of general communication has become a very important capability.”³⁵¹ At first glance, tech companies’ connectivity agendas appear aligned with human-centered development by anticipating an expansion of individual capabilities resulting from newly available infrastructure. Google executives share the view that the extension of connectivity will lead to positive developments; the states of being educated, healthy, employed, and politically active are no doubt valued functionings of many people in both the developed and developing world.³⁵² However, as the Section above demonstrated, it should not be taken for granted that the introduction of ICTs ensures capabilities expansion. Moreover, although some ICT4D projects have noted positive effects on individual capabilities, studies also highlight

³⁴⁷ See *Universal Service Funds in Africa*, *supra* note 31, at 626 (“[A]lthough USF across Africa are used to construct telecentres in disadvantaged areas, more often than not, they are not sustainable as a result of factors such as the relevance of the services offered and the lack of electricity to power computers.”).

³⁴⁸ See Heeks, *supra* note 322, at 6 (noting three areas for innovation, including “new, low-cost devices for local electricity generation; better ways to store, carry and transmit electricity; and lower power consumption by ICT devices.”). In India, a village area network called DakNet is free from dependable power sources by using solar panels or a generator attached to a bicycle wheel. See Shivani Harnal & Jasbir Kaur, *Daknet (The Village Area Network)*, 6 INT’L J. ADVANCED RES. IN COMPUTER SCI. & SOFTWARE ENGINEERING, 303 (2016).

³⁴⁹ See Roztocki & Weistroffer, *supra* note 285, at 543.

³⁵⁰ See GIGLER, *supra* note 293, at 3–46; Coeckelbergh, *supra* note 300, at 84–85.

³⁵¹ See Amartya Sen, *Human Rights and Capabilities*, 6 J. HUM. DEV. 160 (2005).

³⁵² See *supra* note 270 and accompanying text.

that ICT use may negatively impact capabilities.³⁵³ This Section briefly examines the potential impact of newly introduced ICTs on existing and newly acquired capabilities by surveying ways that ICTs can work for and against capabilities.

Many ICT4D projects have noted positive effects on individual capabilities in the areas of education, health, and innovation. ICTs are commonly believed to enhance learning, especially in areas where teachers may lack adequate qualifications.³⁵⁴ Development practitioners also cite examples of ICTs enhancing the capability of innovation in the agricultural context, resulting in better living conditions and higher yields from improved agricultural knowledge.³⁵⁵ For example, Internet access often results in greater profits if farmers have pricing information for crops in different markets.³⁵⁶ Scholars commenting on the impact of ICTs on health tend to agree that the use of ICTs in medicine for knowledge management and service delivery can improve health outcomes.³⁵⁷

Nevertheless, critiques of the ICT4D agenda call into question the magnitude of positive effects of ICTs on individual capabilities. For instance, some question how much ICTs contribute to learning given that online content is not always available in a local language.³⁵⁸ Using the language of opportunity costs, scholars also view ICT investments as “tak[ing] away scarce resources from more urgent and direct development priorities, such as improving poor people’s access to education, water and sanitation, or health.”³⁵⁹

Beyond questioning the magnitude of positive effects, others note how ICTs may in fact exacerbate existing inequalities and impede the equitable development of individual capabilities. Sometimes, those who stand to benefit most from ICTs may be left out completely, especially if they lack digital literacy skills. This may perpetuate or exacerbate existing digital divides in

³⁵³ See Grunfeld et al., *supra* note 312, at 13 (noting that in a Cambodian ICT4D project, “[p]articipants had a strong inclination towards viewing ICT skills, the capability of learning in general and about agriculture, health, and education in Khmer, English, and typing, in particular, as the main capabilities to which iREACH had contributed.”); *but see* Hamel, *supra* note 272, at 58–59 (“ICT4D therefore only represents a potential for increasing opportunities and capabilities through technology, which can also increase inequality around the world and benefit only those that are able to gain from the new opportunities that ICTs facilitate. . .”).

³⁵⁴ See Grunfeld et al., *supra* note 312, at 14.

³⁵⁵ See *id.* at 16, 20; GIGLER, *supra* note 293, at 3–46.

³⁵⁶ GIGLER, *supra* note 293, at 8.

³⁵⁷ See Hamel, *supra* note 272, at 27–33.

³⁵⁸ See *id.* at 36.

³⁵⁹ GIGLER, *supra* note 293, at 6 (citing Gordon Wilson & Richard Heeks, *Technology, Poverty, and Development*, in *POVERTY AND DEVELOPMENT: INTO THE 21ST CENTURY* (T. Allen & A. Thomas Eds., OUP 2000)).

ICT usage, which may be because of gender, age, or social status.³⁶⁰ If marginalized groups are excluded from the information society and do not stand to reap the benefits of ICT use, capabilities will only be enhanced for those on the right side of the divide.

ICT usage may also have unintended consequences, including the undermining of individual capabilities. Although Internet access may yield greater profits for those in certain sectors like agriculture, there is also the concern about potential displacement of old workers with new technology, contributing to labor insecurity.³⁶¹ Moreover, the role of ICTs in connecting citizens to their government is often viewed as integral to enhancing democracy, promoting good governance via accountability, and boosting rule of law.³⁶² However, ICT usage by citizens also creates more opportunities for increased government surveillance, may trigger political instability, and can facilitate channels for encouraging violence.³⁶³ For example, the post-election violence in Kenya in 2007, where SMS was used to engage supporters in violence, shows how technology can contribute to both the worst and best of political and societal activism.³⁶⁴ In sum, ICTs do not always enhance capabilities, but the general belief remains that ICTs are a “crucial enabling infrastructure.”³⁶⁵ Nonetheless, technology is not neutral. It can lead to positive development impacts but may also disrupt communities and existing social and institutional arrangements.

4. *Community-led Development under the Capability Approach*

The importance of community involvement in infrastructure development cannot be understated. As much as communities and the social context can be macro-level barriers to the adoption of ICTs, discussed *supra*, communities can also be key proponents of ICT usage by encouraging uptake and adoption. ICT4D literature reporting on tech development projects notes

³⁶⁰ See Richard Heeks, *ICT4D 2016: New Priorities for ICT4D Policy, Practice and WSIS in a Post-2015 World*, 26 (Development Informatics, Working Paper Series No. 59, 2014).

³⁶¹ See Hamel, *supra* note 272, at 8 (“Many studies show that ICTs and the changes that accompany them are ‘demonstrably disruptive’ for many people in developing countries despite the wealth that they generate.”).

³⁶² See GIGLER, *supra* note 293, at 11; Boettiger, *supra* note 341, at 141.

³⁶³ See generally NYABUGA & BOOKER, *supra* note 45, at 7, 41 (discussing post-election violence); Mark I. Wilson and Kenneth E. Corey, *The Role of ICT in Arab Spring Movements*, 26 NETWORKS & COMM. STUD. 343 (2012) (explaining how ICTs played a major role in the Arab Spring as well as the Tunisian uprising); Policy Brief, CIPESA, *The Growing Trend of African Governments’ Requests for User Information and Content Removal From Internet and Telecom Companies*, (July 2017) https://cipesa.org/?wpfb_dl=248 (discussing state surveillance) [<https://perma.cc/JV56-38WV>].

³⁶⁴ See NYABUGA & BOOKER, *supra* note 45, at 7.

³⁶⁵ See SEN, *supra* note 301, at xii; Thompson, *supra* note 272, at 1.

that ICT usage and development of capability sets may result from community acceptance. Despite the emphasis by capability scholars on individual variables, some studies indicate that “[c]ommunity-level socioeconomic variables are more important determinants of ICT use than individual-level variables in rural communities in Bolivia.”³⁶⁶ This lends further support for the claim that ICTs need to be locally appropriated to ensure expansion of individual capabilities. In other words, communities need to value ICTs.

For ICTs to be accepted by communities and enhance capabilities, ICT use would need to be deemed a basic capability by the community. Some scholars reason “community capabilities represent all the valuable functionings and opportunities that ‘should’ be guaranteed for all members of the community” and should be the central focus for analysis and policy planning.³⁶⁷ Despite the emphasis by Sen on individual capabilities, the capability approach is compatible with group interpretations of well-being and governance processes whereby community priorities are set, which reflect capabilities and functionings that the group has reason to value. Though the capability approach promotes the use of agency by individuals and is “theoretically underspecified” with respect to groups,³⁶⁸ Sen sees a role for collective action when defining basic capabilities. Instead of defining a list of basic capabilities like Nussbaum, he encourages this list to be determined by public reasoning.³⁶⁹ Such an evaluative exercise encourages individuals to hold public discussions about their values and needs, as well as the use of democratic processes.

A community-led approach to infrastructure development may also inform the effective design of ICT programs, strengthen political freedom, and fortify the local tech industry. Such an approach would encourage local participation and, in developing countries, ensure projects are designed with the specific resources and demands of poor communities in mind. Bottom-up initiatives may help achieve universal access objectives if the infrastructure is

³⁶⁶ GIGLER, *supra* note 293, at xxxii.

³⁶⁷ Mario Biggeri & Andrea Ferrannini, *Opportunity Gap Analysis: Procedures and Methods for Applying the Capability Approach in Development Initiatives*, 15 J. HUM. DEV. & CAPABILITIES 60, 63 (2014).

³⁶⁸ Björn-Sören Gigler, *Indigenous Peoples, Human Development and the Capability Approach* 18 (Paper for the 5th International Conference on the Capability Approach, August 8, 2005), <https://ssrn.com/abstract=2924106> [<https://perma.cc/8YV3-J3UF>].

³⁶⁹ GIGLER, *supra* note 293, at 24–25, 26 (noting that Sen encourages this list “to be defined by the local context and people’s priorities.”); Sen, *supra* note 351, at 157.

not only locally appropriated, but also affordable and accessible.³⁷⁰ More importantly, initiatives that stem from the community give rise to the practice of political freedom if communities are involved in setting priorities with regards to capability development. From a capability perspective, infrastructure projects that simultaneously work to create and strengthen democratic institutions stand to gain more than short-term impacts.³⁷¹ A community-led approach to tech infrastructure would also further the long-term objective of creating digital transformation ecosystems in developing countries by strengthening the local tech industry.³⁷² Beyond using the Internet, the hope is for developing countries to one day enhance their tech capacity in order to invent new technologies, create local content, and innovate technological solutions to their development problems.³⁷³

Moreover, community-led development is desirable to resolve potential conflicts regarding the use of resources and improve the governance of infrastructure projects. Given that infrastructure resources are intermediate inputs and “shared means to many ends,”³⁷⁴ community engagement regarding desired infrastructure provision and capabilities expansion is crucial, especially because infrastructure goals often conflict.³⁷⁵ To resolve these conflicts, effective governance mechanisms are essential. Governance challenges are commonly cited for infrastructure projects: not only do elites and technical experts often make decisions, but the “potential combination of private interests, weak accountability mechanisms, and lack of transparency means that [infrastructure] goals might be implemented without balancing natural environment and wellbeing goals, and in a way that exacerbates

³⁷⁰ Universal service projects sometimes have been categorized as either top-down or bottom-up. “Top-down projects refer to government-led initiatives, with clearly defined targets in terms of area and service characteristics . . . Bottom-up initiatives are those in which the operators, nongovernmental organizations or other entities propose programs and projects and request funds to cover them based on a business plan.” Emiliani, *supra* note 25, at 20, 30.

³⁷¹ See Frediani et al., *supra* note 298, at 10.

³⁷² See Nagy K. Hanna, *How Can Developing Countries Make the Most of the Digital Revolution*, WORLD BANK: DIGITAL DEVELOPMENT (March 3, 2017), <https://blogs.worldbank.org/digital-development/how-can-developing-countries-make-most-digital-revolution> [<https://perma.cc/MJ33-LVEE>].

³⁷³ See Meghnad Desai et al., *Measuring the Technology Achievement of Nations and the Capacity to Participate in the Network Age*, 3 J. HUM. DEV. 95, 97 (2002).

³⁷⁴ See *supra* note 273 and accompanying text.

³⁷⁵ Jeff Waage et al., *Governing the UN Sustainable Development Goals: Interactions, Infrastructures, and Institutions*, 3 LANCET 251, 252 (2015) (“Infrastructure goals draw on common natural resources and realising them suggests some conflict with other goals at the same and different levels. For instance, achieving the energy or agriculture goal will have clear benefits for health and education but might be most easily and quickly achieved by actions that undermine biodiversity and climate change goals.”).

contemporary and intergenerational inequalities.”³⁷⁶ In fact, the SDGs encourage multi-stakeholder partnerships, which draw on democratically accountable actors, the private sector, and civil society, to resolve “social, economic, and development-related problems and challenges.”³⁷⁷ Engaging the community that the infrastructure projects are aimed to serve, providing avenues for deliberation, and abiding by transparency standards would further combat the skepticism about the “political economy” of ICT4D projects and whose interests such projects truly promote.

One such community-led infrastructure solution that addresses governance concerns is being implemented by a telecommunications nonprofit organization. Rhizomatica serves the state of Oaxaca, Mexico, where many rural villages are ignored by the major telecoms because they are deemed not profitable. Rhizomatica’s mission is to enlist direct community involvement and participation to develop decentralized telecommunications infrastructure.³⁷⁸ With Rhizomatica’s help, small indigenous communities in Mexico install, own, and operate base stations built with less expensive software and hardware that was once used by major telecom companies. Not only does the community have complete control over the network, but it also has control over the affordability of access. In one community, subscribers to the community network paid about \$2 per month for all local calls and texts, with the town retaining profits after paying electricity and maintenance costs.³⁷⁹ Such community-led infrastructure development improves acceptance of ICTs and promotes capabilities expansion, especially when a common barrier to access is tackled: cost. Moreover, with community owned and operated networks, the community is actively determining which functionalities ought to be guaranteed for all members.³⁸⁰ Some other positive impacts include local technical knowledge development, decreased dependence on foreign tech solutions, and community-governed technology

³⁷⁶ *See id.*

³⁷⁷ Alice Wanjira Munyua, *Exploring the Multi-Stakeholder Experience in Kenya*, 1 J. CYBER POL’Y 206, 207 (2016) (“There is no universally accepted definition of multi-stakeholder governance. It is considered a vehicle for collaboration and cooperation in the resolution of social, economic, and development-related problems and challenges . . .”). The UN defines partnerships as “voluntary and collaborative relationships between various parties, both state and non-state, in which all participants agree to work together to achieve a common purpose or undertake a specific task and to share risks and responsibilities, resources and benefits.” *Id.* (internal quotations and citations omitted).

³⁷⁸ *About Rhizomatica*, RHIZOMATICA, <https://www.rhizomatica.org/about/> (accessed Oct. 27, 2019) [<https://perma.cc/DM3Z-DWHY>].

³⁷⁹ Lizzie Wade, *Where Cellular Networks Don't Exist, People Are Building Their Own*, WIRED, Jan. 14, 2015, <https://www.wired.com/2015/01/diy-cellular-phone-networks-mexico/> [<https://perma.cc/6EUB-WTDD>].

³⁸⁰ Biggeri & Ferrannini, *supra* note 367, at 63.

infrastructure that reinforces the community's values and ways of association.³⁸¹

V. CONCLUSION

After years of significant financing of universal service funds to bridge the digital divide, many universal service missions remain unsolved and have created opportunities for the private sector to step in with innovative last mile solutions.³⁸² Nevertheless, as companies deploy innovative broadband infrastructure to blanket the globe with high-altitude connectivity, they should not ignore existing policies and institutional arrangements. Given the network structure of telecommunications, public policy will likely continue to play a significant role in telecom infrastructure production and regulation.³⁸³ Accordingly, innovators should work with governments (and their existing infrastructure development plans and universal service agendas), international development actors, and the communities the infrastructure is aimed to serve.³⁸⁴ As the world steps into a new age of interconnectedness with 5G networks and potentially dramatic societal challenges, such collaboration may help harness the Internet of Things for future global development. Without the support of other actors already on a mission to bridge the global digital divide, the potentially breakthrough infrastructure solutions to the global last mile may deliver unexceptional development outcomes.

There does not exist a template for bringing disruptive technology to market, let alone to rural markets, and many regulatory challenges concerning national airspace, outer space, and spectrum are guaranteed given the

³⁸¹ *What We Do*, RHIZOMATICA, <https://www.rhizomatica.org/what-we-do/> [<https://perma.cc/N2K7-E46X>].

³⁸² Frieden, *supra* note 99, at 448–49.

³⁸³ See Hausman et al., *supra* note 17, at 1.

³⁸⁴ There has been an uptake of universal service programs in developing countries and a recent surge in national broadband plans, promoting affordable, ubiquitous, and high-speed access, and programs designed specifically to teach ICT skills. See *The Benefits of Applying Universal Service Funds to Support ICT/Broadband Programs*, INTEL (2011), <https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/usf-support-ict-broadband-programs-paper.pdf> [<https://perma.cc/U9KN-H396>]; INT'L TELECOMM. UNION, *Universal Access/Service: Assessment Report, Harmonization of ICT Policies, Legislation and Regulatory Procedures in the Caribbean* (2013). The World Bank and ITU are keen to connect schools to the Internet and have been contemplating how to utilize national universal service or access funds for an ICT for education initiative. See INT'L TELECOMM. UNION, *THE UNIVERSAL SERVICE FUND AND DIGITAL INCLUSION FOR ALL STUDY* (2013); Michael Trucano, *Universal Service Funds & Connecting Schools to the Internet Around the World*, WORLD BANK: EDUTECH, (Feb. 26, 2015), <http://blogs.worldbank.org/edutech/universal-service-funds-connecting-schools-Internet-around-world> [<https://perma.cc/K7TS-MB3J>].

transnational and global nature of the infrastructure. A human-centered approach to commercialization plans for high-altitude connectivity infrastructure may help achieve positive development outcomes. Instead of assuming socioeconomic development in the aggregate will automatically follow the introduction of broadband infrastructure, companies seeking to deploy connectivity infrastructure should consider the real impact the infrastructure will have on individuals and communities, who are the intended beneficiaries of Internet access. This will require thinking beyond the next telecom partnership or deal, and instead will require engaging with the development literature, specifically ICT4D projects that have historically been carried out in many rural communities. Using Sen's capability approach, companies should not take the responsibility lightly to understand how newly introduced broadband infrastructure may or may not deliver "development as freedom."

HEALTH DATA AT YOUR FINGERTIPS: FEDERAL REGULATORY PROPOSALS FOR CONSUMER-GENERATED MOBILE HEALTH DATA

Jianyan Fang*

CITE AS: 4 GEO. L. TECH. REV. 125 (2019)

TABLE OF CONTENTS

I. INTRODUCTION	126
II. BIG HEALTH DATA, BIG CONCERNS, AND MHEALTH APP DATA	132
A. Categories of mHealth Apps and Working Definitions	132
B. Context Matters: Defining Health Data in the Era of Big Data	135
C. Big Health Data, Big Concerns	138
D. Under-regulated mHealth App Data	140
III. CURRENT FEDERAL STATUTORY LANDSCAPE AND GAPS	143
A. HIPAA: A Closed, Downstream Approach	144
B. FTC's Section 5 Power: Only If You Break Your Promise	148
IV. GAP-FILLING PROPOSALS	151
A. Necessity of Governmental Action to Address Regulatory Gaps... 151	
1. <i>Privacy in the Digital Age: Shifted Expectations or Trade-off?</i> 151	
a. Shifted Expectations of Privacy	152
b. Trade-off as a Fallacy	155
2. <i>The Illusion of Successful Self-Regulation</i>	157
B. Comprehensive or Sectoral: A Pragmatic Perspective	161
1. <i>All Previous Attempts of Comprehensive Solutions Failed</i>	162
2. <i>Constitutional Challenges</i>	163
3. <i>Stakeholder Concerns</i>	165
4. <i>Health Privacy Exceptionalism</i>	166

* LL.M. graduate, Harvard Law School. I am grateful to Professor I. Glenn Cohen, who supervised this article's preparation and offered insightful comments. I also benefitted from the helpful discussions with Professor Jonathan Zittrain and Professor William W. Fisher and the valuable comments from Jane Fair Bestor, Oren Tamir, and editors of this journal. Finally, I appreciate the research support provided by the staff of the Harvard Law School Library, especially Jennifer Allison.

C. A Two-Prong Solution.....	167
1. <i>First Things First: Categorizing mHealth App Data</i>	168
a. Different Approaches in Defining Regulated Health Data .	168
(1) Regulated Health Data under HIPAA.....	168
(2) Regulated Health Data under GDPR	169
(3) Regulated Health Data under EU mHealth Code	170
b. Proposed Approach in Categorizing Regulated Data	170
2. <i>Expanding HIPAA to Cover mHealth Data</i>	171
3. <i>FTC-led Co-regulation: Taking It a Step Further</i>	175
a. Success Stories of Co-regulation	175
b. Proposed Co-Regulation Approach for mHealth Consumer Data.....	177
V. CONCLUSION.....	179

“DATA IS THE NEW OIL.”¹

“HEALTH DATA IS VALUABLE: YOUR EMPLOYER WANTS IT, YOUR INSURERS WANT IT, AND YOU’RE ONLY TOO HAPPY TO GIVE IT AWAY TO APPS FOR FREE.”²

I. INTRODUCTION

Are you male or female? When were you born? How tall are you? How much do you weigh? You may have encountered these questions when signing up with a fitness mobile application (app) like MyFitnessPal.³ After you quickly provide this personal information, and consent to the app’s privacy terms and cross-border data transfer policy, the app would be able to track, store, and analyze your diet, steps, exercise, and other daily activities to help you lose weight and promote a healthy lifestyle.

When you search terms such as “health,” “fitness,” “wellness,” and “diagnosis” in the App Store or Google Play Store, hundreds of apps like MyFitnessPal will pop up—all making promises of better health. These apps target a variety of users such as consumers, doctors, and medical students.

Statistics show that around 318,000 mobile health (mHealth) apps are now available in major app stores, and the global mHealth app market is

¹ Nicolas P. Terry, *Will the Internet of Things Transform Healthcare*, 19 VAND. J. ENT. & TECH. L. 327, 337 (2016).

² Angela Lashbrook, *There Is a Reason Apps Make It So Fun to Track Your Health the Outline*, THE FUTURE (Feb. 1, 2019, 1:30 PM), <https://theoutline.com/post/7039/there-is-a-reason-apps-make-it-so-fun-to-track-your-health> [<https://perma.cc/FB4W-GHAM>].

³ MYFITNESSPAL, <https://www.myfitnesspal.com/> [<https://perma.cc/QEV3-E56D>].

expected to reach 111 billion U.S. dollars by 2025.⁴ Half a billion smartphone users have installed at least one mHealth app on their phones.⁵ Continuous market growth and widespread use of these apps have made mHealth an important segment of the health industry and part of our daily lives. mHealth technologies are allowing the devices we take everywhere to constantly collect and share our health data.

Many believe mHealth has the potential to strengthen the “iron triangle of health care” by enhancing quality, decreasing cost, and improving access.⁶ However, these benefits are not without privacy risks. In November 2018, DeepMind Health, a London-based health app team focusing on artificial intelligence (AI) research and mobile tools,⁷ announced that it was joining Google Health.⁸ This news has caused wide concerns over DeepMind Health’s independence from Google in dealing with health data.⁹ Only weeks later, the French regulator fined Google nearly \$57 million for failing to properly disclose its data collection across its various services in accordance with the European Union (EU) General Data Protection Regulation (GDPR),¹⁰ suggesting that such concerns are not without basis. In the United States, Facebook was reported to have received health data from third-party health

⁴ 11 *Surprising Mobile Health Statistics*, MOBIUS MD, <https://www.mobius.md/blog/2019/03/11-mobile-health-statistics/> [<https://perma.cc/9GP6-CJKT>].

⁵ Dov Greenbaum, *Avoiding Overregulation in the Medical Internet of Things*, in *BIG DATA, HEALTH LAW, AND BIOETHICS* 129, 132 (I. Glenn Cohen et al. eds., 2018).

⁶ Cheng-Kai Kao & David M. Liebovitz, *Consumer Mobile Health Apps: Current State, Barriers, and Future Directions*, 9 *PHYSICAL MED & REHABILITATION* 106, 106 (2017).

⁷ Dominic King, *DeepMind’s Health Team Joins Google Health*, DEEPMIND, (Sept. 18, 2019), <https://deepmind.com/blog/announcements/deepmind-health-joins-google-health> [<https://perma.cc/D4JA-TXRG>].

⁸ Demis Hassabis, Mustafa Suleyman & Dominic King, *Scaling Streams with Google*, DEEPMIND (Nov. 13, 2018), <https://deepmind.com/blog/announcements/scaling-streams-google> [<https://perma.cc/4W7J-XL2P>].

⁹ Margi Murphy, *Privacy Concerns as Google Absorbs DeepMind’s Health Division*, TELEGRAPH (Nov. 13, 2018, 10:40 PM), <https://www.telegraph.co.uk/technology/2018/11/13/privacy-concerns-google-absorbs-deepminds-health-division/> [<https://perma.cc/867Y-ZZ2P>].

¹⁰ Adam Satariano, *Google Is Fined \$57 Million Under Europe’s Data Privacy Law*, N.Y. TIMES (Jan. 21, 2019), <https://www.nytimes.com/2019/01/21/technology/google-europe-gdpr-fine.html> [<https://perma.cc/GS5T-WVAA>].

and fitness apps, potentially without notifications to users;¹¹ the data transmitted included diet information, exercise activities, ovulation cycle, and intention to get pregnant.¹² Five months later, Facebook was fined \$5 billion by the Federal Trade Commission (FTC) based on its repeated breaches of its previous 2011 privacy protection settlement with the regulator.¹³

Unlike the EU and many other peer countries such as Canada, Israel, and Japan,¹⁴ the United States is not keen on comprehensive protection of personal data. In the absence of an overarching data protection framework, the United States sticks with a “sectoral” privacy regulatory system¹⁵ where only certain sensitive, high-stakes sectors receive exceptional statutory privacy protection. In terms of general privacy concerns, the United States relies on the competent regulator’s case-by-case, light-touch enforcement actions from the perspective of consumer protection.¹⁶

Unsurprisingly, healthcare has been singled out as one of the sectors receiving exceptional privacy protection. Health data is traditionally protected by the Health Insurance Portability and Accountability Act (HIPAA), which is enforced by the U.S. Department of Health and Human Services (HHS).¹⁷

¹¹ Sam Schechner & Mark Secada, *You Give Apps Sensitive Personal Information. Then They Tell Facebook.*, WALL ST. J. (Feb. 22, 2019, 11:07 AM), <https://www.wsj.com/articles/you-give-apps-sensitive-personal-information-then-they-tell-facebook-11550851636> [<https://perma.cc/GC8T-8UHC>]; see also Nick Statt, *App Makers Are Sharing Sensitive Personal Information with Facebook but Not Telling Users*, VERGE (Feb. 22, 2019, 2:00 PM), <https://www.theverge.com/2019/2/22/18236398/facebook-mobile-apps-data-sharing-ads-health-fitness-privacy-violation> [<https://perma.cc/R6QW-MXDR>].

¹² *Id.*

¹³ Levi Sumagaysay, *Facebook Settlement Confirmed: Are \$5 Billion Fine and Limits on Zuckerberg Enough?*, MERCURY NEWS (July 24, 2019, 3:34 PM), <https://www.mercurynews.com/2019/07/24/facebook-settlement-confirmed-are-5-billion-fine-and-limits-on-zuckerberg-enough/> [<https://perma.cc/DPS7-WD85>].

¹⁴ Nuala O’Connor, *Reforming the U.S. Approach to Data Protection and Privacy*, COUNCIL ON FOREIGN REL. (Jan. 30, 2018), <https://www.cfr.org/report/reforming-us-approach-data-protection> [<https://perma.cc/EDN6-9TXC>].

¹⁵ Anne Marie Helm & Daniel Georgatos, *Privacy and MHealth: How Mobile Health Apps Fit into a Privacy Framework Not Limited to HIPAA*, 64 SYRACUSE L. REV. 131, 140 (2014).

¹⁶ See generally Woodrow Hartzog & Daniel J. Solove, *The Scope and Potential of the FTC Data Protection*, 83 GEO. WASH. L. REV. 2230 (2015) (discussing the scope of FTC’s authority in the area of data protection and its limits, arguing that FTC’s current modest enforcement only focuses on the most egregious violations, and urging FTC to strengthen and improve its enforcement actions); Fed. Trade Comm’n, *A Brief Overview of the Federal Trade Commission’s Investigative and Law Enforcement Authority*, FTC.GOV (revised Oct. 2019) <https://www.ftc.gov/about-ftc/what-we-do/enforcement-authority> (providing an overview of FTC’s various enforcement powers including, among others, its consumer protection power under Section 5 of the FTC Act) [<https://perma.cc/66SK-WJV6>]; see also discussions *infra* Part III

¹⁷ Margaret Foster Riley, *Big Data, HIPAA, and the Common Rule Time for Big Change?*, BIG DATA, HEALTH LAW, AND BIOETHICS 251, 260 (I. Glenn Cohen et al. eds., 2018).

However, mHealth has posed new challenges for HIPAA. Enacted in pre-mHealth times, HIPAA regulates health data only to the extent that it is disclosed and used by covered entities¹⁸ and business associates¹⁹ as defined under HIPAA.²⁰ As the collectors and custodians of HIPAA-regulated health data, most HIPAA-defined covered entities and business associates are traditional players within the healthcare sector; they typically include health plans, healthcare clearinghouses, healthcare providers and independent contractors acting on their behalf.²¹ Because most consumer-facing mHealth apps are developed and used without the involvement of HIPAA-defined covered entities or business associates, health data collected and generated by these consumer grade apps usually falls outside the purview of HIPAA.²²

In addition to HHS, FTC—as mentioned in the Facebook example above—and the Food and Drug Administration (FDA) are two of the primary federal regulators of the mHealth industry identified by the American Health Information Management Association.²³ Both FTC and FDA have intervened in this industry to varying degrees, but have failed to provide complete protection for mHealth app data. FTC’s enforcement power is broad,²⁴ but the Commission must establish “deceptiveness”²⁵ or “unfairness”²⁶ before policing any breach of mHealth app data. FDA, whose focus is to supervise the efficacy of mHealth apps and protect public health,²⁷ is not concerned with protecting privacy.²⁸

The development of the mHealth industry has outpaced the existing federal health data protection regime. To fill the regulatory gaps,

¹⁸ 45 C.F.R. § 160.103 (2019).

¹⁹ *Id.*

²⁰ 45 C.F.R. § 160.102 (2019); *see also* discussion *infra* Part III.A.

²¹ 45 C.F.R. § 160.102 (2019).

²² See Jessica Davis, *HHS Clarifies HIPAA Liability Around Third-Party Health Apps*, XTELLIGENT HEALTHCARE MEDIA, (Apr. 12, 2019), <https://healthitsecurity.com/news/hhs-clarifies-hipaa-liability-around-third-party-health-apps> [<https://perma.cc/46QY-HJ84?type=image>].

²³ Y. Tony Yang & Ross D. Silverman, *Mobile Health Applications: The Patchwork of Legal and Liability Issues Suggests Strategies to Improve Oversight*, 33 HEALTH AFF. 222, 223 (2014) (noting that in addition to HHS, the FTC, and the FDA, the National Institute of Standards and Technology and the Federal Communications Commission are also likely to be involved in regulating mHealth apps).

²⁴ Hartzog & Solove, *supra* note 16, at 2246.

²⁵ 15 U.S.C. § 45(a)(1).

²⁶ *Id.*

²⁷ Matthew R. Langley, *Hide Your Health: Addressing the New Privacy Problem of Consumer Wearables*, 103 GEO. L.J. 1641, 1650 (2015).

²⁸ FUTURE PRIVACY FORUM, BEST PRACTICES FOR CONSUMER WEARABLES & WELLNESS APPS & DEVICES, 1 n.2 (2016), <https://fpf.org/wp-content/uploads/2016/08/FPF-Best-Practices-for-Wearables-and-Wellness-Apps-and-Devices-Final.pdf> [<https://perma.cc/R3S6-HJ8U>].

commentators have made various recommendations. Many have proposed expanding and updating HIPAA to include and accommodate “non-HIPAA” health data.²⁹ Some have suggested passing a specific, standalone statute to protect non-HIPAA health data.³⁰ Others have considered self-regulation and

²⁹ See generally Alexis Guadarrama, *Mind the Gap: Addressing Gaps in HIPAA Coverage in the Mobile Health Apps Industry*, 55 HOUS. L. REV. 999 (2018) (discussing the gap of HIPAA coverage in regulating mobile health apps and suggesting expanding HIPAA to cover private health data regardless of who holds it); Latena Hazard, *Is Your Health Data Really Private: The Need to Update HIPAA Regulations to Incorporate Third-Party and Non-Covered Entities*, 25 CATH. U. J. L. & TECH 447 (2017) (discussing, among others, HIPAA rules and their effect on health apps, and arguing that HIPAA needs to be adjusted to incorporate non-covered entities); Nicolas P. Terry, *Regulatory Disruption and Arbitrage in Health-Care Data Protection*, 17 YALE J. HEALTH POL’Y L. & ETHICS 143 (2017); Grant Arnow, *Apple Watching You: Why Wearable Technology Should Be Federally Regulated*, 49 LOY. L.A. L. REV. 607 (2016) (proposing establishment of a standalone, cabinet-level department to align and unify national Internet data protection efforts and suggesting updating HIPAA); Elizabeth A. Brown, *The Fitbit Fault Line: Two Proposals to Protect Health and Fitness Data at Work*, 16 YALE J. HEALTH POL’Y L. & ETHICS 1 (2016) (discussing the growth of employee-facing health and fitness apps and increased employee monitoring, examining the current data protection regulations and agency actions, and proposing the adoption of a mandatory privacy labeling law for health-related devices and apps, and suggesting expanding HIPAA’s protection); Nicolas P. Terry, *Big Data Proxies and Health Privacy Exceptionalism*, 24 HEALTH MATRIX 65 (2014); Nicolas Terry, *Health Privacy is Difficult But Not Impossible in a Post-HIPAA Data-driven World*, 146 CHEST 835 (2014) (taking the position that health-care-data exceptionalism remains a valid policy and the current HIPAA protection model should be maintained and re-calibrated to consider the upstream and point-of-use protections and protect healthcare data residing outside of the traditional health-care domain).

³⁰ See generally Michelle M. Christovich, *Why Should We Care What Fitbit Shares—A Proposed Statutory Solution to Protect Sensitive Personal Fitness Information*, 38 HASTINGS COMM. & ENT. L.J. 91 (2016) (suggesting, among others, that Congress should adopt a statutory scheme modeled after HIPAA to adequately protect users’ privacy rights against the misuse of fitness information).

co-regulation approaches for protection of mHealth app data, or more generally, consumer privacy.³¹

This Article joins these discussions to explore how mHealth app data can be better protected on the federal level. It aims to contribute to the academic literature in two ways. First, it engages in a detailed rethinking of the concept of health data and frames a two-step approach to define health data in the era of Big Data. Second, unlike those scholarly discussions proposing one-size-fits-all solutions, this Article envisions a single two-prong solution to accord different levels of protection to different categories of mHealth app data and to balance the necessity of privacy protection with commercial needs. The first prong is to expand HIPAA to cover high-stakes mHealth app data that qualifies as health data warranting extra protection under the new definition proposed by this Article. The second prong is to adopt a co-regulation approach to govern the less sensitive mHealth app data that does not fall under the proposed new definition of health data.

This Article focuses exclusively on federal-level regulations; state-level regulatory solutions are therefore not reviewed. Furthermore, this Article treats protection of mHealth app data as a consumer privacy issue and will discuss this issue only from the perspective of the private sector. It will not consider the regulation of the U.S. government's collection and use of health data.

The remainder of this Article proceeds as follows: Part II portrays the current state of mHealth apps and mHealth app data and proposes a two-step approach to define health data in the era of Big Data. Part III assesses the current U.S. federal regulatory system of health data and identifies the regulatory gaps for protection of mHealth app data. Part IV argues that regulation is necessary for mHealth app data, and that the government needs

³¹ See generally Bradyn Fairclough, *Privacy Piracy: The Shortcomings of the United States' Data Privacy Regime and How to Fix It*, 42 J. CORP. L. 461 (2016) (discussing the key role of collaborative governance in creating a regulatory framework that both protects consumers and the businesses); J. Frazee, M. Finley & J.J. Rohack, *mHealth and Unregulated Data: Is This Farewell to Patient Privacy*, 13 IND. HEALTH L. REV. 384 (2016) (proposing a voluntary labelling system and arguing that allowing voluntary adoption of the labelling system rather than mandating increased privacy protections would allow companies to provide free options to consumers while providing privacy conscious consumers with a meaningful choice); Dennis D. Hirsch, *Going Dutch: Collaborative Dutch Privacy Regulation and the Lessons It Holds for U.S. Privacy Law*, 2013 MICH. ST. L. REV. 83 (2013) (introducing the Dutch code of conduct approach, and discussing the lessons the US can draw from the Dutch experience in maximizing strengths and minimizing weaknesses of the code of conduct approach); Ira S. Rubinstein, *Privacy and Regulatory Innovation: Moving Beyond Voluntary Codes*, 6 ISJLP 355 (2011) (discussing different types of self-regulation and exploring co-regulatory approaches in which government sets requirements and imposes sanctions for non-compliance for the purpose of protecting online privacy).

to play a role in such regulation. The final part of this Article discusses in detail the proposed two-prong solution for regulating mHealth app data.

II. BIG HEALTH DATA, BIG CONCERNS, AND MHEALTH APP DATA

A. Categories of mHealth Apps and Working Definitions

In general, any medical and public health practice supported by mobile communication devices such as smartphones, tablets, wearable technology, and other wireless devices falls under the definition of “mobile health” or “mHealth.”³² mHealth apps can be downloaded and installed on mobile devices to perform designated medical or health-related functions.³³

Scholars have developed various typologies for mHealth apps, most of which are functionality-based.³⁴ As one example, Nathan Cortez, a leading figure in the regulation of mobile health technologies and FDA regulation, divides mHealth apps into the following categories based on their respective functions: (1) connectors which connect mobile devices to FDA-regulated devices and thus amplify such regulated devices’ functionalities; (2) replicators which turn mobile devices into FDA-regulated devices; (3) automators and customizers which use different methodologies including questionnaires, algorithms, formulae and medical calculators to aid clinical decisions; (4) informers and educators which primarily inform and educate users; (5) administrators which automate office functions such as identifying insurance billing codes or scheduling patient appointments; and (6) loggers and trackers which allow users to log, record, and make decisions about general health and wellness.³⁵

³² WORLD HEALTH ORG., MHEALTH: NEW HORIZONS FOR HEALTH THROUGH MOBILE TECHNOLOGIES, 6 (2011), https://www.who.int/goe/publications/goe_mhealth_web.pdf [<https://perma.cc/6MBA-9TT7>]; see also Nathan Cortez, *The Mobile Health Revolution*, 47 U.C. DAVIS L. REV. 1173, 1176 (2014) (defining “mobile health” as “the use of mobile communications devices like smartphones and tablet computers for health or medical purposes, usually for diagnosis, treatment, or simply well-being and maintenance”); Anne Marie Helm & Daniel Georgatos, *Privacy and MHealth: How Mobile Health Apps Fit into a Privacy Framework Not Limited to HIPAA*, 64 SYRACUSE L. REV. 131, 134 (2014) (noting that “mHealth occurs when a provider of healthcare services uses connected and interactive mobile computing to produce, access, transmit, or store data for the provision of healthcare services to patients, or when a patient or consumer uses connected and interactive mobile computing to produce, access, transmit, store, or otherwise share data for a health-related purpose”).

³³ Guadarrama, *supra* note 29, at 1002.

³⁴ See e.g., Nicholas P. Terry & Lindsay F. Wiley, *Liability for Mobile Health and Wearable Technologies*, 25 ANNALS HEALTH L. 62, 68 (2016); Cortez, *supra* note 32, at 1176.

³⁵ *Id.* at 1182–89.

In regulating the efficacy of mHealth apps, FDA has also adopted a similar functionality-based typology, but with more of a focus on risk control.³⁶ In its Policy for Device Software Functions and Mobile Medical Applications, FDA divides mHealth apps into three tiers.³⁷ The top tier represents “mobile medical apps” that operate as extensions of FDA-regulated medical devices, turn mobile platforms³⁸ into FDA-regulated devices, or function like FDA-regulated medical devices for analysis, diagnosis, and treatment purposes.³⁹ They qualify as “devices” defined under Section 201(h) of the Federal Food, Drug, and Cosmetic Act (FDCA),⁴⁰ and would pose a risk to a patient’s safety if they do not function as intended.⁴¹ The middle tier represents those mobile apps that *may* meet the definition of devices under the FDCA, but are not regulated as mobile medical apps because of their low risks.⁴² The bottom tier represents mHealth apps that are not medical devices and therefore are not administered by FDA.⁴³ FDA claims that it intends (1) to apply oversight to the top tier of mHealth apps, (2) to exercise enforcement discretion (meaning it will not impose FDCA requirements)⁴⁴ over the middle tier, and (3) not to regulate the bottom tier.⁴⁵

Some mHealth apps are consumer-facing, some require expertise, and many do not differentiate amongst target users at all.⁴⁶ Such user-based differentiation is meaningful for the purpose of this Article, and from a regulatory perspective as well. Because professional-facing mHealth apps are usually used or prescribed by HIPAA-regulated entities, such as healthcare

³⁶ See generally U.S. FOOD & DRUG ADMIN., POLICY FOR DEVICE SOFTWARE FUNCTIONS AND MOBILE MEDICAL APPLICATIONS GUIDANCE FOR INDUSTRY AND FOOD AND DRUG ADMINISTRATION STAFF (2019), <https://www.fda.gov/media/80958/download> (providing detailed explanations in classifying mobile apps based on their potential risks posed to public health) [<https://perma.cc/2WRD-5RHB>].

³⁷ See generally *id.*

³⁸ *Id.* at 4 (Mobile platforms include smart phones, tablet computers, or other portable computers.).

³⁹ *Id.* at 11–12.

⁴⁰ 21 U.S.C. § 321(h) (2018) (defining a “device” in relevant part as: “an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including any component, part, or accessory, which is . . .intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals, or intended to affect the structure or any function of the body of man or other animals”).

⁴¹ U.S. FOOD & DRUG ADMIN., *supra* note 36, at 2.

⁴² *Id.* at 9.

⁴³ *Id.* at 16.

⁴⁴ *Id.* at 2.

⁴⁵ *Id.*

⁴⁶ Cortez, *supra* note 32, at 1177.

professionals,⁴⁷ and data collected and generated by these apps is generally within the reach of HIPAA, they are not a concern of this Article. In contrast, consumer-facing mHealth apps, which are at the heart of this Article, are usually developed and used by entities not governed by HIPAA. Data collected and generated by these mHealth apps is by and large in a regulation-free zone.

However, under the current regulatory framework, this classification is not absolutely binary because the applicable data protection rules may vary as data changes hands. For example, a patient's blood pressure stored in a hospital's Electronic Health Record (EHR) is initially regulated by HIPAA, but it may escape from HIPAA's regulation once the patient downloads and inputs the same information into a consumer-facing mHealth app. Conversely, a user's blood pressure collected by a health management mobile app is free from regulation by HIPAA initially, but it will end up in EHR and be governed by HIPAA if the app user later transmits the blood pressure results to his or her physician for diagnosis or treatment purposes. This convertible situation will be further discussed in Part II of this Article.

To facilitate the subsequent discussions, unless otherwise stated, the term "mHealth app" hereinafter refers to consumer-facing mHealth apps used by individual consumers without the direct involvement of conventional healthcare providers⁴⁸ or other HIPAA-defined covered entities and business associates. The term "mHealth app data" refers to data of any nature that is collected and generated by consumer-facing mHealth apps and is not subject to the governance of the current HIPAA.⁴⁹ According to the various functions of mHealth apps discussed earlier, mHealth app data typically includes three groups of data: (1) health-related data, including medical history, test results, and clinical data, that has inherent medical significance and is used to identify an app user's particular condition, (2) biometrics information and lifestyle data that are somewhat related to an app user's body or general health and wellness, and (3) data that is normally not health-related, such as an app user's geolocation, identity, contact list, payment records, and social or recreational information.

⁴⁷ Nicolas P. Terry & Tracy D. Gunter, *Regulating Mobile Mental Health Apps*, 36 BEHAV. SCI. L. 136, 139 (2018).

⁴⁸ Nicolas P. Terry, *Mobile Health: Assessing the Barriers*, 147 CHEST 1429, 1430 (2015).

⁴⁹ *But see* Nicolas P. Terry, *Regulatory Disruption and Arbitrage in Health-Care Data Protection*, 17 YALE J. HEALTH POL'Y L. & ETHICS 143, 182 (2017) (mentioning that although the majority of mHealth apps are operating in the HIPAA-free zone, some developers of apps or wearables are beginning to advertise HIPAA-compliance).

B. Context Matters: Defining Health Data in the Era of Big Data

Big Data has obscured the distinction between different categories of data,⁵⁰ and is making it difficult to define precisely what health data encompasses. This Section will first discuss the concerns and problems brought about by Big Data analytics, and then propose a new definition of health data in the era of Big Data.

There are different understandings of Big Data. Some stress the characteristics of the data at issue and refer to Big Data as “large volumes of high-velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information.”⁵¹ Others focus on effects of the application of Big Data technologies and define Big Data as “a confluence of factors, including the nearly ubiquitous collection of consumer data from a variety of sources, the plummeting cost of data storage, and powerful new capabilities to analyze data to draw connections and make inferences and predictions.”⁵² Despite the nuances, the key implication is that Big Data analytics can shift the context of data use and creates a world where the significance and sensitivity of a single set of data will vary with the changing context. As a result, mundane data that is not inherently health-related could reveal health-related correlations or conclusions if aggregated and analyzed with other datasets.

Powered by Big Data analytics, ordinary consumer data is now widely used in health-related contexts. For example, a high school girl’s purchasing data with the department store Target has been used to predict that she was pregnant – even before her father found out;⁵³ a childless man who does online clothing shopping, spends a lot on cable TV, and drives a minivan is inferably overweight;⁵⁴ a woman who purchases plus-size clothing is considered at risk

⁵⁰ Tal Z. Zarsky, *Incompatible: GDPR in the Age of Big Data*, 47 SETON HALL L. REV. 995, 1013 (2017).

⁵¹ Riley, *supra* note 17, at 252.

⁵² FED. TRADE COMM’N, *BIG DATA A TOOL FOR INCLUSION OR EXCLUSION: UNDERSTANDING THE ISSUES 1* (2016), <https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf> [<https://perma.cc/6P3X-4PA3>].

⁵³ Kashmir Hill, *How Target Figured Out a Teen Girl Was Pregnant Before Her Father Did*, FORBES (Feb. 16, 2012, 11:02 AM), <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#4ee5d5286668> [<https://perma.cc/XBD4-VLRK>].

⁵⁴ FRANK PASQUALE, *THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION* 30 (2015).

of depression,⁵⁵ and people downsizing homes tend to incur higher healthcare costs.⁵⁶ The former FTC Chairwoman Edith Ramirez has termed these predictions and inferences “data determinism” saying:

[Persons are judged] . . . not because of what they’ve done, or what they will do in the future, but because inferences or correlations drawn by algorithms suggest they may behave in ways that make them poor credit or insurance risks, unsuitable candidates for employment or admission to schools or other institutions, or unlikely to carry out certain functions.⁵⁷

Directed by data determinism, consumer categories such as “Expectant Parent,” “Diabetes Interest,” and “Cholesterol Focus” are created based on various ordinary consumer data that would not have been medically meaningful but for Big Data analytics.⁵⁸ Insurance companies,⁵⁹ credit-card companies,⁶⁰ and pharmaceutical companies⁶¹ all utilize Big Data analytics to predict costs, calibrate rates, evaluate risks, and optimize target advertising.

Health professionals also seem to accept this blurred divide between health data and other data. Recognizing that many non-health social determinants and indicators are outside of the medical system,⁶² the Institute of Medicine (IOM) recommended several years ago that federal health information technology (IT) policymakers should add new social and

⁵⁵ Marshall Allen, *Health Insurers Are Vacuuming Up Details About You—And It Could Raise Your Rates*, PROPUBLICA (Jul. 17, 2018, 5:00 AM), <https://www.propublica.org/article/health-insurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates> [<https://perma.cc/82P9-EFJY>].

⁵⁶ *Id.*

⁵⁷ Nicolas P. Terry, *Big Data Proxies and Health Privacy Exceptionalism*, 24 HEALTH MATRIX 65, 79 (2014), <http://scholarlycommons.law.case.edu/cgi/viewcontent.cgi?article=1005&context=healthmatrix> [<https://perma.cc/ZQ8X-MGSH>].

⁵⁸ FED. TRADE COMM’N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY 47 (2014), <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf> [<https://perma.cc/J7CP-P937>].

⁵⁹ Kaiser Health News, *Health Insurers Are Vacuuming Up Consumer Data That Could Be Used to Raise Rates*, HEALTHLEADERS (Jul. 17, 2018), <https://www.healthleadersmedia.com/finance/health-insurers-are-vacuuming-consumer-data-could-be-used-raise-rates> [<https://perma.cc/7YSG-Q92S>].

⁶⁰ Jay Hancock, *Is Your Private Health Data Safe In Your Workplace Wellness Program?*, PBS NEWSHOUR (Sep. 30, 2015, 6:07 PM), <https://www.pbs.org/newshour/health/many-workplace-wellness-programs-dont-follow-health-privacy-laws> [<https://perma.cc/D9C4-Q7D3>].

⁶¹ Guadarrama, *supra* note 29, at 1013.

⁶² Frank Pasquale, *Redescribing Health Privacy: The Importance of Information Policy*, 14 HOUS. J. HEALTH L. & POL’Y 95, 126 (2014).

behavioral information as part of the EHR to improve clinical research and healthcare delivery.⁶³ The recommended determinants and indicators include individuals' financial resource strain, level of physical activity, level of stress, educational status,⁶⁴ dietary patterns, employment, sexual orientation, and even neighborhood and community compositional characteristics.⁶⁵ As an example, according to IOM, the geographic location where an individual lives or works has no health-related significance on its face, but it can become medically revealing for the purpose of individual healthcare delivery or public health policies when it interacts with other datasets such as air pollution, the availability of sidewalks, public transportation, and healthy food options.⁶⁶

A question follows: If the distinction between health data and other data is fading, is it even possible to define health data in this evolving setting? This Article argues that it is still possible to define health data because this blurring or confluence is not absolute but rather context-specific.

As discussed above, in the era of Big Data, aside from data's inherent nature, the context in which data is used is also a factor determining whether it will be considered health data in a particular circumstance. Accordingly, this Section proposes a two-step approach for defining health data.

First, the data in question should always be considered health data if it is intrinsically of medical significance, regardless of its sources, contexts for use, and ultimate purposes. For example, insulin and blood glucose levels should always be regarded as health data, whether it is collected by a caregiver at a clinic or by a patient through a diabetes mobile app, whether it is analyzed to evaluate the treatment effect or recorded to observe the development of diabetes, or whether it is ultimately for dividing different consumer groups or treating a particular patient.

Second, if the data fails the first step, it should be considered health data only to the extent that it is intended to be used when collected, or is actually used later on, to conduct health-related analysis, determine health-related correlations, draw health-related conclusions, or make health-related predictions. For example, one's diet pattern is nothing special if it is only used to record daily life activities and recommend a good lifestyle. However, the same diet pattern should be considered health data when used to evaluate your

⁶³ Joseph Conn, *IOM Panel Urges More EHR Collection of Social, Behavioral Data*, MODERN HEALTHCARE (Nov. 13, 2014, 12:00 AM), <https://www.modernhealthcare.com/article/20141113/NEWS/311139943/iom-panel-urges-more-ehr-collection-of-social-behavioral-data> [<https://perma.cc/8GAM-NMVP>].

⁶⁴ *Id.*

⁶⁵ See generally INST. OF MED., CAPTURING SOCIAL AND BEHAVIORAL DOMAINS IN ELECTRONIC HEALTH RECORDS: PHASE 1 (2014); INST. OF MED., CAPTURING SOCIAL AND BEHAVIORAL DOMAINS IN ELECTRONIC HEALTH RECORDS: PHASE 2 (2014).

⁶⁶ INST. OF MED., CAPTURING SOCIAL AND BEHAVIORAL DOMAINS IN ELECTRONIC HEALTH RECORDS: PHASE 1, 42 (2014).

risk of hypertension, no matter whether that risk evaluation is for purposes of diagnosing a particular disease by a caregiver or for rating health risks by an insurance company.

C. Big Health Data, Big Concerns

The expansion of health data is not always a blessing for consumers. If misused in combination with Big Data analytics, health data could operate against consumers' interests. The three main concerns to be discussed in this Section are: (1) psychological harms associated with unauthorized disclosure, (2) algorithmically imposed new discrimination threats, and (3) facilitation of fraud and identity theft because of easier data aggregation.

The first concern is the psychological harm associated with the unauthorized disclosure of sensitive personal information. Health data is recognized as the most private personal information.⁶⁷ Eighty-one percent of the respondents in a Pew survey believed that health data was "sensitive," with fifty-five percent considering it "very sensitive."⁶⁸ Dr. Richard Harding, a former president of the American Psychiatric Association, believes that disclosures of medical information could cause personal disgrace as well as discrimination;⁶⁹ he notes that:

These disclosures can jeopardize our careers, our friendships, and even our marriages. And if such disclosures occur, there are truly few meaningful remedies. Seeking redress will simply lead to further dissemination of the highly private information that the patient wished to keep secret, nor can a financial settlement do much to compensate the individual for these highly personal losses. For all of these reasons, very tight restrictions on access as well as disclosure of medical records information is essential.⁷⁰

Similarly, in enacting relevant privacy rules, HHS stated that a health privacy breach could cause significant impacts beyond one's physical health, including alienation of family and friends and public humiliation.⁷¹ Granted,

⁶⁷ Zarsky, *supra* note 50, at 1012.

⁶⁸ Lori Andrews, *A New Privacy Paradigm in the Age of Apps*, 53 WAKE FOREST L. REV. 421, 460–61 n.291 (2018).

⁶⁹ *Financial Privacy: Hearing on H.R. 10 Before the Subcomm. On Financial Institutions and Consumer Credit of the Comm. on Banking and Financial Services*, 106th Cong. 100 (1999) (statement of Donald J. Palmisano, M.D., J.D., A.M.A.), <https://archive.org/details/financialprivacy00unit/page/n1> [<https://perma.cc/26Z6-4LDM>].

⁷⁰ *Id.*

⁷¹ See Paul Ohm, *Sensitive Information*, 88 S. CAL. L. REV. 1125, 1151 (2015).

people may have only traditional health data on their minds when responding to surveys and making comments and it is true that not all health data under this broadened Big Data definition carries the same level of sensitivity. Nevertheless, the underlying rationale is not limited to traditional health data. That is, people feel strongly about when, how, to whom, and to what extent their health data should be disclosed. They care about the resulting judgments made about their health problems and detriments to their reputations no matter if the information revealed is traditional health data or a result of combined health datasets.

In addition, because the resulting damages often extend beyond data subjects' expectations, misuse of health data in combination with Big Data analytics is probably even more psychologically harmful than misuse of health data in the traditional sense. What if the girl in the Target example above did not want her father to know about her pregnancy? She would of course be upset if her pregnancy was accidentally disclosed by a breach of her EHR, but she might be more distressed upon realizing that her purchasing data at Target turned out to be the information source.

Sharona Hoffman, a leading scholar in the areas of health information technology and civil rights, has noted another concern caused by Big Data analytics: the new health-related discrimination threats; employers, financial institutions, marketers, and educational institutions are all involved as potential stakeholders in this respect.⁷² Data subjects could pay more for their health insurance plans, be denied certain financial transactions, and be excluded from certain products just because of the correlations and inferences arising from misuse of health data in the context of Big Data.⁷³ The White House raised the same concern in 2014 and pointed out that Big Data analytics could result in discriminatory use of personal information, endangering the civil rights protections in various domains including the healthcare sector.⁷⁴ If treatment, eligibility, inclusion, or access are determined by data, not only will

⁷² See Sharona Hoffman, *Big Data's New Discrimination Threats: Amending the Americans with Disabilities Act to Cover Discrimination Based on Data-Driven Predictions of Future Disease*, in *BIG DATA, HEALTH LAW, AND BIOETHICS* 85, 85 (I. Glenn Cohen et al. eds., 2018) (noting that employers are keen to obtain prospective employees' medical information to determine whether they will develop serious illnesses for purposes of employment decisions; financial institutions are eager to collect individuals' health data to screen out applicants with a high risk of defaulting on loans because of medical difficulties; some educational institutions may be interested in applicants' health data to determine whether they are likely to have abbreviated careers and limited earnings because of medical challenges and become successful professors or otherwise bring honors to the institutions).

⁷³ See discussion *supra* Part II.B.

⁷⁴ See JOHN PODESTA ET AL., *BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES* 45–47 (2014), https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf [<https://perma.cc/96S3-7TR6>].

the well-established U.S. anti-discrimination values be significantly compromised, but also one's health, choices, life chances, and "general autonomy and human dignity"⁷⁵ may be undermined.

Worse, algorithmically imposed decisions about treatment or access may be based on correlations and inferences which are imperfect and can be easily misleading. Imagine you are studying issues related to multiple sclerosis (M.S.): You have conducted some online searches and subscribed to an online recommendation engine to look up physicians; then you receive an invite to a meeting of M.S. patients.⁷⁶ Why? Because some data handler has mistakenly profiled you as an M.S. patient and shared that profile with other marketers.⁷⁷ Wrong advertisements might only be annoying and superficially damaging, but if the recipient of an incorrect profile is an insurer or an employer with a discriminatory view against pre-existing conditions, the resulting harm could be more significant—you may lose a job opportunity or be denied a service based on a mistaken inference by a third-party, without your fault or knowledge.

Another emerging issue is Big Data's role in facilitating health related fraud and identity theft. Consider this example from 2007: A data broker named InfoUSA came up with a list of "Suffering Seniors" by aggregating health data about cancer and Alzheimer's disease, and sold it to telemarketers who targeted senior citizens.⁷⁸ Then the telemarketers raided those senior citizens' bank accounts by tricking them into revealing their bank information.⁷⁹ Even if the data transferred to perpetrators is insignificant, Big Data analytics makes it much easier for wrongdoers to devise similar lists and harm data subjects.

D. Under-regulated mHealth App Data

As one of the most significant groups of health data in the Big Data context, mHealth app data, unlike its counterparts in the traditional healthcare setting, currently receives no more protection than other consumer data. Because mHealth apps are mostly designed to fulfill health-related functions, such as treating diseases, managing chronic conditions, and promoting a

⁷⁵ Wolfie Christl, *How Companies Use Personal Data Against People* 17 (Working Paper by Cracked Labs, 2017), https://crackedlabs.org/dl/CrackedLabs_Christl_DataAgainstPeople.pdf [<https://perma.cc/WB7S-6YV9>].

⁷⁶ Natasha Singer, *When Your Data Wanders to Places You've Never Been*, N.Y. TIMES (Apr. 27, 2013), <https://www.nytimes.com/2013/04/28/technology/personal-data-takes-a-winding-path-into-marketers-hands.html> [<https://perma.cc/MJ7U-UMB8>].

⁷⁷ *Id.*

⁷⁸ Ashley Kuempel, *The Invisible Middlemen: A Critique and Call for Reform of the Data Broker Industry*, 36 NW. J. INT'L L. & BUS. 207, 221 (2016).

⁷⁹ *Id.*

healthy lifestyle, the bulk of data generated by mHealth apps is intrinsically medically significant or is very likely to become health-related given the health-related context these apps are designed and used in. Accordingly, mHealth app data is overall more medically revealing and sensitive than consumer data from other sources. Baby kicks recorded by a pregnancy tracker app, for example, will always be a more direct source of data compared with Big Data inferences like Target's use of purchasing data discussed above.

However, because no regulation has been promulgated to provide extra protection, mHealth app data is, as with other consumer data, completely open to decontextualized, discriminatory, or criminal uses—as discussed earlier. Like data generated by other consumer-facing mobile apps, mHealth app data suffers from excessive collection from data subjects and under-regulated transmissions among app developers, data brokers, and other interested players.

Without compulsory standards, the only potential limitation is the notice-and-consent mechanism embedded in mobile apps' self-imposed privacy policies, breach of which, as discussed in Part III, triggers FTC's enforcement. However, most of these privacy policies are so one-sided that app developers can modify their terms at will and leave consumers with a consent-or-abandon choice.⁸⁰ In addition, privacy policies are usually long and complicated, and many users do not read or understand them prior to agreeing to the terms.⁸¹

Margaret Jane Radin, a leading legal scholar focusing on legal issues in cyberspace and exploring freedom of choice in the information society, has noted that “free consent involves a knowing understanding of what one is doing in a context in which it is actually possible for one to do otherwise, and an affirmative action in doing something, rather than a merely passive acquiescence in accepting something.”⁸² Unfortunately, free consent is difficult to find in many consumer-facing apps including mHealth apps. Although questions remain about the best way to seek consumers' consent, be it the traditional notice-and-consent mechanism or the opt-in and opt-out approaches, the chosen method should allow users to make meaningful choices, instead of operating as a shield for potential privacy violations.

Moreover, some mHealth apps, together with many other consumer apps, do not even have privacy policies. It is reported that approximately twenty six percent of free mobile apps and forty percent of paid health mobile

⁸⁰ Fred H. Cate & Viktor Mayer-Schönberger, *Notice and Consent in a World of Big Data*, 3 INT'L DATA PRIVACY LAW 67, 67 (2013).

⁸¹ See e.g., J. Frazee et al., *mHealth and Unregulated Data: Is This Farewell to Patient Privacy*, 13 IND. HEALTH L. REV. 384, 407 (2016).

⁸² Margaret Jane Radin, *Humans, Computers, and Binding Commitment*, 75 IND. L.J. 1125, 1125–26 (1999).

apps do not have privacy policies.⁸³ Another survey indicates that only about thirty percent out of the six hundred most-used mHealth apps had privacy policies.⁸⁴ In addition, most data brokers, as the primary downstream players, do not set any contractual obligations with their data sources to ensure that they or their counterparties provide consumers with notice of information sharing with third-parties and an opportunity to opt out of such sharing.⁸⁵ In conclusion, the whole industrial chain is taking advantage of the current liberal regulatory environment.

What is more worrisome is that mHealth app data is more “popular” on the market, and therefore more liable to misuses. First, compared with other consumer apps, health and fitness apps sent sensitive data to more third-party domains.⁸⁶ FTC’s earlier study revealed that twelve mHealth apps and devices transmitted information to seventy-six different third parties without consumers’ knowledge, with eighteen third parties receiving device-specific identifiers, fourteen receiving consumer-specific identifiers, and twenty-two receiving other key health data.⁸⁷

Second, in contrast with other consumer data, mHealth app data is more likely to become a pricy “commodity” on the market: Its price is ten times that of other personal data on the data market,⁸⁸ and is fifty times that of credit card information on the black market.⁸⁹

Some may argue that mHealth app data does not necessarily have major sensitivities because data can be de-identified. This is largely unfounded. First and foremost, there are few proper regulations for the collection, transmission and transaction of mHealth app data, much less a statutory requirement for its de-identification. Second, even if a single set of mHealth app data is de-identified, re-identification is not a difficult task when analyzed in conjunction with other datasets, thanks again to Big Data analytics.⁹⁰ For example, the latest findings show that by matching daily step data collected by activity trackers, smartwatches, and smartphones to

⁸³ Stacy-Ann Elvy, *Commodifying Consumer Data in the Era of the Internet of Things*, 59 B.C. L. REV. 423, 439 (2018).

⁸⁴ Guadarrama, *supra* note 29, at 1016.

⁸⁵ FED. TRADE COMM’N, *supra* note 58, at 16.

⁸⁶ Jinyan Zang et al., *Who Knows What About Me? A Survey of Behind the Scenes Personal Data Sharing to Third Parties by Mobile Apps*, TECH. SCI. 1, 17 (2015), <https://techscience.org/a/2015103001/download.pdf> [<https://perma.cc/8TE3-R9PA>].

⁸⁷ Sarah Kellogg, *Every Breath You Take: Data Privacy and Your Wearable Fitness Device*, 72 J. MO. B. 76, 77 (2016).

⁸⁸ Lashbrook, *supra* note 2.

⁸⁹ Andrews, *supra* note 68, at 431.

⁹⁰ See Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1717–22 (2010).

demographic data, individuals can be re-identified.⁹¹ So it is possible for Facebook to re-identify you if it gathers your step data from your mHealth app, buys another set of data from another company, and then matches the two datasets.⁹²

In conclusion, compared with other consumer data, mHealth app data is generally more sensitive, more revealing, and more liable to Big Data misuse against the interests of data subjects and yet they are in an almost regulation-free zone.

III. CURRENT FEDERAL STATUTORY LANDSCAPE AND GAPS

Commentators have criticized the U.S. data protection legal system as “fragmented,”⁹³ “patchwork,”⁹⁴ and “haphazard.”⁹⁵ These criticisms are based on the fact that there is no unified, baseline protection statute for consumer data on the federal level. Under the existing legal framework, whether privacy problems can be addressed mainly depends on the industry at issue and the actors involved.⁹⁶

As discussed earlier, the primary federal regulators of mHealth apps are HHS, FDA, and FTC, but no single regulator has complete jurisdiction over mHealth app data. HIPAA only regulates “protected health information” (PHI) to the extent that it is held by HIPAA-defined covered entities and business associates, but many users and processors of mHealth app data do not meet HIPAA’s definitions of covered entities and business associates. Also, HIPAA does not regulate health data collection, despite excessive collection being a primary concern for mHealth app data. FDA, on the other hand, focuses its oversight on the efficacy and safety of mHealth apps, and does not regulate privacy issues. Lastly, FTC, relying on Section 5 of the Federal Trade Commission Act (FTC Act), is now an active regulator for mHealth app data, but its authority is limited, unclear, and liable to challenges.

⁹¹ *Artificial Intelligence Advances Threaten Privacy of Health Data*, EUREKALERT! (Jan. 3, 2019), https://www.eurekaalert.org/pub_releases/2019-01/uoc--aia010319.php [<https://perma.cc/UKP6-TEJV>].

⁹² Jessica Kim Cohen, *AI Can Re-Identify De-Identified Health Data, Study Finds*, BECKER’S HEALTH IT & CIO REP. (Jan. 3, 2019), <https://www.beckershospitalreview.com/artificial-intelligence/ai-can-re-identify-de-identified-health-data-study-finds.html> [<https://perma.cc/5H4V-EAQJ>].

⁹³ Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 587 (2014).

⁹⁴ Helm & Georgatos, *supra* note 15, at 140.

⁹⁵ Ohm, *supra* note 71, at 1140.

⁹⁶ Helm & Georgatos, *supra* note 15, at 140.

Because FDA is not currently engaged in privacy issues, the following Section will only discuss the regulatory approaches adopted by HIPAA and FTC, and their flaws.

A. HIPAA: A Closed, Downstream Approach

The most well-known federal statute for health data protection is HIPAA.⁹⁷ After several amendments and updates, the existing HIPAA rules (HIPAA Rules) are primarily composed of the Privacy Rule, the Security Rule, and the Breach Notification Rule, incorporating relevant requirements posed by other acts such as the Health Information Technology and Economic Clinical Health Act (HITECH Act).⁹⁸

Promulgated with several purposes such as combating health care fraud and improving access to insurance coverage,⁹⁹ HIPAA was initially silent on health data protection; instead, HHS was delegated to issue separate regulations to protect personal health information absent a Congressionally enacted comprehensive privacy legislation within three years of HIPAA's enactment.¹⁰⁰ Therefore, when Congress missed its deadline, HHS promulgated the Privacy Rule in 2002 and the Security Rule in 2003.¹⁰¹ The HITECH Act, enacted in 2009, strengthened HIPAA's protection by expanding the definition of business associate, extending application of the Privacy Rule and the Security Rule to business associates,¹⁰² and incorporating new breach notification requirements.¹⁰³ On January 25, 2013, by publishing the HIPAA Omnibus Rule, HHS further incorporated the relevant data

⁹⁷ There are some other statutes regulating use of health data, but they are more focused on prevention of health-based discrimination by certain specified actors, rather than regulating the flow of health data. Therefore, this Article does not examine these statutes. For example, the Genetic Information Nondiscrimination Act (GINA) of 2008 protects Americans from discrimination based on their genetic information by health insurers and employers. The Americans with Disabilities Act (ADA) prohibits discrimination against individuals with disabilities in areas of public life, such as jobs, schools, and transportation. *See* Genetic Information Nondiscrimination Act, Pub. L. No. 110-233, 122 Stat. 881–922 (2008); Americans with Disabilities Act of 1989, Pub. L. No. 101-336, 104 Stat. 327–378 (1989).

⁹⁸ U.S. DEP'T OF HEALTH & HUMAN SERVS., EXAMINING OVERSIGHT OF THE PRIVACY & SECURITY OF HEALTH DATA COLLECTED BY ENTITIES NOT REGULATED BY HIPAA 12 (2016), https://www.healthit.gov/sites/default/files/non-covered_entities_report_june_17_2016.pdf [<https://perma.cc/XB3P-D4XD>].

⁹⁹ Stacey A. Tovino, *The HIPAA Privacy Rule and the EU GDPR: Illustrative Comparisons*, 47 SETON HALL L. REV. 973, 975 (2017).

¹⁰⁰ *Id.* at 976.

¹⁰¹ STEPHEN S. WU, A GUIDE TO HIPAA SECURITY AND THE LAW 2 (2d ed. 2016).

¹⁰² Nicolas Terry, *Health Privacy is Difficult but Not Impossible in a Post-HIPAA Data-Driven World*, 146 CHEST 835, 836 (2014).

¹⁰³ *Id.*

protection provisions of the HITECH Act and the Genetic Information Nondiscrimination Act (GINA) into the HIPAA Rules.¹⁰⁴

The HIPAA Rules regulate the flow of PHI within a closed community among individuals, covered entities, and business associates. Some definitions are needed to better understand the HIPAA mechanism.

As used in the HIPAA Rules, PHI means information relating to an individual's health conditions, health care received, or health care related payment that (1) is created or received by covered entities or business associates, and (2) identifies or can be reasonably used to identify such individual.¹⁰⁵ De-identified health information is thus not protected.¹⁰⁶ "Covered entity" includes a health plan, a health care clearinghouse, or a health care provider who transmits electronic health information in relation to a HIPAA-covered transaction;¹⁰⁷ "business associate" refers to: (1) a person providing PHI transmission services to a covered entity and requiring routine access to PHI, (2) a person offering personal health records to individuals on behalf of a covered entity, and (3) a subcontractor that deals with PHI on behalf of the business associate.¹⁰⁸

To the extent that PHI is held by a covered entity or a business associate, the HIPAA Rules provide relatively robust protections in the sense that data subjects can obtain and correct their PHI; can know, designate, and limit their PHI's recipients; and can be informed of and complain about breaches of PHI.¹⁰⁹ Specifically, data subjects' written authorizations from individuals must be obtained before use or disclosure of their PHI,¹¹⁰ except that PHI may be used or disclosed without individuals' authorization¹¹¹ for

¹⁰⁴ Modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules Under the Health Information Technology for Economic and Clinical Health Act and the Genetic Information Nondiscrimination Act; Other Modifications to the HIPAA Rules, 78 Fed. Reg. 5566 (proposed Jan. 25, 2013) (to be codified at 45 C.F.R. § 160 and 164).

¹⁰⁵ 45 C.F.R. § 160.103 (2019).

¹⁰⁶ U.S. DEP'T OF HEALTH & HUMAN SERVS., SUMMARY OF THE HIPAA PRIVACY RULE 4 (2013), <https://www.hhs.gov/sites/default/files/privacysummary.pdf> [<https://perma.cc/A7CL-F844>].

¹⁰⁷ 45 C.F.R. § 160.103 (2019).

¹⁰⁸ *Id.*

¹⁰⁹ U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 98, at 14.

¹¹⁰ *Id.*

¹¹¹ U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 98, at 14; *but see* INST. OF MED., CAPTURING SOCIAL AND BEHAVIORAL DOMAINS IN ELECTRONIC HEALTH RECORDS: PHASE 2 325–326 (2014) (noting that psychotherapy notes recorded in any medium by a mental health professional documenting or analyzing the contents of conversation during a private counseling session or a group, joint or family counseling session are separated from the rest of the individual's medical record. A covered entity is required to obtain the patient's express written authorization for any use or disclosure of psychotherapy notes, except for the following: (1) the treatment uses by the originator of the notes; (2) use or disclosure in mental

HIPAA-permitted purposes such as health care operations¹¹² and public health activities.¹¹³

Also, HIPAA embraces a principle of “minimum necessary” disclosure.¹¹⁴ Unless under specified circumstances, a covered entity or business associate is required to “make reasonable efforts to limit protected health information to the minimum necessary to accomplish the intended purpose of the use, disclosure, or request.”¹¹⁵ These robust protections do not function well when applied to mHealth app data for two reasons.

First, mHealth app data—however sensitive—is not subject to HIPAA’s governance because mHealth apps are consumer-grade products not typically offered by a covered entity or business associate.¹¹⁶ The fundamental issue under HIPAA is that the custodians of the data matter more than the data itself.¹¹⁷ That is, whether an individual’s health data is protected by HIPAA depends on *who* is holding the data, rather than *what* the data is.¹¹⁸ HIPAA focuses on how health data should be channeled, instead of how the private interests attached to health data should be safeguarded.¹¹⁹

This custodian-centric approach has led to two unreasonable scenarios. First, health data of the same nature may receive different treatment as a result of the different settings where it is processed. For instance, a patient’s glucose level measured at a brick-and-mortar healthcare provider, such as a clinic or a

health professional training programs; (3) use by the covered entity to defend itself in a lawsuit brought by the individual who is the subject of the notes; (4) disclosures required by law; (5) uses related to oversight of the originator of the notes; (6) disclosures to coroners and examiners to help determine cause of death; and (7) disclosures to prevent an imminent threat to health or safety).

¹¹² 45 C.F.R. § 164.501 (2019) (“Health care operations” include a covered entity’s various activities necessary to run its business and to support the core functions of treatment and payment, such as quality assessment and improvement activities, internal performance evaluation, medical review, legal services, auditing functions, business planning and development, business management and general administrative activities).

¹¹³ 45 C.F.R. § 164.512 (2019) (A covered entity may use or disclose PHI for the purpose of public health activities. For example, a covered entity may disclose PHI to a public health authority for purpose of preventing or controlling disease, injury, or disability and receiving reports of child abuse or neglect or disclose to FDA-governed persons for the purpose of activities related to the quality, safety or effectiveness of FDA-regulated products or activities).

¹¹⁴ 45 C.F.R. § 164.502(b) (2019).

¹¹⁵ *Id.* (For example, the minimum necessary standard does not apply to disclosures by a health care provider for treatment purposes, or disclosures permitted or compelled by applicable law).

¹¹⁶ Guadarrama, *supra* note 29, at 1005.

¹¹⁷ See Terry, *Regulatory Disruption*, *supra* note 49, at 164.

¹¹⁸ U.S. DEP’T OF HEALTH & HUMAN SERVS., *supra* note 98, at 12.

¹¹⁹ Theodore Rostow, *What Happens When an Acquaintance Buys Your Data: A New Privacy Harm in the Age of Data Brokers*, 34 YALE J. REG. 667, 677 (2017).

hospital, is protected by HIPAA. However, if the patient enters the same dataset into an mHealth app at home, the data is not protected under HIPAA.

Second, HIPAA-protected health data may lose its protection solely due to a change of hands. If consumers use a third-party health management app that is not offered or otherwise in association with their healthcare providers and then add certain health data from their HIPAA-protected EHR to that app, that health data is now outside the reach of HIPAA.¹²⁰ This counterintuitive outcome is related to HIPAA's initial legislative intention. Initially, HIPAA was promulgated to standardize the electronic exchange of health data involved in financial and administrative transactions covered by HIPAA. Privacy protection was not initially included within the statute.¹²¹ Consequently, the focus of the HIPAA Rules on the doctor-hospital-insurer ecosystem is natural. In a traditional healthcare system, doctors, hospitals and insurers—as the data holders captured by the HIPAA—are almost the only players. It seems sufficient for lawmakers to only regulate conducts of these players. Therefore, the regulatory gap before the emergence of mHealth products was not that significant. Currently, mHealth technology is reshaping the healthcare sector and has presented new issues as to (1) under what circumstances and for what purposes health data is collected, transmitted and processed, and (2) who has the opportunity and right to touch and control health data. HIPAA's custodian-centric approach needs to be updated—if not abandoned—in order to respond to these new issues.

HIPAA's second flaw is that it only regulates the flow of health data, not its collection.¹²² Nicolas P. Terry, a widely recognized leading academic in the field of health information and technology laws, describes this privacy protection approach as a “downstream” model, which only seeks to contain the collected data within the healthcare system by prohibiting its transmission to non-health-care parties, but imposes no limitation on data collection at the outset.¹²³ Terry traces this to the traditional culture of medicine, which favors collecting as much information as possible.¹²⁴ The rationale is that universal collection and free flow of health data could maximize patient health and achieve more public health goals.¹²⁵ This suggests a lag in regulatory rationale in the face of new information technologies.

¹²⁰ Timothy Newman & Jennifer Kreick, *The Impact of HIPAA (and Other Federal Law) on Wearable Technology*, 18 SMU SCI. & TECH. L. REV. 429, 446 (2015).

¹²¹ WU, *supra* note 101, at 2.

¹²² Nicolas Terry, *Health Privacy is Difficult but Not Impossible in a Post-HIPAA Data-Driven World*, 146 CHEST 835, 837 (2014).

¹²³ *Id.*

¹²⁴ Terry, *supra* note 49, at 165.

¹²⁵ Terry, *supra* note 57, at 68.

Data collection may not be a major concern if healthcare is only conducted by traditional healthcare providers in a “simple health-care data exchange scenario”¹²⁶ that is already supported by confidentiality rules in exchange for more effective treatment.¹²⁷ A patient sitting in front of a physician knows exactly what information is being collected and does not need to worry about unnecessary data collection. However, data collection becomes a problem when data is collected by mHealth apps. mHealth apps have broadened the scope of data collection, and users are often in the dark about what data is being collected, who is collecting it, and how it is being used. An mHealth app could—if turned on by a person and left without interruption—collect an abundance of data and “achieve 24/7 monitoring in order to create a digital doppelganger of the person.”¹²⁸ The scope of data collection further expands when an app goes beyond its proclaimed purpose and unnecessarily collects users’ other data such as geolocation, contact lists, or search history—as is often the case. This over-collection concern would not be addressed by the current HIPAA Rules even if those rules did cover mHealth app data because the HIPAA Rules have no complementary limit on data collection despite its existing requirement for minimization of data sharing.

B. FTC’s Section 5 Power: Only If You Break Your Promise

Because the HIPAA Rules cannot capture most mHealth app data, and FDA is hands-off to protection of mHealth app data, FTC is becoming a leading regulator in this field.¹²⁹ HHS acknowledges that the FTC Act¹³⁰ is currently the primary federal statute regulating health data not covered by HIPAA.¹³¹

FTC’s authority to regulate mHealth app data stems from its general power to protect consumer privacy under Section 5 of the FTC Act.¹³² FTC’s Section 5 power extends to persons, partnerships, or corporations, except banks, savings and loan institutions, federal credit unions, common carriers, air carriers, and packers that are subject to specialized laws and regulations.¹³³ Therefore, all data handlers, whether covered by HIPAA or not, are subject to Section 5 as long as they do not fall into one of those exceptional categories.

¹²⁶ Terry, *supra* note 49, at 165.

¹²⁷ *Id.*

¹²⁸ Andrews, *supra* note 68, at 426.

¹²⁹ Hartzog & Solove, *supra* note 16, at 2267.

¹³⁰ Federal Trade Commission Act, 15 U.S.C. §§ 41–58 (2018).

¹³¹ U.S. DEP’T OF HEALTH & HUMAN SERVS., *supra* note 98, at 17.

¹³² Solove & Hartzog, *supra* note 93, at 604.

¹³³ 15 U.S.C. § 45(a) (2018).

Instead of focusing on data's intrinsic characteristics or data protection,¹³⁴ FTC's Section 5 authority prevents an entity "from using unfair methods of competition in or affecting commerce and unfair or deceptive acts or practices in or affecting commerce."¹³⁵ To justify its power to enforce privacy protection, FTC reasons that false or misleading promises about privacy and security or inadequate security measures are likely to injure consumers, and should therefore be covered by Section 5.¹³⁶

According to FTC, unauthorized access to information by an app in violation of its app's own privacy policies is "deceptive,"¹³⁷ and an app's failure to maintain adequate data security is simply "unfair", regardless of its own data security promise.¹³⁸

In enforcing Section 5, FTC has developed some general data protection principles, which include (1) adhering to promised privacy practices; (2) informing data subjects of uses and disclosures of material personal information; (3) notifying data subjects of data sharing with third parties outside the direct consumer relationship; (4) implementing reasonable and appropriate security measures; and (5) safeguarding private information based on the type of information and the risk presented to consumers.¹³⁹

If FTC reasonably believes that any person is engaging in any unfair or deceptive act or practice, FTC may initiate a hearing upon serving on the person a complaint and a notice of a hearing;¹⁴⁰ if the person complained of fails to justify its act or practice at the hearing, FTC may then make a written report and issue a cease and desist order.¹⁴¹ The person receiving a cease and desist order may petition for a review of that order by a competent court, and the court may then affirm, modify, or setting aside FTC's order, and enforce the same to the extent that such order is affirmed.¹⁴²

However, FTC's Section 5 authority is limited. For a deceptiveness prong claim, FTC follows a "broken promises" approach.¹⁴³ That is, only when an app developer makes false or misleading claims about its privacy or

¹³⁴ Terry & Gunter, *supra* note 47, at 139.

¹³⁵ Helm & Georgatos, *supra* note 15, at 159.

¹³⁶ Guadarrama, *supra* note 29, at 1011.

¹³⁷ Helm & Georgatos, *supra* note 15, at 160.

¹³⁸ Hartzog & Solove, *supra* note 16, at 2275.

¹³⁹ MAXIMUS FED. SERVS., NON-HIPAA COVERED ENTITIES: PRIVACY AND SECURITY POLICIES AND PRACTICES OF PHR VENDORS AND RELATED ENTITIES REPORT 3–4 (2012), https://www.healthit.gov/sites/default/files/maximus_report_012816.pdf [<https://perma.cc/Y73U-RXJK>].

¹⁴⁰ Federal Trade Commission Act, 15 U.S.C. § 45(b) (2018).

¹⁴¹ *Id.*

¹⁴² 15 U.S.C. § 45(c) (2018).

¹⁴³ ARI EZRA WALDMAN, PRIVACY AS TRUST: INFORMATION PRIVACY FOR AN INFORMATION AGE 82 (2018).

data security procedures can FTC trigger an enforcement action. Because there is currently no compulsory standard or robust self-regulation regime for privacy policies and security measures in the mHealth app industry, app developers can decide how they draft their own privacy policies and formulate security procedures. If they choose to set a lower bar, FTC has no power to impose a higher one.

Compared to the limited—but straightforward—“broken promises” approach, claims under the unfairness-prong appear more challenging because the standard of proof is high. FTC must prove that (1) the practice in question has caused or is likely cause significant harms to consumers, (2) such harms cannot be reasonably avoided by customers, and (3) such harms are not outweighed by the resulting benefits.¹⁴⁴

FTC has used this three-part test to take enforcement actions against companies that have poor data security practices regarding health data,¹⁴⁵ but not always with success.¹⁴⁶ The U.S. Court of Appeals for the Eleventh Circuit recently struck down FTC’s cease and desist order against LabMD, a cancer detection facility holding sensitive health data.¹⁴⁷ Because one of LabMD’s employees used a file-sharing app and inadvertently made available health data to third parties, a hacker company managed to access sensitive information of about 9,000 patients.¹⁴⁸ FTC investigated and asserted in its complaint that LabMD’s data security practices were “unreasonably lax,” making them unfair under Section 5.¹⁴⁹

The court found that FTC failed to cite explicitly the source of the standard of unfairness and failed to specify the unfair acts or practices engaged in by LabMD.¹⁵⁰ Although fact-specific, the court’s holding indicates that compared with the deceptiveness prong claim, the standard of proof of the unfairness prong claim is higher and requires FTC to stretch its interpretation of Section 5 even further to justify its enforcement actions.

As the leading, catchall regulator of privacy-related practices of mobile apps, FTC seems to be in a good position to regulate mHealth app data. However, because the FTC Act does not explicitly grant FTC the authority to

¹⁴⁴ 15 U.S.C. § 45(n) (2018).

¹⁴⁵ U.S. DEP’T OF HEALTH & HUMAN SERVS., *supra* note 98, at 18.

¹⁴⁶ See *FTC Rebuked in LabMD Case: What’s Next for Data Security?*, WILEY REIN LLP (Jun. 7, 2018), https://www.wileyrein.com/newsroom-articles-FTC_Rebuked_in_LabMD_Case_Whats_Next_for_Data_Security.html [<https://perma.cc/5WU6-MY3U>].

¹⁴⁷ *Id.*

¹⁴⁸ *Id.*

¹⁴⁹ *Id.*

¹⁵⁰ *LabMD, Inc. v. Fed. Trade Comm’n*, 894 F.3d 1221, 1230–31 (11th Cir. 2018).

protect privacy,¹⁵¹ the effect of FTC's regulation is limited. The deceptiveness prong ultimately counts on industry players' self-discipline, and the unfairness prong relies on FTC's stretched interpretation of "fairness," neither of which are guaranteed by any rulemaking power or enforceable, compulsory standards. Unless these limitations are overcome, FTC is not ready to completely address the privacy concerns posed by mHealth app data.

IV. GAP-FILLING PROPOSALS

A. Necessity of Governmental Action to Address Regulatory Gaps

The next inquiries are: Do these regulatory gaps need to be addressed? If so, how? On the general topic of consumer privacy protection, views differ as to whether privacy protection could best be achieved through agency regulation or self-regulation.¹⁵² There are two major arguments supporting the self-regulation approach. First, regulation for privacy protection is not worth discussing, as people either no longer have expectations of privacy or are willing to exchange privacy for more significant benefits. Second, self-regulation is sufficient for privacy protection and there is no need for any government-involved regulation. This Part will examine and counter these two arguments in turn.

1. *Privacy in the Digital Age: Shifted Expectations or Trade-off?*

Many doubt the intrinsic value of privacy in the digital age,¹⁵³ and ask whether regulation of any form is needed at all. This Section argues that the expectations of privacy have shifted over time, but this shift has not led, and should not lead, to the demise of privacy in the digital age.

Those opposed to privacy rights might base their propositions on the assumption that privacy is somewhat equal to secrecy and ends when disclosure is made. Under their reasoning, because data collection, disclosure, aggregation, and analysis are inevitable in the digital age, there is no way to

¹⁵¹ U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-13-663, INFORMATION RESELLERS: CONSUMER PRIVACY FRAMEWORK NEEDS TO REFLECT CHANGES IN TECHNOLOGY AND THE MARKETPLACE 11 (2013), <https://www.gao.gov/assets/660/658151.pdf> [<https://perma.cc/FV5R-G59C>].

¹⁵² *Id.* at 16.

¹⁵³ See, e.g., Peter Friedman, *Is Privacy Dead in The Digital Age? What To Do About It: Part I*, MEDIAPOST (July 26, 2018), <https://www.mediapost.com/publications/article/322686/is-privacy-dead-in-the-digital-age-what-to-do-abo.html> [<https://perma.cc/E38D-Y63W>].

hide information forever, and the protection of privacy as a right is therefore questionable.¹⁵⁴

Alternatively, privacy opponents might argue that although privacy remains of significant value, there is, at least in the digital world, a trade-off about privacy.¹⁵⁵ Consumers are willing to sacrifice their data and privacy in exchange for considerable benefits,¹⁵⁶ such as increased economic efficiency, improved security, better personalization of services, increased availability of information, innovative platforms for communication,¹⁵⁷ and free services. In other words, data subjects are willing to accept, and have accepted, the costs incurred due to the technological advancements, and do not care about privacy as much as advocates and regulators believe. For example, in a survey of over 4,000 individuals about online purchases, sixty-five percent of the respondents reported that they seldom read privacy policies.¹⁵⁸ This seems to show consumers' indifference to privacy. Those opposed further reason that, now that consumers' expectations and notions of privacy have changed, there is no need for strict privacy control mandated by law.¹⁵⁹ The next Section provides some theoretical and empirical counterarguments to this anti-privacy perspective.

a. Shifted Expectations of Privacy

The doubters of privacy rights miss the mark because they regard privacy as a static notion. They have failed to consider another general proposition from the famous article by Samuel Warren and Louis Brandeis: "Political, social and economic changes entail the recognition of new rights."¹⁶⁰ The right to privacy, like all other rights, is ever-changing, but has never faded. At first, we had privacy in physical space, and then privacy relating to making choices, and now, information privacy¹⁶¹ as a right to

¹⁵⁴ Art Caplan, *Why Privacy Must Die*, HEALTH CARE BLOG (Dec. 19, 2016), <https://thehealthcareblog.com/blog/2016/12/19/goodbye-privacy-we-hardly-knew-ye/> [<https://perma.cc/KX7P-UZLD>].

¹⁵⁵ See, e.g., Jon Reily, *Privacy or Convenience: What's the Tradeoff?*, PUBLICIS SAPIENT, <https://www.publicissapient.com/insights/privacy-or-convenience--what-s-the-tradeoff> [<https://perma.cc/T4VV-2DN9>].

¹⁵⁶ Frazee et al., *supra* note 81, at 409.

¹⁵⁷ Robert H. Sloan & Richard Warner, *Beyond Notice and Choice: Privacy, Norms, and Consent*, 14 J. HIGH TECH. L. 370, 385 (2014).

¹⁵⁸ Frazee et al., *supra* note 81, at 393 n.46.

¹⁵⁹ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 151, at 27.

¹⁶⁰ Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 193 (1890).

¹⁶¹ See INST. OF MED., HEALTH DATA IN THE INFORMATION AGE: USE, DISCLOSURE, AND PRIVACY 143–54 (1994), (providing a review of the concept of privacy in America, and noting that the development of the concept of privacy in the U.S. encompasses three clusters of ideas:

privacy in the sharing of personal information.¹⁶² Best depicted by the eight “Fair Information Practice Principles” (FIPPs) adopted in the 1970s, the contents of informational privacy include openness, individual access, individual participation, collection limitation, use limitation, disclosure limitation, information management, and accountability.¹⁶³

Entering the digital age, many commentators have posited structured elaborations of the right to privacy from various perspectives. Alan Westin, a pioneering legal scholar and political scientist in the field of consumer data privacy and data protection, associates privacy with one’s control over personal information and defines privacy as “the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.”¹⁶⁴

Jack Balkin, a prominent legal scholar in the field of Constitutional law and new information technologies, approaches the right to privacy with a focus on relationship.¹⁶⁵ He has put forward the concept of an “information

(1) decisional privacy, which embodies autonomy interests concerning the exercise of fundamental constitutional liberties with respect to private behavior, such as decisions relating to marriage, procreation, contraception, family relationships, and child-rearing; (2) spatial privacy, which protects people against surveillance or intrusion such as unlawful searches of one’s home or person and unauthorized wiretapping; and (3) informational privacy, which represents an individual’s control over the dissemination, use, and access by others of information that relates to himself or herself); *see also* NAT’L RESEARCH COUNCIL, ENGAGING PRIVACY AND INFORMATION TECHNOLOGY IN A DIGITAL AGE 58–87 (2007) (going through the long intellectual history of the notion of privacy in the United States from the perspectives of philosophy, economics, and sociology).

¹⁶² Elizabeth A. Rowe, *Sharing Data*, 104 IOWA L. REV. 287, 301 (2018), <https://ilr.law.uiowa.edu/print/volume-103-issue-6/sharing-data/> [<https://perma.cc/BP9K-E6VN>].

¹⁶³ Griffin Drake, *Navigating the Atlantic: Understanding EU Data Privacy Compliance Amidst a Sea of Uncertainty*, 91 S. CAL. L. REV. 163, 167 (2017), <https://southern.californialawreview.com/2017/11/01/navigating-the-atlantic-understanding-eu-data-privacy-compliance-amidst-a-sea-of-uncertainty-note-by-griffin-drake/> [<https://perma.cc/4VN6-RBWK>].

¹⁶⁴ *See* NAT’L RESEARCH COUNCIL, ENGAGING PRIVACY AND INFORMATION TECHNOLOGY IN A DIGITAL AGE 59 (2007).

¹⁶⁵ *See* Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C.D. L. REV. 1183, 1234 (2016) (arguing that new classes of digital information fiduciaries should be created to require platform owners to respect the free speech and privacy of end users in return for special legal status and benefits); *see also* Benjamin Wittes & Wells C. Bennett, *Database and a Trusteeship Model of Consumer Protection in the Big Data Era*, BROOKINGS: GOVERNANCE STUD. (2014), https://www.brookings.edu/wp-content/uploads/2016/06/Wittes-and-Bennett_Database.pdf (proposing a negative right to privacy based on “trusteeship” and expectations of no “database,” which is defined as “the malicious, reckless, negligent, or unjustified handling, collection, or use of a person’s data in a fashion adverse to that person’s interests and in the absence of that person’s knowing consent”) [<https://perma.cc/SCP2-49VL>].

fiduciary” to define certain non-contractual, relationship-based duties of data service providers towards data subjects.¹⁶⁶ Similar to other professional relationships, such as the lawyer-client relationship and the physician-patient relationship, these fiduciary duties would require data service providers to, even absent an express contractual promise, act in the best interest of data subjects.¹⁶⁷ Specifically, this means information fiduciaries could not use information obtained in the course of the relationship to the disadvantage of data subjects or to create conflicts of interest with data subjects.¹⁶⁸

Ari Ezra Waldman, a leading thought leader on online privacy and safety, frames privacy as a trust-based social norm and argues that “we should conceptualize information privacy in terms of relationships of trust and leverage law to protect those relationship.”¹⁶⁹

Helen Nissenbaum, who has developed the “contextual integrity” theory, associates adequate privacy protection with norms of specific contexts.¹⁷⁰ In her view, the gathering, dissemination, and use of data should be context-specific, and the governing norms of distribution within different contexts should be obeyed.¹⁷¹

European scholars go even further to propose a “legitimate interest” argument. They argue that privacy is adequately protected if a legitimate interest is served in favor of data subjects in all stages of the life cycle of personal data, including collection, use, further use, and destruction.¹⁷²

¹⁶⁶ Balkin, *supra* note 165, at 1234.

¹⁶⁷ *Id.*

¹⁶⁸ *See id.*

¹⁶⁹ WALDMAN, *supra* note 143, at 4–5; *see also* Eugenio Mantovani et al., *Towards a Code of Conduct on Privacy for mHealth to Foster Trust Amongst Users of Mobile Health Applications*, DATA PROTECTION AND PRIVACY: (IN)VISIBILITIES AND INFRASTRUCTURES 81, 84 (Ronald Leenes et al. eds., 2017) (noting that privacy is based on the trust that the other party will behave responsibly and will not attempt to exploit the vulnerabilities of the user).

¹⁷⁰ *See* Helen Nissenbaum, *Privacy as Contextual Integrity*, 79 WASH. L. REV. 119, 158 (2004); *see also* Helen Nissenbaum, *PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE* (2010) (proposing a model of informational privacy where contextual integrity is adopted as the benchmark for prescribing restrictions on the collection, use, and dissemination of information, variables of which include the nature of the context, the nature of the information in relation to that context, the roles of agents receiving information and their relationships with information subjects, and how the information is shared and further disseminated).

¹⁷¹ *Id.*

¹⁷² *See generally* Lokke Moerel & Corien Prins, *Privacy for the Homo Digitalis: Proposal for a New Regulatory Framework for Data Protection in the Light of Big Data and the Internet of Things*, SSRN ELEC. J. (2016), <https://papers.ssrn.com/abstract=2784123> (asking whether the existing purpose-centric EU legal regime is effective and legitimate in a future driven by data, and proposing a legitimate interest test where regulations will not determine the conditions for which data processing is allowed, but will only set standards in cases where data are used in a manner that should be qualified as abuse) [<https://perma.cc/D4MC-4Z7S>].

Together, these different theories of privacy indicate that scholarly and individual perceptions of privacy lie on a spectrum. This Section does not attempt to argue that we should place mHealth app data on any particular point on that spectrum. Instead, it argues that these doctrinal efforts have reflected the shifted expectations of privacy and have defeated the argument that the right to privacy has been vitiated.

b. Trade-off as a Fallacy

The privacy trade-off argument proves to be a fallacy as well.¹⁷³ Most adults, as reported by Harris Poll, are willing to allow people to access and use their personal information only to the extent that they know the reasons for data use, see the tangible benefits for doing so, and believe care is taken to prevent misuse.¹⁷⁴ A 2015 Nielsen survey revealed that fifty-three percent of respondents were concerned that their data might be shared without their knowledge.¹⁷⁵ Although these surveys seem to be conducted on a debatable assumption that personal data is property owned and controlled by data subjects,¹⁷⁶ it is at least fair to conclude from these surveys that, in principle, people are both cautious about how their personal data is transmitted and used and afraid of the risks and harms associated with the unexpected misuse of personal data.

Why, however, do some surveys suggest that users are indifferent to their personal data? This gap, termed by some as the “privacy paradox,” results from several factors, including unawareness, resignation, psychological distortions, and overestimation of existing protections.¹⁷⁷

Both unawareness and resignation arise out of information asymmetry. As discussed earlier, due to the complexity of data processing technologies

¹⁷³ See generally JOSEPH TUROW ET AL., *THE TRADEOFF FALLACY: HOW MARKETERS ARE MISREPRESENTING AMERICAN CONSUMERS AND OPENING THEM UP TO EXPLOITATION* (Annenberg School for Communication ed., 2015) (going beyond the cost-benefit analyses and discussing the issues such as resignation and information asymmetry).

¹⁷⁴ Sloan & Warner, *supra* note 157, at 384–85.

¹⁷⁵ See Elizabeth A. Brown, *The Fitbit Fault Line: Two Proposals to Protect Health and Fitness Data at Work*, 16 *YALE J. HEALTH POL'Y L. & ETHICS* 1, 10 (2016).

¹⁷⁶ See, e.g., I. Glenn Cohen, *Is There a Duty to Share Healthcare Data?*, in *BIG DATA, HEALTH LAW, & BIOETHICS* 209 (I. Glenn Cohen et al., eds., 2018); Jorge L. Contreras & Francisca Nordfalk, *Liability (and) Rules for Health Information*, 29 *HEALTH MATRIX* 179 (2019); Mark Rodwin, *Patient Data: Property, Privacy & the Public Interest*, 36 *AM. J. L. & MED.* 586 (2010).

¹⁷⁷ See e.g., Sonja Utz & Nicole C. Krämer, *The Privacy Paradox on Social Network Sites Revisited: The Role of Individual Characteristics and Group Norms*, 3 *CYBERPSYCHOLOGY: J. OF PSYCHOL. RES. ON CYBERSPACE* (2009), <https://cyberpsychology.eu/article/view/4223> [<https://perma.cc/7V6P-LEXQ>].

and unfriendly, take-it-or-leave-it style privacy policies,¹⁷⁸ users know little about how privacy policies apply to their personal data¹⁷⁹ or the alternatives available to them. As a result, users may feel forced to sacrifice privacy¹⁸⁰ and resigned to giving up control of their data.¹⁸¹ Back in the 1990s, Jeffrey Rothfeder perfectly described the feeling of powerlessness in the face of privacy intrusion:

Increasingly, people are at the whim of not only pressure groups, but also large organizations - direct marketers, the credit bureaus, the government, and the entire information economy - that view individuals as nothing but lifeless data floating like microscopic entities in vast electronic chambers, data that exists [sic] to be captured, examined, collated, and sold, regardless of the individual's desire to choose what should be concealed and what should be made public.¹⁸²

Information asymmetry has limited users' ability to make choices and take actions in accordance with their true desire to protect privacy.¹⁸³

Furthermore, as indicated by research in social psychology and behavioral economics, even having all the information necessary for an informed choice, one would sometimes end up behaving against better judgment due to the plight of immediate gratification.¹⁸⁴ As long as the privacy threat is not imminent, it is easier for most people to accept a default choice already made for them rather than to make a different choice, even if the default is less advantageous.¹⁸⁵

¹⁷⁸ Sloan & Warner, *supra* note 157, at 400.

¹⁷⁹ Nicolas A. Ozer, *Putting Online Privacy above the Fold: Building a Social Movement and Creating Corporate Change*, 36 N.Y.U. REV. L. & SOC. CHANGE 215, 224 (2012).

¹⁸⁰ Patrick Myers, *Protecting Personal Information: Achieving a Balance between User Privacy and Behavioral Targeting*, 49 U. MICH. J. L. REFORM 717, 731 (2016).

¹⁸¹ TUROW ET AL., *supra* note 173, at 3.

¹⁸² INST. OF MED., *supra* note 161, at 138 (quoting JEFFERY ROTHFEDER, *PRIVACY FOR SALE* (1992)).

¹⁸³ TUROW ET AL., *supra* note 173, at 3.

¹⁸⁴ Alessandro Acquisti, *Privacy in Electronic Commerce and the Economics of Immediate Gratification*, in *PROCEEDINGS OF THE 5TH ACM CONFERENCE ON ELECTRONIC COMMERCE* 21, 24 (2004), <https://www.heinz.cmu.edu/~acquisti/papers/privacy-gratification.pdf> (stating the plight of immediate gratification when users face privacy sensitive decisions, they hardly have all information necessary for an informed choice; however, even if they had, they would be likely unable to process it and even if they could process it, they may still end behaving against our own better judgment) [<https://perma.cc/3AYL-EAE5>].

¹⁸⁵ NAT'L RESEARCH COUNCIL, *supra* note 164, at 76.

The last, ironic reason for this paradox is that users often overestimate the protections available to them.¹⁸⁶ Some assume that a privacy policy is sufficient. For example, sixty-five percent of the responding users reported believing that having a privacy policy meant that the service provider would not share their information with others without their permission.¹⁸⁷ Others believed that if they paid a fee, their privacy protection would be guaranteed.¹⁸⁸ However, paid mHealth apps are found to be no better than free apps in terms of data collection and sharing.¹⁸⁹

Some may further question that if people really value privacy, why are products with enhanced privacy protections not dominating the market. The quick answer is that the privacy paradox has in turn hindered the development of a healthy market and rendered the market test ineffective. Consumers are simply stuck in this privacy paradox. Without sufficient demands for privacy-protecting goods and services, the market share of such products remains too small to make any impact.¹⁹⁰

To sum up, there are many doctrinal efforts and empirical studies around privacy going on. It can be concluded that individuals' expectations of privacy remain, but information asymmetry and psychological limitations have distorted consumers' true desire and hindered the formation of a healthy market. That is exactly why we need regulation to correct the distortion.

2. *The Illusion of Successful Self-Regulation*

The opponents' second major argument challenges the necessity of any governmental involvement in the regulation. Industrial players insist that the industry will work out on its own over time, and the government should not

¹⁸⁶ TUROW ET AL., *supra* note 173, at 4.

¹⁸⁷ *Id.*

¹⁸⁸ Frazee et al., *supra* note 81, at 410.

¹⁸⁹ *Id.*

¹⁹⁰ NAT'L RESEARCH COUNCIL, *supra* note 164, at 76.

interfere in the self-regulation and stifle innovation.¹⁹¹ The general marketing and information reseller industries argue that the regulatory gaps in consumer privacy protections are not that significant¹⁹² and flexible industry self-regulation should be able to adapt to rapid changes in technology and meet consumers' expectations.¹⁹³ Relatedly, health IT players claim that any well-meant new set of regulations may "have unintended consequences or lead to a regulatory land grab."¹⁹⁴

This "let the market play out" stance is generally in line with the U.S. government's long-standing desire to maintain a free market economy and a limited government,¹⁹⁵ and has long prevailed in guiding the U.S.'s privacy policies. For example, the Clinton administration advocated for industrial self-regulation in privacy protection as opposed to governmental regulation.¹⁹⁶ FTC also supported self-regulation as an alternative to "baseline legislation when it comes to privacy protection."¹⁹⁷

While the role of self-regulation in general consumer privacy regulation is beyond the scope of this Article, this Article argues that at least in terms of regulating mHealth app data, a self-imposed regulatory approach is unlikely to be successful if that self-regulatory mechanism only involves trade associations, third-party organizations, and companies within the industry,¹⁹⁸ and is not backed by any governmental agency.

¹⁹¹ See U.S. GOV'T ACCOUNTABILITY OFF., *supra* note 151, at 29–30 (2013) ("industry self-regulation was flexible and could adapt to rapid changes in technology and consumer expectations, whereas legislation and government regulation could be inflexible and quickly become outdated in an era of rapidly evolving technologies; imposing privacy protections by law or regulation, rather than through self-regulatory means, would raise compliance costs for businesses, with these increased costs falling hardest on small operators and start-up companies"); Bradyn Fairclough, *Privacy Piracy: The Shortcomings of the United States' Data Privacy Regime and How to Fix It*, 42 J. CORP. L. 461, 463 (2016) (mentioning the argument that businesses are in the best position to decide what are best for them and their consumers and the government should not infringe on the marketplace and burden the data-driven economy); FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESS AND POLICYMAKERS 11 (2012), <https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf> (mentioning the argument that the government is impeding the industry's ability to keep up with the rapidly changing marketplace) [<https://perma.cc/EB64-7FNV>].

¹⁹² U.S. GOV'T ACCOUNTABILITY OFF., *supra* note 151, at 29–30.

¹⁹³ *Id.* at 29.

¹⁹⁴ Natalie R. Bilbrough, *The FDA, Congress, and Mobile Health Apps: Lessons from DSHEA and the Regulation of Dietary Supplements*, 74 MD. L. REV. 921, 936 (2015).

¹⁹⁵ Drake, *supra* note 163, at 177.

¹⁹⁶ *Id.*

¹⁹⁷ FED. TRADE COMM'N, *supra* note 191, at 7–13.

¹⁹⁸ See Ira S. Rubinstein, *Privacy and Regulatory Innovation: Moving Beyond Voluntary Codes*, 6 J. L. & POL'Y FOR INFO. SOC'Y, 355, 356 (2011).

Self-regulation has succeeded in some sectors such as advertising,¹⁹⁹ but not in the field of consumer privacy protection, let alone the protection of mHealth app data.²⁰⁰ A paper published by the World Privacy Forum concluded that privacy self-regulation carried out in the past has failed for lack of transparency, credibility, sincerity, staying power, and meaningful enforcement.²⁰¹ FTC has similarly noted that “efforts to address privacy through self-regulation have been too slow, and have failed to provide adequate and meaningful protection.”²⁰² HHS also conceded that despite the recent best efforts, no widely adopted voluntary code of conduct has emerged in the area of non-HIPAA health data protection.²⁰³

Successful self-regulatory initiatives share several features.²⁰⁴ The following two are critical: (1) sufficient motivation to participate, and (2) effective monitoring and enforcement mechanisms.²⁰⁵ Unfortunately, the status quo of self-regulation of mHealth app data fails on both counts.

A fundamental problem with self-regulation is that it can only regulate those actors motivated or principled enough to participate.²⁰⁶ Without governmental interference, mHealth industry players’ incentive to take part in self-regulatory programs is quite low because there is no pressure along the industrial chain. Neither customers nor the downstream and upstream players are in a position to require changes. Specifically, the current mHealth app data market suffers two flaws.

¹⁹⁹ See Maureen K. Ohlhausen, Comm’r, Fed. Trade Comm’n, Address at the Better Business Bureau Self-Regulation Conference—Success in Self-Regulation: Strategies to Bring to the Mobile and Global Era (June 24, 2014), https://www.ftc.gov/system/files/documents/public_statements/410391/140624bbbself-regulation.pdf [<https://perma.cc/2BKA-ZQ93>].

²⁰⁰ See generally NAT’L TELECOMM. & INFO. ADMIN., U.S. DEP’T OF COMMERCE, *Chapter 1: Theory of Markets and Privacy*, in PRIVACY AND SELF-REGULATION IN THE INFORMATION AGE (1997), <https://www.ntia.doc.gov/page/chapter-1-theory-markets-and-privacy> (providing comprehensive discussions on the roles, limitations and failures of market, government and self-regulation in protecting personal information) [<https://perma.cc/6NFR-MUBW>].

²⁰¹ See generally ROBERT GELLMAN & PAM DIXON, WORLD PRIVACY F., *MANY FAILURES: A BRIEF HISTORY OF PRIVACY SELF-REGULATION IN THE UNITED STATES* (2011), <http://www.worldprivacyforum.org/wp-content/uploads/2011/10/WPFselfregulationhistory.pdf> (conducting a review of the first wave of privacy self-regulatory efforts in the area of digital privacy protection and finding that the “majority of these industry self-regulatory programs failed,” and “many disappeared entirely”) [<https://perma.cc/9499-UCTM>].

²⁰² John Schinasi, *Practicing Privacy Online: Examining Data Protection Regulations through Google’s Global Expansion*, 52 COLUM. J. TRANSNAT’L L. 569, 585 (2014).

²⁰³ U.S. DEP’T OF HEALTH & HUMAN SERVS., *supra* note 98, at 31.

²⁰⁴ See Ohlhausen, *supra* note 199.

²⁰⁵ *Id.*

²⁰⁶ A. Michael Froomkin, *The Death of Privacy*, 52 STAN. L. REV. 1461, 1527–28 (2000).

First, information asymmetry is severe between data collectors, processors and service providers on one side and consumers on the other side. As discussed earlier, information asymmetry prevents customers from making meaningful judgment on data processing and privacy threats, differentiating credible service providers from the inferior, making informed choices about whom they can trust, and eventually acting in their best interest. Thus, data handlers cannot benefit from participating in and complying with any self-regulatory initiative as long as consumers lack the knowledge to seek out those who comply.

Second, in the mHealth app industry, transacting data is much more lucrative than rendering the underlying services where the data comes from. This has shaped the whole data-driven industry. In the absence of external pressure, any “unnecessary” self-imposed limitation will only affect a player’s own profitability and its interaction with others within the industry. When a clear conflict of interest arises, and no external pressure exists, low incentive to participate becomes the first hurdle to self-regulation. In the end, profits remain the primary, if not the only, driver of businesses.²⁰⁷

Lack of effective monitoring and enforcement actions is the second major obstacle. As noted by many critics, self-regulation lacks accountability and suffers weak oversight and enforcement.²⁰⁸

To illustrate, this paragraph examines some recent self-regulation initiatives within the mHealth industry. The first initiative is from the CARIN Alliance, a multi-sector alliance of more than sixty providers, payers, consumers, electronic health record vendors, pharmaceutical companies, consumer platform companies, digital health companies, and consumer advocates.²⁰⁹ The CARIN Alliance developed a “voluntary code of conduct for entities not covered by HIPAA for handling health care data accessed via application programming interfaces” which requires obtaining “informed, proactive consent from users” and “giving consumers complete access and control over the use of their health care data,” and encourages consumer platform companies to “adopt the code as part of their consumer-facing application’s registration and onboarding process.”²¹⁰ The Consumer Electronics Association (CEA) issued its “Guiding Principles on the Privacy and Security of Personal Wellness Data” in October 2015, but adopting these principles is not compulsory for CEA members.²¹¹ Likewise, Xcertia, a self-

²⁰⁷ Bradyn Fairclough, *Privacy Piracy: The Shortcomings of the United States’ Data Privacy Regime and How to Fix It*, 42 J. CORP. L. 461, 464 (2016).

²⁰⁸ Rubinstein, *supra* note 198, at 356.

²⁰⁹ *Voluntary Code Established For Handling Health Care Data Not Covered by HIPAA*, 26 GUIDE TO GOOD CLINICAL PRACTICE NEWSL. 10 (2019).

²¹⁰ *Id.*

²¹¹ U.S. DEP’T OF HEALTH & HUMAN SERVS., *supra* note 98, at 31.

described “joint mHealth app collaborative effort” between app developers and healthcare organizations, recently finalized its 2019 version of mHealth App Guidelines, including the app privacy guidelines in its first section.²¹² The app privacy guidelines provide requirements in several aspects, such as “notice of use and disclosure, data retention, access mechanisms,” and compliance with HIPAA, Children’s Online Privacy Protection Act (COPPA), and GDPR.²¹³ In releasing an earlier draft of the 2019 Xcertia guidelines, Xcertia Board Chair Michael Hodgkins noted the “pressing need to establish a framework to evaluate mobile health apps “and said that “Xcertia encourages stakeholders [in the industry] to ... support the implementation of these guidelines in the market.”²¹⁴ By their own terms these standards are only “encouraged.”²¹⁵ Having standards is good, but standards only have teeth when outliers can be disciplined. Unfortunately, all the existing self-regulation initiatives are toothless.

Judging from the history and the current unsuccessful initiatives, complete self-regulation of privacy standards for mHealth app data is unlikely to be satisfactory. While “big government” intervention may go too far, the government still has a role to play. If there are valid, enforceable government-backed standards, companies that do not intend to undermine consumer privacy can be reassured that they will not face liability. Further, bad actors that undermine consumer privacy may see a legal penalty imposed to restrain from committing any breach of consumer privacy.

B. Comprehensive or Sectoral: A Pragmatic Perspective

From the perspective of governmental regulation, there have been widespread discussions among mHealth industry players, scholars and practitioners over how to regulate non-HIPAA health data. These parties have stressed the importance of reconciling existing data protections with new regulatory changes. Additionally, these parties argue that new regulations must provide sufficient protection while avoiding over-regulation.²¹⁶ Kirk Nahra, one of the leading practitioners in the field of privacy and cybersecurity, notes that there may be three options to achieve this balance:

²¹²XCERTIA, *2019 Board Approved Xcertia Guidelines*, XCERTIA MHEALTH APP GUIDELINES 2019, 3–9 (2019), <https://xcertia.org/wp-content/uploads/2019/08/xcertia-guidelines-2019-final.pdf> [<https://perma.cc/22VW-W5F9>].

²¹³*Id.*

²¹⁴ Eric Wicklund, *Xcertia Releases New mHealth App Guidelines, Adding 3 Categories*, MHEALTHINTELLIGENCE (Feb. 13, 2019), <https://mhealthintelligence.com/news/xcertia-releases-new-mhealth-app-guidelines-adding-3-categories> [<https://perma.cc/A7CE-EVLG>].

²¹⁵*Id.*

²¹⁶ See *infra* notes 240–42 and accompanying text.

[1] a specific set of principles applicable only to “non-HIPAA health care data” (with an obvious ambiguity about what “health care data” would mean); [2] a set of principles (through an amendment to the scope of HIPAA or otherwise) that would apply to all health care data; or [3] a broader general privacy law that would apply to all personal data (with or without a carve-out for data currently covered by the HIPAA rules).²¹⁷

An intuitive option would be for Congress to pass a comprehensive consumer data protection law that covers all data, including mHealth app data, regardless of the entities and platforms. This would offer all-inclusive protections, and may integrate all regulatory efforts under a single authority and simplify enforcement actions. Unfortunately, a federal comprehensive solution will not easily succeed due to political stakeholder concerns and potential constitutional challenges.

Therefore, policymakers should take a more pragmatic approach. A more efficient, effective solution for regulating mHealth app data would be to take a sectoral approach based on the federal health privacy exceptionalism policy.²¹⁸ Rather than waiting for Congress to pass a comprehensive privacy statute, federal agencies should instead refine the existing health data protection regulatory regime and address the immediate privacy concerns posed by mHealth app data.

1. *All Previous Attempts of Comprehensive Solutions Failed*

Following the implementation of GDPR, members of Congress have introduced a flood of privacy bills, including the Data Care Act,²¹⁹ proposed by Senators Brian Schatz, Amy Klobuchar, and Cory Booker,²²⁰ and the Consumer Data Protection Act proposed by Senator Robert Menendez.²²¹

²¹⁷ KIRK NAHRA, WILEY REIN LLP, MOVING TOWARD A NEW HEALTH CARE PRIVACY PARADIGM 4 (2014), (paper submitted to the Office of the National Coordinator for Health Information Technology’s Privacy and Security Working Group Virtual Hearing of Dec. 8, 2014), https://www.healthit.gov/sites/default/files/facas/PSWG_Background_Kirk_Nahra_Health_Care_Privacy_Paradigm_2014-12-08.pdf [<https://perma.cc/3JKW-AEA9>].

²¹⁸ See Terry, *supra* note 57 at 93–97.

²¹⁹ Data Care Act of 2018, S. 3744, 115th Cong. (2018).

²²⁰ Tom Davies, *New Data Protection Bill Could Strengthen Data Privacy Rights of US Consumers*, GDPR: REPORT (Dec. 17, 2018), <https://gdpr.report/news/2018/12/17/new-data-protection-bill-could-strengthen-data-privacy-rights-of-us-consumers/> [<https://perma.cc/GPF5-WPF3>].

²²¹ Consumer Data Protection Act, S. 2188, 115th Cong. (2018).

While these legislative actions are encouraging, the prospect of passing a comprehensive federal privacy law remains an uphill battle.

Past attempts to pass a federal general privacy protection regime have never succeeded. The BEST Practices Act, the Commercial Privacy Bill of Rights Act, and the Consumer Privacy Protection Act of 2011 were all abandoned by Congress.²²² The Obama administration's 2015 Consumer Privacy Bill of Rights also failed due to criticism from both industry players, who claimed it imposed an undue burden to the business, and privacy advocates, who claimed the privacy protections were inadequate.²²³

2. Constitutional Challenges

There are two key constitutional concerns preventing a comprehensive federal privacy statute. First, unlike in other countries (and some states), privacy is not a fundamental right written into the U.S. Constitution. By contrast, the EU has recognized that citizens have a fundamental right to data protection.²²⁴ The right to privacy is recognized in the Charter of Fundamental Rights of the European Union and Treaty on the Functioning of the European Union.²²⁵ Therefore, there was little doubt that adopting comprehensive data protection like GDPR was permissible because it protects a recognized fundamental right.²²⁶ Likewise, some states like California²²⁷ have the right to privacy specifically written into the state's constitution. This in a way explains why "states are leading the way on data privacy" in the United States.²²⁸

At the federal level, the U.S. Supreme Court has held that privacy is not an enumerated right, but is a protected social value coming from the "penumbras formed by emanations from those guarantees in the Bill of

²²² Schinasi, *supra* note 202, at 581.

²²³ Paul Bischoff, *What is the Consumer Privacy Bill of Rights and How Has it Evolved?*, COMPARITECH (Nov. 27, 2018), <https://www.comparitech.com/blog/vpn-privacy/consumer-privacy-bill-of-rights/> [<https://perma.cc/YTB6-XTPC>].

²²⁴ Drake, *supra* note 163, at 173.

²²⁵ Minke D. Reijneveld, *Quantified Self, Freedom, and GDPR*, 14 SCRIPTED 285, 288 (2017).

²²⁶ *But see* Bart van der Sloot, *Legal Fundamentalism: Is Data Protection Really a Fundamental Right?*, in DATA PROTECTION AND PRIVACY: (IN)VISIBILITIES AND INFRASTRUCTURES 3–5, 8 (Ronald Leenes et al. eds., 2017) (arguing that "data protection" has gradually been disconnected from the right to privacy and should be regarded as a consumer right).

²²⁷ Margaret Betzel, *Privacy Law Developments in California*, 2 I/S: J.L. AND POL'Y 831, 834 (2006).

²²⁸ Andrew Burt, *States Are Leading the Way On Data Privacy*, HILL (Aug. 21, 2018, 10:30 AM), <https://thehill.com/opinion/technology/402775-states-are-leading-the-way-on-data-privacy> [<https://perma.cc/5PRX-7MJT>].

Rights.”²²⁹ Thus, the Supreme Court has embraced a right to privacy, though the Court has only recognized the right in select areas such as marriage and abortion.²³⁰ As one commentator has noted, this constitutional patchwork parallels the legislative sectoral approach to data privacy.²³¹

The second consideration is the tension between freedom of commercial speech under the First Amendment and privacy protection. The Supreme Court’s decision in *Sorrell v. IMS Health* is instructive.²³² In *Sorrell*, the Court struck down a Vermont statute restricting the sale and use of pharmaceutical data, holding that the government could not engage in “content” or “viewpoint” discrimination against marketers by prohibiting the commercial use of data while permitting non-commercial use.²³³ In that opinion, the Court approved of HIPAA’s universal opt-in approach, which specifically regulates healthcare industry players’ use of personal data and requires notification to all consumers of how their personal data will be used.²³⁴ The Court’s ruling in *Sorrell* supports the sectoral approach, which regulates select sensitive contexts, and rejects approaches where specific harmful uses are carved out from general permissible use of data. This perhaps explains why previous context-specific statutes such as HIPAA and GINA were enacted without facing significant constitutional challenges.

Certainly, the First Amendment is not an absolute bar to a universal data protection regime, but it could be deployed as a powerful argument against such a regime.²³⁵ For instance, Jeff Joseph, President and CEO of the Software & Information Industry Association, recently criticized the California Consumer Privacy Act of 2018 (CCPA), claiming that it lacks a compelling public purpose and engages in content-based discrimination.²³⁶ He further argued that any federal-level privacy legislation had to “be structured to accomplish their important purpose in a way that is the least restrictive of

²²⁹ *Griswold v. Connecticut*, 381 U.S. 479, 484 (1965).

²³⁰ Michael C. James, *A Comparative Analysis of the Right to Privacy in the United States, Canada and Europe*, 29 CONN. J. INT’L L. 257, 268 (2013).

²³¹ *Id.* at 270.

²³² *Sorrell v. IMS Health Inc.*, 564 U.S. 552 (2011).

²³³ *Id.* at 565.

²³⁴ Katie Booth, *The All-or-Nothing Approach to Data Privacy: Sorrell v. IMS Health, Citizens United, and the Future of Online Data Privacy Legislation*, HARV. J. OF L. & TECH.: JOLT DIGEST (Vivian Tao ed., Aug. 17, 2011), <https://jolt.law.harvard.edu/digest/the-all-or-nothing-approach-to-data-privacy-sorrell-v-ims-health-citizens-united-and-the-future-of-online-data-privacy-legislation> [<https://perma.cc/5KST-KYJA>].

²³⁵ See Christopher Mohr, *Data is Speech: The Constitution Has a Role in Informational Privacy II*, SIIA (Oct. 11, 2018), <http://www.sii.net/blog/index/Post/76979/Data-is-Speech-The-Constitution-Has-a-Role-in-Informational-Privacy-II> [<https://perma.cc/73ZP-LEXL>].

²³⁶ Donald Gilliland, *We Need a National Privacy Law That Respects the First Amendment*, HILL (Mar. 13, 2019, 1:30 PM), <https://thehill.com/opinion/technology/433621-we-need-a-national-privacy-law-that-respects-the-first-amendment> [<https://perma.cc/S26P-XLQ9>].

speech.”²³⁷ Legislators need to address these potential arguments and polish the wording of any future privacy statute; they must differentiate between “permissible” and “harmful” uses of data and strike a careful balance between commercial needs and data protection.²³⁸

3. *Stakeholder Concerns*

The reactions of certain classes of stakeholders will have an impact on whether a given comprehensive privacy measure will succeed. In response to any comprehensive privacy bill, industry players will likely to be the first to fight. According to Karsten Weide, Media and Entertainment Program Vice President at International Data Corporation (IDC), “growing demand would cause data vendor sales to more than triple to \$10.1 billion by 2022, compared with \$3.1 billion in 2017.”²³⁹ If any general privacy legislation is adopted, data vendors stand to suffer, and the survival of the broader data-selling industry would be at risk. As one commentator said: “prepare for a new privacy lobbying battle.”²⁴⁰

Beyond industrial stakeholders, the federal government has consistently been reluctant to either make significant structural changes to the existing sectoral framework or to pass a comprehensive privacy law. For example, while proposing a comprehensive consumer privacy bill to Congress, the Obama administration limited the proposal to “commercial sectors that are not subject to existing federal data privacy laws.”²⁴¹ FTC also cautioned that overlapping or duplicative requirements should be avoided and suggested that a general data protection law should not apply to the entities that are subject to sector-specific laws like HIPAA.²⁴² The Trump administration said that the sectoral approach provided strong, focused

²³⁷ *Id.*

²³⁸ See generally Marsha Cope Huie, Stephen F. Larabee & Stephen D. Hogan, *The Right to Privacy in Personal Data: The EU Prods the U.S. and Controversy Continues*, 9 TULSA J. COMP. & INT’L L. 391, 470 (2002) (providing a historical review and discussions about conflict between freedom of commercial speech and privacy).

²³⁹ Aliya Ram & Madhumita Murgia, *Data Brokers: Regulators Try To Rein In The “Privacy Deathstars,”* FIN. TIMES (Jan. 7, 2019), <https://www.ft.com/content/f1590694-fe68-11e8-aebf-99e208d3e521> [<https://perma.cc/2ASM-SA5G>].

²⁴⁰ David Meyer, *In the Wake of GDPR, Will the U.S. Embrace Data Privacy?*, FORTUNE (Nov. 29, 2018), <http://fortune.com/2018/11/29/federal-data-privacy-law/> [<https://perma.cc/5XMT-7CQ7>].

²⁴¹ Terry, *supra* note 57, at 96.

²⁴² *Id.* at 97.

protections that should be maintained, and called for actions addressing consumer privacy in areas that are not currently covered by sectoral laws.²⁴³

Even if passed, comprehensive legislation would most likely only address baseline privacy issues and leave sector-specific details to industry regulators. In the case of health data, comprehensive legislation might simply carve out HIPAA-covered data. In that case, it would not be a problem to first fill the identified gaps of the current sectoral regulations.

4. *Health Privacy Exceptionalism*

A sectoral approach would not be a mere second-best solution. The policy of health privacy exceptionalism, as frequently discussed by Nicolas P. Terry,²⁴⁴ should be continued for regulation of mHealth app data. If a policy of health privacy exceptionalism is extended to mHealth app data, a sectoral approach would be justified. This Section argues that compared with other consumer data, which can to some extent give way to commercial innovation, mHealth app data warrants the government's exceptional regulatory attention, and should be covered by the long-standing health privacy exceptionalism.

First, as discussed in Part I, mHealth apps are the main, if not the largest, source of non-HIPAA health data. Some highly sensitive mHealth app data can pose significant harms to data subjects if misused in combination with Big Data analytics, just like traditional HIPAA-regulated health data.

Second, existing statutes cannot fully mitigate the harms posed by misuse of mHealth app. This further justifies the necessity of imposing robust *ex ante* regulations for mHealth app data and the continuance of health privacy exceptionalism.

For example, there are already anti-discrimination laws that cover the use of health data, but they do not cover most mHealth app data. The Americans with Disabilities Act (ADA) is supposed to prevent disability-based discrimination, but it does not prohibit predictions of future disability on the basis of data regarding things like health habits, stress level, and exposure to environmental pollutants.²⁴⁵ GINA better protects consumers by prohibiting discrimination based on genetic information and gene-based inference and suspicion, but the law only covers genetic information and

²⁴³ STEPHEN MULLIGAN ET AL., CONG. RESEARCH SERV., R45631, DATA PROTECTION LAW: AN OVERVIEW, 52 (2019) <https://crsreports.congress.gov/product/pdf/R/R45631> [<https://perma.cc/2GF6-7Z7P>]; see also Developing the Administration's Approach to Consumer Privacy, 83 Fed. Reg. 48,600, 48,601 (Sept. 26, 2018), <https://www.federalregister.gov/documents/2018/09/26/2018-20941/developing-the-administrations-approach-to-consumer-privacy> [<https://perma.cc/SKF3-PR5Z>].

²⁴⁴ See Terry, *supra* note 57.

²⁴⁵ Hoffman, *supra* note 72, at 93.

applies only to health insurers and employers.²⁴⁶ Both statutes fail to completely prevent discriminatory use of mHealth app data because mHealth app data may involve more kinds of health data (inherently or context-based, descriptive or predictive), more areas of health data (e.g., genetic information, illness, chronic conditions), and more actors (e.g., marketers, credit-card companies).

Actually, the government has adopted this exceptionalism approach to deal with other similarly sensitive data, when a stringent, general regulation is not advisable. For instance, in response to the pressing needs of regulating the problematic gathering and use of children's personal data as found in a 1998 FTC survey,²⁴⁷ FTC designed protections for children's online privacy and passed COPPA to govern the collection and use of personal information from minors.²⁴⁸ In 2013, COPPA was revised to enhance the regulatory protection by including an expanded definition of "personal information," and the definition of "commercial website operator."²⁴⁹ The exceptionalism mindset behind COPPA is the same: Although commercial innovation and development are valid considerations and should not be unnecessarily impeded, the social costs and potential risks associated with misuse of certain especially sensitive data are too high to tolerate.

To conclude, it seems prudent, at least at the outset, for the government to adopt a conservative approach to the regulation of mHealth app data and determine whether a lighter touch is possible in the future, when the industry is accustomed to, and customers are well educated on, good privacy practices. Granted, mHealth app data may meanwhile include some less sensitive data that does not necessarily deserve exceptional protections. However, that is something left for regulators to consider when designing the regulatory framework, not a pretext to deny robust regulations for mHealth app data altogether.

C. A Two-Prong Solution

Following the sectoral approach, this Section proposes a two-prong solution based on the existing HIPAA-FTC framework to grant mHealth app

²⁴⁶ *Id.* at 94.

²⁴⁷ FED. TRADE COMM'N, IMPLEMENTING THE CHILDREN'S ONLINE PRIVACY PROTECTION ACT—A REPORT TO CONGRESS 3 (2007), https://www.ftc.gov/sites/default/files/documents/reports/implementing-childrens-online-privacy-protection-act-federal-trade-commission-report-congress/07coppa_report_to_congress.pdf [<https://perma.cc/SHN3-8BCM>].

²⁴⁸ MULLIGAN ET AL., *supra* note 243, at 52.

²⁴⁹ Christina Scelsi, *Care and Feeding of Privacy Policies and Keeping the Big Data Monster at Bay: Legal Concerns in Healthcare in the Age of the Internet of Things*, 39 NOVA L. REV. 391, 411 (2015).

data different levels of protection according to their risk levels. Such a balanced proposal can capture all mHealth app data, but avoid over-regulation in the meantime.

Under this proposal, mHealth app data is classified into two defined categories: “mHealth data” and “mHealth consumer data.” mHealth data is mHealth app data qualifying as health data under the two-step definition discussed in Part I, and mHealth consumer data is mHealth app data other than mHealth data. mHealth data, the related apps, and the entities handling them will be governed by HIPAA after its proper expansion. mHealth consumer data, the related apps, and the entities handling them will be regulated through an FTC-led co-regulation mechanism embedded with principles such as data minimization, purpose specification, and contextual consistency.

1. *First Things First: Categorizing mHealth App Data*

Before discussing the proposed regulatory approach, it is necessary to determine how mHealth app data should be categorized for regulation purposes.

a. Different Approaches in Defining Regulated Health Data

(1) Regulated Health Data under HIPAA

The current HIPAA Rules regulate any information, whether oral or recorded in any form or medium, that is created or received by a covered entity or a business associate and relates to (a) the past, present, or future physical or mental health or condition of an individual, (b) the provision of health care to an individual, or (c) the past, present, or future payment for the provision of health care to an individual.²⁵⁰ “Health care” means “care, services or supplies related to the health of an individual,” including (1) preventive, diagnostic, therapeutic, rehabilitative, maintenance, or palliative care, and counseling, service, assessment, or procedure with respect to the physical or mental condition, or functional status, of an individual or that affects the structure or function of the body, and (2) sale or dispensing of a drug, device, equipment, or other item in accordance with a prescription.²⁵¹ By such a broad definition, regardless of the nature of data, all data identifiable to an individual collected or processed by a HIPAA-covered entity is PHI governed by HIPAA.²⁵² As discussed earlier, it is the type of data holders (rather than the sensitivity of the data) that is the dispositive factor in determining whether certain health

²⁵⁰ U.S. DEP’T OF HEALTH & HUMAN SERVS., *supra* note 98, at 28.

²⁵¹ 45 C.F.R. § 160.103 (2019).

²⁵² U.S. DEP’T OF HEALTH & HUMAN SERVS., *supra* note 98, at 3.

data will be regulated by HIPAA. In that sense, HIPAA is both over-inclusive and under-inclusive.

(2) Regulated Health Data under GDPR

By contrast, GDPR regulates “data concerning health,” and defines data concerning health as “personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status.”²⁵³ This definition is further clarified by Recital 35 of GDPR by providing examples of what the regulators intend to cover.²⁵⁴ Recital 35 states that:

Personal data concerning health should include all data pertaining to the health status of a data subject which reveal information relating to the past, current or future physical or mental health status of the data subject. This includes information about the natural person collected in the course of the registration for, or the provision of, health care services to that natural person; a number, symbol or particular assigned to a natural person to uniquely identify the natural person for health purposes; information derived from the testing or examination of a body part or bodily substance, including from genetic data and biological samples; and any information on, for example, a disease, disability, disease risk, medical history, clinical treatment or the physiological or biomedical state of the data subject independent of its source, for example from a physician or other health professional, a hospital, a medical device or an in vitro diagnostic test.²⁵⁵

Regardless of the sources and custodians, GDPR is more focused on the inherent characteristics of the data at issue than HIPAA, but GDPR’s definition is over-inclusive and may be unmanageable.

²⁵³ Mantovani et al., *supra* note 169, at 90.

²⁵⁴ *Id.*

²⁵⁵ Regulation (EU) 2016/679 2016 O.J (L 119/6) (General Data Protection Regulation).

(3) Regulated Health Data under EU mHealth Code

The draft Privacy Code of Conduct on Mobile Health (mHealth) Apps facilitated by the European Commission²⁵⁶ (the EU mHealth Code) and its former supervising body the Article 29 Working Party have provided a context-based, fact-specific approach in categorizing health data.²⁵⁷ The draft EU mHealth Code states that “the context of processing, and particularly the purpose for which the app is made available or whether the data is made available through the app to a member of the medical community, is also relevant to determine whether data should be qualified as data concerning health.”²⁵⁸

The Article 29 Working Party further contended that to determine whether certain data is health data, not only the type of the data, but also the intended use of the data and its combination with other datasets should be considered; specifically, personal data is health data when (1) the data is inherently or clearly medical data; (2) the data is raw sensor data that can be used in itself or in combination with other data to draw a health-related conclusion; or (3) conclusions are drawn about a person’s health status or health risk.²⁵⁹

The EU mHealth Code’s approach generally fits the two-step definition discussed in Part I, but without a workable tool, this approach could be open-ended and unmanageable, especially for the context test in the second step.

b. Proposed Approach in Categorizing Regulated Data

To avoid over-inclusion or under-inclusion and to provide a workable approach, this Section suggests that mHealth app data be categorized by a combination of FDA’s classification of mHealth apps and the two-step definition of health data discussed in Part I. Specifically, FDA’s classification of mHealth apps has provided a practical tool in applying the test of health-

²⁵⁶ Press Release, European Commission, Code of Conduct on Privacy for mHealth Apps Has Been Finalised (Jun. 7, 2016), <https://ec.europa.eu/digital-single-market/en/news/code-conduct-privacy-mhealth-apps-has-been-finalised> (stating that the Privacy Code of Conduct on mobile health (mHealth) apps aims to promote trust among users of mHealth apps and will provide a competitive advantage for those who sign up to it in the future) [<https://perma.cc/886M-EAGD>].

²⁵⁷ Mantovani et al., *supra* note 169, at 90–91.

²⁵⁸ EUROPEAN COMMISSION, DRAFT CODE OF CONDUCT ON PRIVACY FOR MOBILE HEALTH APPLICATIONS 2 (2016), http://ec.europa.eu/newsroom/dac/document.cfm?action=display&doc_id=16125 (click the “Code” hyperlink in the press release text) [<https://perma.cc/3R85-X33P>].

²⁵⁹ Mantovani et al., *supra* note 169, at 90–91.

related context in the second step of the definition. This combined approach creates two categories of data:

First, “mHealth data” refers to data generated by the first and second categories of mHealth apps under FDA’s classification, namely those mHealth apps qualifying as devices under the FDCA. As discussed previously, such data is either inherently medically meaningful under the first step of the new definition or collected and used in a medical context under the second step of the new definition.

Second, “mHealth consumer data” refers to data generated by the third category of mHealth apps. Such data usually (1) does not by itself point to any specific disease or condition, and (2) does not incur any medical context or proclaim any medical-related purpose. The caveat is that if any data generated by the third category of mHealth apps is inherently medically meaningful, it should, according to the first step of the new definition, always be considered mHealth data. For example, if one app is designed to check your ancestry, not to identify diseases, the genetic data so collected and generated should still be considered mHealth data because it is inherently health-related, although the data is used outside a medical context.

2. *Expanding HIPAA to Cover mHealth Data*

HIPAA should expand its reach to include mHealth data and enhance the data minimization requirement to address the concern of excessive data collection. Because mHealth data is sensitive data generated in non-HIPAA health-related settings it should receive the same treatment as HIPAA-covered health data presently. The current HIPAA exceptions, such as free use for public health research purposes, should also apply to mHealth data. The proposed expansion has two aspects.

First, HIPAA should change its covered entity paradigm and regulate all entities and mHealth apps that collect, use, and process mHealth data. Although mHealth data is collected and processed outside the traditional medical setting, the nature of this data is similar to data held by the HIPAA-defined covered entities and business associates in terms of sensitivity and risks.

In response to this call to expand HIPAA, HHS stated that the HIPAA Rules work only for healthcare providers and insurers, and a simple extension of HIPAA may not be practical.²⁶⁰ This argument might not be convincing.

The current HIPAA Rules do not only apply to healthcare providers and insurers, but also their business associates—including any non-medical entity providing, among other things, legal, accounting, and IT services. In particular, some mHealth data collected on behalf of health providers by apps in the capacity of business associates have already been regulated under HIPAA. For instance, PHR vendors—who qualify as business associates—may provide online accounts linked to wearable health and fitness devices to collect PHI on behalf of covered entities so that health care providers can then advise their patients about their health based on the data collected.²⁶¹ The activity of PHR vendors is comparable to that of developers of consumer-facing mHealth apps, except they are engaged by covered entities and qualify as business associates.

Some provisions of the HIPAA Rules may need to be refined to fit mHealth data collectors. For example, the current rules rely on Institutional Review Boards (IRB) and privacy boards that are appropriate for large institutions and require significant resources and expertise.²⁶² These boards may not be compatible with or appropriate for consumer-generated mHealth data because of the overhead and relatively lower professional standards involved, but they are only one part of the HIPAA Rules. Most of the HIPAA Rules are privacy-related requirements for general use which could be adopted to govern smaller developers. With necessary refinements to address these resource issues the expansion of HIPAA to cover these developers should be practical.

Second, HIPAA should expand its regulatory scope by enhancing data subjects' control over their health data, especially upstream protection regarding data collection. This would be a necessary accommodation because of the combination of mHealth technology and Big Data analytics. As discussed earlier, because the HIPAA Rules were only contemplating the traditional clinical setting at their enactment, disclosure of personal information by data subjects was assumed. However, the current technologies have incentivized excessive data collection and posed pressing privacy

²⁶⁰ NAT'L COMM. VITAL & HEALTH. STATISTICS, U.S. DEP'T HEALTH & HUMAN SERVS., HEALTH INFORMATION PRIVACY BEYOND HIPAA: A 2018 ENVIRONMENTAL SCAN OF MAJOR TRENDS AND CHALLENGES 51 (2017), https://ncvhs.hhs.gov/wp-content/uploads/2018/05/NCVHS-Beyond-HIPAA_Report-Final-02-08-18.pdf (“The same rules may not work in the same way for other health data processors so that any simple extension of HIPAA may not be practical.”) [<https://perma.cc/VQ7H-DMAN>].

²⁶¹ WU, *supra* note 101, at 31.

²⁶² 45 C.F.R. § 164.512(i) (2019).

concerns. After examining GDPR, CCPA, and FIPPS, the following principles should be added to strengthen HIPAA's upstream protections.

(1) *Data minimization for data collection.* Currently HIPAA imposes a minimum necessary requirement to the downstream (post-collection) area of health data. Specifically, the minimum necessary standard requires covered entities to allow access to and disclosure of PHI on a need-to-know basis.²⁶³ A covered entity then must set up and implement mechanisms for role-based access and use of PHI among its members of the workforce.²⁶⁴ Such a minimum necessary standard is now equally important at the data collection stage if mHealth data is to be covered. The rationale is that only data necessary to implement mHealth apps' functions for the stated purposes should be collected, and relevant policies and procedures should similarly be developed to implement this standard.

To be clear, the proposed upstream limitations are not to unreasonably limit data collection and frustrate mHealth apps' purposes, but to ensure that data collection takes place only to the extent necessary and in a transparent way. Because of the lax regulatory environment, current extensive data collection practices may be the core part of the business model and the main profit source for many mHealth apps. Had there been upstream limitations, these apps would perhaps never have been created. Some may argue that if upstream limitations were imposed, there would undoubtedly be a loss to these businesses. However, that is exactly one of the issues that should be tackled by data protection regulations: mHealth app data is supposed to fulfill meaningful health-related functions in exchange for reasonable returns, rather than profiting from mining, exploiting, and selling data against data subjects' interests. If an mHealth app's primary purpose is to take advantage of consumers' privacy, it makes sense to have the statutory upstream limitations to screen it out. A pure financial loss argument from commercial actors should not defeat the necessity of adding upstream limitations into HIPAA to prevent unreasonable collection, mining, and exploitation of health data.

(2) *Right to deletion.* The right to deletion is the other side of the coin of data minimization. Similar rights have been embraced by GDPR and CCPA. For example, GDPR gives a data subject the right of erasure of his or her personal data without undue delay when, among other criteria, (i) the personal data is no longer necessary in relation to the purposes for which it was collected or otherwise processed; (ii) the data subject objects to processing and there are no overriding legitimate grounds for the processing; (iii) the personal data has been unlawfully processed; or (iv) the personal data has to

²⁶³ U.S. DEP'T OF HEALTH & HUMAN SERVS., OCR HIPAA PRIVACY, MINIMUM NECESSARY (2003), <https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/minimumnecessary.pdf> [<https://perma.cc/YG6E-Y2EE>].

²⁶⁴ U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 106, at 10.

be erased for the data subject to be in compliance with a legal obligation.²⁶⁵ Of course, such a right is not unlimited. GDPR does have exceptions to the right to erasure that address situations in which retention is desirable for, among others, public health or scientific archiving reasons.²⁶⁶ The right to deletion set forth in CCPA seems to be more business-friendly because it expressly provides that the right to deletion is not exercisable if the data in question is needed for completing transactions, maintaining business relationships, performing contracts, and solely for reasonable internal use.²⁶⁷

HIPAA takes the opposite approach. As part of the documentation requirements, covered entities are required to retain the data for “6 years from the date of its creation or the date when it last was in effect, whichever is later.”²⁶⁸ This Article suggests that HIPAA should similarly grant data subjects a right to request deletion of their health data if such data is no longer relevant or necessary, subject to reasonable exceptions such as use for public health research or for business necessity.

Although the public health exception for the right to deletion could easily follow the existing HIPAA standards, the applicability of the reasonable business necessity exception would invite more debate, because it must be crafted to support legitimate business functions without permitting more invasive data mining. This Article does not attempt to go into the details regarding the possible interpretations of this exception but proposes two premises on which this exception should be based.

First, the business necessity of the data collection should be made clear to data subjects so that they are able to make an informed choice about participation before the data is collected. As an example, consider a gene testing app that presents two options to users. One option is a one-off test to check ancestry or disease risks; the other is a five-year gene research program in exchange for a discount. If a user only opts in the one-off test, there is no basis for the app to claim business necessity even if the app’s primary business model is based on data aggregation. This lack of business necessity is an incentive for users to opt for the one-off test—just as much as the discount is an incentive for them to opt for the five-year program. This information should be presented to users before they select an option.

Second, reliable de-identification measures should be in place if an app invokes the business necessity exception to retain users’ data so that the

²⁶⁵ Tovino, *supra* note 99, at 990–91.

²⁶⁶ *Id.* at 991.

²⁶⁷ David Kessler & Anna Rudawski, *CCPA Extends “Right to Deletion” to California Residents*, NORTON ROSE FULBRIGHT BLOG NETWORK: DATA PROTECTION REPORT (Sept. 27, 2018), <https://www.dataprotectionreport.com/2018/09/ccpa-extends-right-to-deletion-to-california-residents> [<https://perma.cc/RSM6-FQVF>].

²⁶⁸ 45 C.F.R. § 164.316 (2019).

exposures to data-related harms can be reduced to a minimal level. That is, if the data invokes the exception and does not delete the data, it should at least be de-identified with robust technologies to protect against re-identification.

3. *FTC-led Co-regulation: Taking It a Step Further*

For regulation of mHealth consumer data, this Article suggests that FTC take a step forward to turn self-regulation into co-regulation. As a collaborative process, co-regulation is a middle ground between complete self-regulation and strict agency regulation.

a. Success Stories of Co-regulation

Co-regulation is not a new creature, but a verified, effective methodology. Before the passage of GDPR, the EU used a co-regulatory approach to protect privacy. The typical process went as follows: after a member state passed a comprehensive data protection statute, sectoral representatives would be invited to draft a code of conduct for different sectors detailing the data protection requirements.²⁶⁹ Once approved by the regulator, the drafted code of conduct would take on the force of law, and data custodians within that sector could be punished for noncompliance.²⁷⁰

Currently, the EU intends to continue this approach to regulate mHealth apps.²⁷¹ The European Commission is now facilitating the passage of the EU mHealth Code.²⁷² If no more comments are received, the current draft will be approved by the European Data Protection Board and be granted general validity across the EU.²⁷³ An app provider intending to obey this code of conduct may apply for a trust mark for its app and, once approved label its app with the trust mark.²⁷⁴ Accordingly, adherence to an approved code of conduct would be “an element to demonstrate compliance” with data protection requirements.²⁷⁵

Co-regulation is not new in the United States either. In fact, the White House acknowledged the merits of such a collaborative, multi-stakeholder

²⁶⁹ Dennis D. Hirsch, *The Law and Policy of Online Privacy: Regulation, Self-Regulation, or Co-Regulation*, 34 SEATTLE U. L. REV. 439, 442 (2011).

²⁷⁰ *Id.*

²⁷¹ Press Release, European Commission, Privacy Code of Conduct on Mobile Health Apps (Dec. 10, 2018), <https://ec.europa.eu/digital-single-market/en/privacy-code-conduct-mobile-health-apps> [<https://perma.cc/XMQ6-Y72A>].

²⁷² *Id.*

²⁷³ *Id.*

²⁷⁴ EUROPEAN COMMISSION, *supra* note 258, at 74.

²⁷⁵ Mantovani et al., *supra* note 169, at 73.

process in a 2012 report.²⁷⁶ The report stated that the federal government intended to convene discussions among stakeholders to develop industry-specific codes of conduct to protect privacy; stakeholders include companies, consumers, privacy advocates, international partners, state attorneys general, federal criminal and civil law enforcement representatives, and academics.²⁷⁷

Federal agencies already have some success stories of co-regulation, although the legal criteria and distribution of responsibilities are different from the EU practice. According to the EU practice, industry-made codes of conduct would be recognized as having direct applicability across the EU, but the U.S. practice is more focused on utilizing the expertise of the industry and relevant associations to come up with and incorporate professional standards. The FDA-USP model is a typical example. United States Pharmacopeia (USP) is a non-profit organization relying on its convention member organizations and their delegates to discuss important industry issues and to carry out critical governance activities. Its members include academic institutions, health practitioners, scientific associations, consumer organizations, manufacturer and trade associations, government bodies, and non-governmental standards-setting and conformity assessment bodies.²⁷⁸ USP developed and published standards for drug substances, drug products, excipients, and dietary supplements in the United States Pharmacopeia-National Formulary, and these standards were officially recognized through the Federal Food and Drug Act of 1906.²⁷⁹ Today, USP's compendial standards remain connected to FDA provisions and other consumer protection laws and regulations.²⁸⁰ Although USP does not have its own enforcement power, FDA enforces any breach of USP standards or provisions.²⁸¹

In the field of health IT, the Office of the National Coordinator for Health Information Technology (ONC) has been successfully operating a voluntary certification program of health IT since 2010.²⁸² This program is a

²⁷⁶ WHITE HOUSE, CONSUMER DATA PRIVACY IN A NETWORKED WORLD: A FRAMEWORK FOR PROTECTING PRIVACY AND PROMOTING INNOVATION IN THE GLOBAL DIGITAL ECONOMY 2 (2012), <https://obamawhitehouse.archives.gov/sites/default/files/privacy-final.pdf> [<https://perma.cc/LZ4V-3MHS>].

²⁷⁷ *Id.*

²⁷⁸ U.S. Pharmacopeia, *Convention Membership*, USP.ORG, <http://www.usp.org/about/convention-membership> (last visited Dec. 21, 2019) [<https://perma.cc/38UJ-X94B>].

²⁷⁹ Sarah Jean Kilker, *Effectiveness of Federal Regulation of Mobile Medical Applications*, 93 WASH. U. L. REV. 1341, 1353 (2016).

²⁸⁰ U.S. Pharmacopeia, *USP and the FDA Working Together to Protect Public Health*, USP.ORG, <http://www.usp.org/about/public-policy/usp-fda-roles> (last visited Dec. 21, 2019) [<https://perma.cc/5ZUC-8CNW>].

²⁸¹ Kilker, *supra* note 279.

²⁸² See generally OFFICE NAT'L COORDINATOR FOR HEALTH INFO. TECH., U.S. DEP'T HEALTH & HUMAN SERVS., HEALTH IT CERTIFICATION PROGRAM

quasi-co-regulation mechanism. First, ONC establishes the underlying program requirements and certification criteria for certified health IT, taking into consideration input from within the federal government, as well as from public and private stakeholders.²⁸³ Second, ONC collaborates with third-party organizations to perform functions such as conformance testing, certification issuing, and surveillance, and health IT developers apply for certification on a voluntary basis.²⁸⁴ Meanwhile, ONC maintains a complementary power of direct review to promote health IT developers' accountability for performance, reliability, and safety of health IT.²⁸⁵

b. Proposed Co-Regulation Approach for mHealth Consumer Data

The co-regulation approach for mHealth consumer data proposed by this Article is a similar collaborative process. Specifically, the approach consists of the following:

(1) *Development of Code of Conduct*. First, a trusted public-private entity representing various interests—including regulators, app developers, app platforms, privacy advocates, end users, and industry experts—should be created or, if feasible, selected from among the existing third-party organizations. The public-private entity would develop a code of conduct for mHealth app data protection through multi-stakeholder negotiations.

(2) *FTC Approval of Code of Conduct*. Second, FTC would review, comment on, and finally approve a code of conduct to ensure that FIPPS principles, FTC's principles summarized from its own enforcement experience, and other widely acknowledged privacy protection practices (such as enhanced notice and express consent for sensitive use, as stated in GDPR) are in place.

(3) *Voluntary Opt-in for Trust Mark*. Third, mHealth app developers intending to adhere to the FTC-approved code of conduct would submit a privacy impact assessment to the public-private entity for review of possible privacy risks and recommendations of any appropriate mitigating measures. If the privacy impact assessment is passed, applicants would be qualified to voluntarily apply for a trust mark from the public-private entity by submitting a declaration of adherence, which would be registered with FTC for record and publication. If a trust mark is granted, the app developers would be entitled

OVERVIEW (2019), <https://www.healthit.gov/sites/default/files/PUBLICHealthITCertificationProgramOverview.pdf> (providing an overview of the Health IT Certification Program, including the program participants, program structure, surveillance of certified health IT, and ONC's direct review of certified health IT) [<https://perma.cc/T6SC-HFGT>].

²⁸³ *Id.* at 2.

²⁸⁴ *Id.* at 1.

²⁸⁵ *Id.* at 4.

to label their apps with such marks. End users, as well as healthcare providers who recommend mHealth apps to their patients, would be educated that mHealth apps with trust marks are government-endorsed and trustworthy.

(4) *FTC's Section 5 Enforcement.* Finally, if an adhering app developer is found to be in violation of the code of conduct as a result of a random check, a verified consumer's complaint, or otherwise, FTC would exercise its existing enforcement power under the deceptiveness prong of Section 5 of the FTC Act and rescind such developer's trust marks. The results of such enforcement would be announced to the public.

As long as appropriately structured, this co-regulation approach would have several advantages over agency regulation and self-regulation. Unlike traditional agency regulation, this mechanism would delegate operational tasks such as developing industry-specific codes of conduct, assessing privacy impacts, and issuing trust marks to a specific public-private entity so that rulemaking and monitoring can be conducted in a relatively neutral, professional, efficient, and customized way.

In the meantime, this government-sponsored initiative can also address the two concerns posed by complete self-regulation discussed earlier: One is that self-regulation lacks sufficient incentive for the players to obey, and the other is that no enforcement is guaranteed.

With respect to the first concern, under this framework, external factors would provide adequate incentive for app providers to adhere to their industry's code of conduct for fear of losing consumers' trust and market competitiveness. Current self-regulation efforts by purely private entities are scattered and without credibility, while this co-regulation approach is centralized, well-structured, and authoritative. The granted trust mark acts as a simple, straightforward, and credible identifier of trustworthy mHealth apps, and places pressures on industry players from the outside. Some may doubt the significance of such a trust mark, but experience has shown that it can calibrate information asymmetry and effectively direct consumers' choices. For example, privacy seals, similar to trust marks, issued by TRUSTe in the 1990s have been a key driver for websites to formulate privacy policies and conform to basic privacy norms.²⁸⁶ An indication of the government's endorsement would only reinforce such function.

As to the second concern, adhering app providers would clearly trigger FTC's existing Section 5 power if opting-in app providers breach consumers' privacy and violate its own declaration of adherence as a result. No additional interpretation or standards would be needed in this respect.

²⁸⁶ Solove & Hartzog, *supra* note 93, at 593.

Proponents of co-regulation see it as “the best of both worlds.”²⁸⁷ Although it may be a stretch to consider co-regulation a perfect solution, it does present a good combination of strengths of the different players, with industries as experts and the government as an enforcer,²⁸⁸ and should at least be an effective stopgap measure to fill the current regulatory gaps.²⁸⁹

V. CONCLUSION

Extensive use of mHealth apps and Big Data analytics have blurred the line between health data and other consumer data and have allowed entities outside the traditional healthcare ecosystem to collect, hold, transmit, and process health data. New privacy concerns arising from mHealth technological developments have posed new challenges to the current U.S. regulatory framework for health data protection.

HIPAA does not regulate most mHealth app data. FDA is hands-off to data protection. FTC stands to regulate consumer data in general, including mHealth app data, but is unfortunately constrained by an absence of clear standards and by limited police power. These regulatory gaps need to be addressed.

It seems that Congress is now considering passing a comprehensive consumer data protection statute, given the increased attention to privacy issues and enhanced enforcement by peer nations. However, whether such an effort will come to fruition remains an open question due to some debatable constitutional concerns and political considerations.

This Article argues that, before any general consumer privacy law is passed, a more effective, practical solution would be to refine the current framework to regulate mHealth app data. It proposes a single two-prong solution to provide different levels of protection to different categories of mHealth app data.

FDA’s functionality-based typology of mHealth apps, in combination with the two-step definition of health data in the era of Big Data proposed by this Article, would help determine the categorization of sensitive mHealth data

²⁸⁷ See Hirsch, *supra* note 269, at 441.

²⁸⁸ *Id.*

²⁸⁹ It is also suggested that under this two-prong regulatory framework, FTC, HHS, and FDA, develop a cross-agency tool to assist mHealth app developers assess the kinds of mHealth app data they collect (mHealth data or mHealth consumer data) and determine whether the developers should comply with the HIPAA Rules, turn to the FTC co-regulation regime for trust marks, or do both, just as these agencies did with the previously-developed Mobile Health Apps Interactive Tool. See FED. TRADE COMM’N, MOBILE HEALTH APPS INTERACTIVE TOOL, <https://www.ftc.gov/tips-advice/business-center/guidance/mobile-health-apps-interactive-tool> [<https://perma.cc/XV5N-FFJV>].

versus less sensitive mHealth consumer data. mHealth data is health data and would be covered by the expanded HIPAA.

mHealth consumer data is less sensitive than mHealth data but overall more health-related and more likely to become health data than other consumer data due to the context set by the mHealth apps in question. As a result, such data would be governed by a co-regulation regime endorsed by FTC and stricter than the current regulation-free situation. This co-regulation approach would include industry-specific codes of conduct representing interests of different stakeholders and would introduce a trust mark mechanism to motivate industry players and address information asymmetry issues among consumers. As a middle ground between complete self-regulation and traditional agency regulation, this co-regulation regime properly combines the expertise of industry players with the enforcement power of the relevant government agency.

NOTES

INCITEMENT AND THE GEOPOLITICAL INFLUENCE OF FACEBOOK CONTENT MODERATION

Sarah Koslov*

CITE AS: 4 GEO. L. TECH. REV. 183 (2019)

TABLE OF CONTENTS

I. INTRODUCTION	184
II. AN OVERVIEW OF FACEBOOK’S COMMUNITY STANDARDS: GUIDELINES, IMPLEMENTATION, AND GOVERNANCE IN ISRAELI AND PALESTINIAN TERRITORIES	187
A. Facebook’s Publicly Available Community Standards Are Articulated in Broad Terms to Facilitate Global Applicability, but Their Acontextual Framing Invites Subjective and Arbitrary Interpretations that Fill Policy Gaps.	188
B. Facebook’s Internal Guidelines Reveal an Evolving, Patchwork Approach to Content Moderation that Contributes to Uneven Enforcement of the Community Standards	190
C. The Complex Legal Landscape Within Israeli and Palestinian Territories Creates Incentives for Facebook to Streamline Content Moderation Through Their Community Standards, Disproportionately and Negatively Impacting Specific Communities.	192
III. CONSIDERING THE ENABLING FACTORS: THE CHALLENGES OF ALLOCATING RESPONSIBILITY AND AGENCY TO A PRIVATE PLATFORM IN EFFORT TO PROTECT INDIVIDUAL EXPRESSIVE INTERESTS	193
A. Suspected Double Standards in Enforcement of the Community Standards Increase Concerns of Targeted Censorship and Undermine Users’ Trust in Facebook.	194
B. Facebook’s Opacity Eschews Democratic, Procedural Norms That Establish the Scope of Permissive Forms of Public Expression. .	197

* J.D. Candidate, Georgetown University Law Center, 2020; B.A., University of South Carolina, 2014. Thank you to Professor Julie Cohen for her guidance, insight, and feedback throughout the development and writing of this note. I am also grateful to Lyn Abdullah and the editors of GLTR for their thoughtful comments and assistance.

C. Algorithmic Flagging and Artificial Intelligence (AI) Tools Heighten, Rather Than Mitigate, Unequal Levels of Suspicion and Surveillance.....	200
D. Secret Flows of Information Between Private and Public Actors for the Purposes of Surveillance and Arrest Thwart Procedural Oversight and Legal Guardrails that Are Normatively Desirable and Functionally Expected.	202
IV. OPPORTUNITIES: DEFINING, DELINEATING, AND CLARIFYING THE ROLE OF CONTENT MODERATION IN GLOBAL AFFAIRS	205
A. Understanding Facebook’s Power As Both Distinct From and In Relation to Government.....	206
B. Transparency Can Improve Facebook’s Accountability to Users by Creating Opportunities to Clarify and Correct.....	208
C. A Human Rights Approach to Content Moderation Creates Space for Competing Narratives, Experiences, and Truths.....	211
V. CONCLUSION.....	214

I. INTRODUCTION

Between 2015 and 2016, a new surge of violence broke out against Israeli civilians and soldiers, typified by lone actors using knives to attack at sudden and opportunistic times.¹ Most often, assailants were Palestinian men without ties to any formal organization or a history of engaging in violence.² During this period, attackers killed thirty-four Israelis while nearly two hundred Palestinians died trying to carry out the stabbings.³ While there remained no clear evidence that the attacks were organized or supported by a particular faction of the Palestinian community, some Palestinians took to social media voicing encouragement and support for the attackers.⁴ Opposing

¹ Peter Beaumont, *What’s Driving the Young Lone Wolves Who Are Stalking the Streets of Israel?*, GUARDIAN (Oct. 17, 2015), <https://www.theguardian.com/world/2015/oct/18/knife-intifada-palestinian-israel-west-bank> [<https://perma.cc/B36X-U6EC>]; William Booth & Ruth Eglash, *Israelis Are Calling Attacks a “New Kind of Palestinian Terrorism,”* WASH. POST (Dec. 25, 2015), https://www.washingtonpost.com/world/middle_east/israelis-are-calling-attacks-a-new-kind-of-palestinian-terror/2015/12/24/e162e088-0953-4de5-992e-adb2126f1dcc_story.html [<https://perma.cc/Z2X8-4YK4>].

² Booth & Eglash, *supra* note 1.

³ Ruth Eglash, William Booth & Darlas Cameron, *A New Kind of Terrorism in Israel*, WASH. POST, <https://www.washingtonpost.com/graphics/world/israel-palestine-deaths/> (last entry Sept. 2016) [<https://perma.cc/FMX6-XCWU>]; see also Pierrick Leurent, *Palestinian “Knife Intifada” Reflects a Generation’s Despair*, FRANCE 24 (June 5, 2016), <https://www.france24.com/en/20160506-reporter-israel-knife-intifada-palestinian-territories-violence> [<https://perma.cc/J3EX-RSLL>].

⁴*Is Palestinian-Israeli violence being driven by social media?*, BBC NEWS (October 22, 2015), <https://www.bbc.com/news/world-middle-east-34513693> [<https://perma.cc/6ZBP-7QUS>].

narratives developed around the “Knife Intifada”⁵ and the motivations behind it. The Israeli government blamed the violence on inciting posts on Facebook, claiming that content encouraging violence against Israelis permeated the platform. Meanwhile, Palestinians argued that the attackers acted out of frustration, desperation, and a loss of hope after years of occupation and displacement.⁶

The Israeli government’s actions following the attacks illustrate how Facebook has emerged as an influential intermediary in politically conflicted territories. In the midst of the Knife Intifada, the Israeli government called on Facebook to aid it in combating lone-wolf attacks.⁷ Israeli Justice Minister Ayelet Shaked explained that the government wanted Facebook “to themselves remove posts by terrorist groups and incitement to terrorism without us having to flag each individual post.”⁸ Frustrated that Facebook was not cooperating satisfactorily,⁹ Public Security Minister Gilad Erdan appeared on television calling Facebook a “monster” and said “some of the victims’ blood is on Zuckerberg’s hands.”¹⁰ In addition, the Knesset began working on a new “Facebook Bill” granting law enforcement officials broad authority to seek court orders compelling Facebook to remove content based on police recommendations.¹¹ Facebook, which usually prefers to wield its power

⁵ Beaumont, *supra* note 1.

⁶ 7AMLEH, FACEBOOK AND PALESTINIANS: BIASED OR NEUTRAL CONTENT MODERATION POLICIES 7 (2018), <https://7amleh.org/wp-content/uploads/2018/10/booklet-final2-1.pdf> [<https://perma.cc/8ZCV-2AV7>]; Beaumont, *supra* note 1; Bethan McKernan, *Israel and Facebook Team Up to Combat Social Media Posts That Incite Violence*, INDEP. (Sept. 14, 2016), <https://www.independent.co.uk/news/world/middle-east/israel-facebook-team-up-social-media-posts-incitement-violence-a7306436.html> [<https://perma.cc/6NCB-V9KT>].

⁷ Amar Toor, *Israel Calls Facebook a “Monster” for Not Helping To Curb Violence*, VERGE (July 4, 2016), <https://www.theverge.com/2016/7/4/12092762/israel-facebook-palestinian-attacks-censorship> [<https://perma.cc/M476-4TB4>].

⁸ *Id.*

⁹ Mark Bergen, *Israel’s Public Security Minister Blames Facebook After Recent West Bank Attacks*, RECODE (July 3, 2016), <https://www.recode.net/2016/7/3/12090532/israel-facebook-west-bank> [<https://perma.cc/6PL2-HU36>]; David Wainer, *Israel Accuses Facebook of Complicity in West Bank Violence*, BLOOMBERG BUS. (July 3, 2016), <https://www.bloomberg.com/news/articles/2016-07-03/israel-accuses-facebook-of-contributing-west-bank-violence> [<https://perma.cc/H7PX-RHEZ>].

¹⁰ Jonathan Lis, *Israeli Minister Slams Facebook: “Terror Victims’ Blood Is on Zuckerberg’s Hands,”* HAARETZ (July 3, 2016), <https://www.haaretz.com/israel-news/israeli-minister-terror-victims-blood-is-on-zuckerberg-s-hands-1.5404675> [<https://perma.cc/9SAS-T7Z7>].

¹¹ Shoshannah Solomon, *Israel’s Facebook Bill May Endanger Democracy, Company Official Implies*, TIMES ISR. (Jan. 18, 2017), <https://www.timesofisrael.com/israels-facebook-bill-may-endanger-democracy-company-official-implies/> [<https://perma.cc/C869-RXVG>].

quietly and behind closed doors,¹² was receptive to this public pressure. By September of 2016, Facebook had reached an informal agreement with the Israeli government to work together to address incitement on its platform.

While Facebook agreed to remove content that incites violence, promotes hate speech, or involves terrorism,¹³ deciding which content falls within the scope of these terms has proven a complicated task. News outlets reported that Israel and Facebook agreed to “create teams that would figure out how best to monitor and remove inflammatory content,” but no additional information was given to the public.¹⁴ This news troubled Palestinian users, who expressed concern that Facebook was “adopting Israeli policy and terminology when it comes to defining what incitement is.”¹⁵ With ninety-six percent of Palestinians reporting that their primary use for Facebook is following the news,¹⁶ decisions around who can access and exchange information on the platform are matters of great importance. They worried that this partnership, in effect, would lead not only to the removal of unlawful speech, but also stifle legitimate forms of dissent that conflict with the Israeli government’s narrative and perspective.

The stakes here are high. Facebook’s decisions about content removal and profile suspensions dictate the terms for participation in the “modern public square”¹⁷ and effectively arbitrate which narratives can reach the global public.¹⁸ In the context of the Israeli-Palestinian conflict, Facebook functions as a proxy battlefield, where longstanding geopolitical disputes are part and

¹² Facebook relies on private ordering and self-governance in shaping policy decisions that fit its business objectives. See discussion *infra* Part IV; see also Tarleton Gillespie, *Platforms Are Not Intermediaries*, 2 GEO. L. TECH. REV. 198, 202 (2018) (recognizing that “[t]oo often, social media platforms discuss content moderation as a problem to be solved—and solved privately and reactively.”).

¹³ Toor, *supra* note 7.

¹⁴ *Shaked: “Penny Has Dropped” for Facebook on Incitement*, TIMES ISR. (Sept. 12, 2016), <http://www.timesofisrael.com/shaked-penny-has-dropped-for-facebook-on-incitement/> [<https://perma.cc/F5WM-B4AZ>].

¹⁵ Matthew Ingram, *Facebook’s Censorship of Palestinian Journalists Raises Serious Questions*, FORTUNE (Sept. 28, 2016), <https://fortune.com/2016/09/28/facebook-censorship-palestinian/> [[perma.cc link unavailable](https://perma.cc/link-unavailable)].

¹⁶ *Palestinians Decry Increase in Arrests for “Incitement To Violence” on Social Media*, INDEPENDENT (Apr. 4, 2018), <https://www.independent.co.uk/news/world/middle-east/palestinians-israel-incitement-arrests-social-media-twitter-facebook-a8288631.html> [<https://perma.cc/33N3-2PEP>].

¹⁷ *Packingham v. North Carolina*, 137 S. Ct. 1730, 1732 (2017) (stating that barring an individual from accessing social media platforms functionally denies her access to “speaking and listening in the modern public square, and otherwise exploring the vast realms of human thought and knowledge”).

¹⁸ While content ranking and curatorial preferences influence these issues, this Note will focus only on content removal and profile suspension.

parcel of content moderation decisions that directly impact the nature of political discourse.

This Note examines Facebook's content moderation practices through the lens of the Israeli-Palestinian geopolitical conflict to highlight how Facebook's Community Standards and business practices—along with its informal relationships with individual nation states—obfuscate traditional allocations of responsibility, accountability, and power in democratic societies. This Note also challenges the inevitability of this construction by first examining the elements of Facebook's current content moderation practices that leave it vulnerable to manipulation, and then identifying opportunities to address these issues. Part II considers Facebook's content moderation practices in terms of its Community Standards and compliance with local law. Part III then locates enabling factors that make Facebook susceptible to biased execution of content moderation in relation to Israeli and Palestinian narratives. Finally, Part IV analyzes several conceptual frameworks to understand the new power dynamic among Facebook, individuals, and governments, positing steps to foster greater accountability and evenhandedness on the platform. Building on this analysis, Part IV also proposes that the platform shift away from a posture of neutrality and towards a stance informed by transparency and established human rights principles.

While the proliferation of extremism online is an important issue, this Note does not address organized terrorist activity on Facebook. Moreover, the experiences and case studies highlighted in this writing are included to illustrate the complex power dynamics at play on the platform; however, they by no means represent the only narratives or experiences impacted by Facebook's content moderation practices in the region. Finally, this Note does not stake a position on a path forward in resolving Israeli-Palestinian territorial disputes. Rather, the analysis here aims to delineate sources of power exercised on Facebook's platform and demonstrate that transparency and an orientation towards human rights can improve the role that platform governance plays in relation to geopolitical conflict.

II. AN OVERVIEW OF FACEBOOK'S COMMUNITY STANDARDS: GUIDELINES, IMPLEMENTATION, AND GOVERNANCE IN ISRAELI AND PALESTINIAN TERRITORIES

Facebook's Community Standards are guidelines that inform content removal and profile suspension decisions on the platform. Their stated goal "is to encourage expression and create a safe environment" for users to

connect and communicate.¹⁹ The Community Standards are organized into six categories: (1) violence and criminal behavior;²⁰ (2) safety;²¹ (3) objectionable content;²² (4) integrity and authenticity;²³ (5) respecting intellectual property;²⁴ and (6) content-related requests.²⁵ While the publicly available Community Standards serve as general guidelines, the company also circulates internal memoranda and training documents to moderators that more concretely instruct moderators' decision-making.²⁶ The Community Standards remain especially important because they are the primary lens through which Facebook considers content moderation decisions absent legal requirements specific to a particular regional jurisdiction. Consequently, the Community Standards are an authority cited to justify restricting speech that are distinct from public law. In an effort to understand the way content moderation decisions are made in geopolitically contested territories, Part II examines Facebook's public Community Standards, internal guidance documents, and the conflation of Facebook's private governance and its compliance with local law.

A. Facebook's Publicly Available Community Standards Are Articulated in Broad Terms to Facilitate Global Applicability, but Their Acontextual

¹⁹ *Community Standards*, FACEBOOK (2019). <https://www.facebook.com/communitystandards/introduction> (accessed Oct. 30, 2019) [<https://perma.cc/C75T-EWH3>].

²⁰ *Violence and Criminal Behavior*, FACEBOOK (2019), https://www.facebook.com/communitystandards/violence_criminal_behavior (accessed Nov. 20, 2019) [perma.cc/429J-NAQB].

²¹ *Safety*, FACEBOOK (2019), <https://www.facebook.com/communitystandards/safety> (accessed Nov. 20, 2019) [<https://perma.cc/EC86-PPYK>].

²² *Objectionable Content*, FACEBOOK (2019), https://www.facebook.com/communitystandards/objectionable_content (accessed Nov. 20, 2019) [<https://perma.cc/M37B-6R94>].

²³ *Integrity and Authenticity*, FACEBOOK (2019), https://www.facebook.com/communitystandards/integrity_authenticity (accessed Nov. 20, 2019) [<https://perma.cc/4GH7-6PDV>].

²⁴ *Respecting Intellectual Property*, FACEBOOK (2019), https://www.facebook.com/communitystandards/respecting_intellectual_property (accessed Nov. 20, 2019) [<https://perma.cc/GRB9-N3UM>].

²⁵ *Content-Related Requests*, FACEBOOK (2019), https://www.facebook.com/communitystandards/content_related_requests (accessed Nov. 20, 2019) [<https://perma.cc/M66M-7QAT>].

²⁶ Max Fisher, *Inside Facebook's Secret Rulebook for Global Political Speech*, N.Y. TIMES (Dec. 27, 2018), <https://www.nytimes.com/2018/12/27/world/facebook-moderators.html> [<https://perma.cc/9759-SHJH>]; Nick Hopkins, *Revealed: Facebook's Internal Rulebook on Sex, Terrorism and Violence*, GUARDIAN (May 21, 2017), <https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence> [<https://perma.cc/A3VC-VC8F>].

Framing Invites Subjective and Arbitrary Interpretations that Fill Policy Gaps.

The Community Standards constitute a single set of rules applied to every country in which Facebook operates, meaning that content moderators endeavor to apply the same set rules to Facebook's 2.3 billion monthly users around the world regardless of social or political context.²⁷ By virtue of their global applicability, Facebook's public-facing documents employ broad, flexible language. While this language is geared toward addressing myriad circumstances, its utility is limited by the lack of context-specific insight needed to implement evenhanded, consistent enforcement.²⁸

The Community Standards' high level of generality and lack of detail leave content moderators ill-equipped to navigate difficult questions in discrete circumstances. The Standards, for example, provide that the platform will "remove content that expresses support or praise for groups, leaders, or individuals involved" in terrorist or criminal activities or organized violence.²⁹ It is not clear, however, how these rules might apply to posts that support or praise non-violent and non-terrorist humanitarian aid efforts that also share a connection to extremist or terrorist groups. For instance, Facebook's guidelines do not provide insight as to how a post that praises programs funded by the Holy Land Foundation, a Muslim charity that supports international relief programs affiliated with Hamas, might fare under its Community Standards.³⁰ These complicated dynamics require context-specific knowledge, which Facebook has not accounted for in its content moderation practices (at least insofar as they have been explained to the public).

The lack of formal standardization is also apparent when considering Facebook's range of possible consequences for violating the Community Standards. Facebook states that users who violate Community Standards may

²⁷ Casey Newton, *The Trauma Floor: The Secret Lives of Facebook Moderators in America*, VERGE (Feb. 25, 2019), <https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona> [<https://perma.cc/2PCE-722C>].

²⁸ This applies both to helping users understand the requirements for compliance to be able to engage on the platform and to instructing moderators on which posts are impermissible in particular circumstances.

²⁹ *Violence and Criminal Behavior*, FACEBOOK (2019), https://www.facebook.com/communitystandards/violence_criminal_behavior (accessed Oct. 30, 2019) [<https://perma.cc/K4GQ-S2G6>].

³⁰ Fazia Patel & Rachel Levinson-Walsman, *Facebook and Free Speech*, JUST SECURITY (May 24, 2018), <https://www.justsecurity.org/56864/facebook-takedown-rules-raise-questions-require-transparency/> [<https://perma.cc/D25F-QX38>]; see generally Holy Land Found. for Relief and Dev. v. Ashcroft, 333 F.3d 156 (D.C. Cir. 2003); Laurie Goodstein, *U.S. Muslims Taken Aback by a Charity's Conviction*, N.Y. TIMES (Nov. 25, 2008), <https://www.nytimes.com/2008/11/26/us/26charity.html> [<https://perma.cc/EQA6-P49P>].

experience different consequences depending on the severity of the offense and the user's "history on the platform."³¹ Here, history seems primarily to refer to the frequency with which an account has violated the Community Standards.³² However, that explanation does not consider how seemingly random or inconsistent enforcement actions may influence frequency, nor does it account for whether erroneous content flagging or removals contribute to one's history on the platform.

In a similar vein, Facebook tells users that, based on severity, the platform may provide content to law enforcement in an effort "to prevent real-world harm."³³ The guidelines, though, do not elaborate on what such instances might entail, other than that law enforcement may be engaged where Facebook deems it is appropriate. This objective is not inherently bad: indeed, one could argue this demonstrates responsible platform governance. However, in the context of Israel and Palestine, provisions like this one remain a source of concern in balancing individuals' safety and autonomy. An effort by Facebook officials to elaborate upon the circumstances in which they coordinate with law enforcement would help mitigate speculation and suspicion discussed later in this analysis.

B. Facebook's Internal Guidelines Reveal an Evolving, Patchwork Approach to Content Moderation that Contributes to Uneven Enforcement of the Community Standards.

While the public Community Standards lack culture-specific considerations, leaked internal guidelines illustrate ways that double standards and biases can shape content moderation patterns.³⁴ Although Facebook's internal content moderation guidelines are more specific than publicly available resources, content moderators report that internal materials are frequently changed on an ad hoc basis.³⁵ Moderators have also claimed that there is not a reliable, uniform handbook, "master file[,] or overarching guide"

³¹ *Community Standards*, *supra* note 19.

³² *Id.*

³³ *Community Standards Enforcement Report*, FACEBOOK (2019), <https://transparency.facebook.com/community-standards-enforcement> (accessed Oct. 30, 2019) [<https://perma.cc/K4AJ-F748?type=image>].

³⁴ Fisher, *supra* note 26.

³⁵ Newton, *supra* note 27 ("While official policy changes typically arrive every other Wednesday, incremental guidance about developing issues is distributed on a near-daily basis."); *see also* Fisher, *supra* note 26.

to reference as new changes are announced.³⁶ Consequently, for any given enforcement decision, moderators find that they have several sources of “truth to consider” from within the organization’s internal guidance documents, which contributes to inconsistent moderation outcomes.³⁷ These constantly evolving and malleable guidelines contribute to the dissimilar treatment of similar content posted on the platform.³⁸ A 2017 ProPublica investigative report found that training materials for Facebook employees “banned posts that praise the use of ‘violence to resist occupation of an internationally recognized state.’”³⁹ It is noteworthy that the inverse scenario, praise of violence against those under occupation, is not expressly banned. Facebook reasoned that it adopted this rule because it did not want content moderators “to be in a position of deciding who is a freedom fighter.”⁴⁰ While the platform reported that it dropped the rule in 2017, its procedural and substantive approach to content moderation reveals how it can be manipulated by either individual preferences or majority-rule mentality.⁴¹

At a macro level, the platform’s fluid approach may seem justifiable due to the challenges inherent to content moderation at this scale or the need to rapidly respond to unanticipated, emerging conflicts. However, absent culturally and politically specific considerations, policy gaps leave an opening for bias to go unchecked. For instance, Facebook’s 2017 training materials instructed moderators to identify hate speech using a simple formula: “protected category + attack = hate speech.”⁴² Here, protected categories included race, gender, and religion, but the guidelines give greater latitude to

³⁶ Fisher, *supra* note 26 (“Facebook says the files are only for training, but moderators say they are used as day-to-day reference materials.”); *see also* Newton, *supra* note 27 (“Often, this guidance is posted to Workplace, the enterprise version of Facebook that the company introduced in 2016. Like Facebook itself, Workplace has an algorithmic News Feed that displays posts based on engagement. During a breaking news event, such as a mass shooting, managers will often post conflicting information about how to moderate individual pieces of content, which then appear out of chronological order on Workplace. Six current and former employees told me that they had made moderation mistakes based on seeing an outdated post at the top of their feed.”).

³⁷ Newton, *supra* note 27.

³⁸ Fisher, *supra* note 26. For research in the context of Israeli and Palestinian experiences, *see* TAMLEH, *supra* note 6.

³⁹ Julia Angwin & Hannes Grassegger, *Facebook's Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children*, PROPUBLICA (June 28, 2017), <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms> [<https://perma.cc/A9TJ-RP8B>].

⁴⁰ *Id.*

⁴¹ *Id.*; *see also* Julie E. Cohen, *Law for the Platform Economy*, 51 U.C. DAVIS L. REV. 133, 151 (2017) (“[P]latform affordances for volatility, polarization, and relativization are easily manipulated for malicious or simply self-interested purposes.”).

⁴² Angwin & Grassegger, *supra* note 39.

posts that only refer to subsets of protected categories. Thus, white men enjoy greater protection than female drivers or black children because gender and race are both protected, but occupation and age are not.⁴³ As such, these policies may afford fewer protections to vulnerable groups than those it provides to less frequent targets of hate speech. Because hate speech is difficult to define acontextually, the internal guidelines can exacerbate, and potentially magnify, power disparities, through their allocation of protection and censorship.

C. The Complex Legal Landscape Within Israeli and Palestinian Territories Creates Incentives for Facebook to Streamline Content Moderation Through Their Community Standards, Disproportionately and Negatively Impacting Specific Communities.

The legal complexity and geopolitical nuances specific to Israeli and Palestinian governance over physical territories create incentives for Facebook to rely on its own policies in making governance decisions in the region. Facebook's content moderation practices flow through two intersecting channels: (1) the Community Standards outlined above; and (2) the enforcement of public law specific to the region in which the platform operates. Put differently, the platform removes content or suspends accounts either because the account violated the Community Standards or because it violated local law.

Further, the variety of legal regimes governing Palestinian territories create a complex and challenging web of legal obligations. Palestinians living in Eastern Jerusalem are subject to laws enforced by Israeli civil courts, which rely on Articles from the penal code to address "incitement to violence and terrorism."⁴⁴ Meanwhile, Palestinian communities under Israeli occupation in regions like Gaza and the West Bank are subject to Israeli military law, along with the laws of the Palestinian Authority.⁴⁵ Israeli military law provides for greater prosecutorial discretion, expanding charges of incitement to include "those who express sympathy with terrorist activities," which grants the military authority to take action against conduct that may otherwise seem lawful.⁴⁶ Economic efficiency and ease of application create strong incentives for Facebook to rely on its own Community Standards rather than apply local

⁴³ *Id.*

⁴⁴ TAMLEH, *supra* note 6.

⁴⁵ *Id.*

⁴⁶ *Id.*

law.⁴⁷ While conflating the Community Standards with legal requirements streamlines Facebook's compliance burdens, it also undermines the protections properly afforded in a democratic society.

Patterns of inconsistency in applying the Community Standards reveal how power dynamics permeate platform governance. In examining Facebook's documents advising moderators on differentiating between hate speech and political expression, investigators found that "at least in some instances, the company's hate-speech rules tend[ed] to favor elites and governments over grassroots activists and racial minorities."⁴⁸ The Community Standards' lack of specificity or situational awareness leave significant policy determinations vulnerable to arbitrariness or the magnification of existing power struggles. As Professor Hannah Bloch-Wehba explains, because the public "lacks key information about how, when, and at whose direction platform governance is taking place, it is extremely difficult for the outside observer to discern what is going on, and to distinguish private action from government pressure."⁴⁹

III. CONSIDERING THE ENABLING FACTORS: THE CHALLENGES OF ALLOCATING RESPONSIBILITY AND AGENCY TO A PRIVATE PLATFORM IN EFFORT TO PROTECT INDIVIDUAL EXPRESSIVE INTERESTS

Democratic institutions and norms are underpinned by procedural and substantive modes of accountability that aim to hold those empowered to

⁴⁷ Decisions to remove content that does not violate Community Standards can stem from Facebook's desire to placate the governments that can regulate or penalize them, whereas protecting legitimate but unpopular speech may inspire backlash or negative PR. "The easiest, cheapest, and most risk-avoidant path for any technical intermediary is simply to process a removal request and not question its validity. A company that takes an 'if in doubt, take it down' approach to requests may simply be a rational economic actor." Daphne Keller, *Empirical Evidence Of "Over-removal" By Internet Companies Under Intermediary Liability Laws*, CTR. INTERNET & SOC'Y STAN. L. SCH. (Oct. 12, 2015), <http://cyberlaw.stanford.edu/blog/2015/10/empirical-evidence-over-removal-internet-companies-under-intermediary-liability-laws> [https://perma.cc/K5BC-SL7T].

⁴⁸ Angwin & Grassegger, *supra* note 39; *see also* Fisher, *supra* note 26 (reporting that "Facebook instructed moderators to 'look out for' the phrase 'Free Kashmir'—though the slogan, common among activists, is completely legal.").

⁴⁹ Hannah Bloch-Wehba, *Global Platform Governance: Private Power in the Shadow of the State*, 72 SMU L. REV. 27, 79 (2019).

government accountable to those living under their rule of law.⁵⁰ The Israeli government is a parliamentary democracy with legislative, judicial, and executive branches.⁵¹ While Facebook is not a sovereign or democratic institution, the company continuously positions and prides itself as a facilitator of free speech and proponent of democratic values.⁵² However, the coordination and amorphous relationship between Facebook and Israel reconceptualizes the exercise of governing power and challenges existing democratic modes of accountability. In connection to Israeli and Palestinian users, suspected double standards, lack of procedural transparency, algorithmic bias, and covert coordination all undermine trust and accountability among users in relation to the platform and the government.

A. Suspected Double Standards in Enforcement of the Community Standards Increase Concerns of Targeted Censorship and Undermine Users' Trust in Facebook.

Absent oversight, content removal and account suspensions can function as extensions of state action hidden under the aegis of Community Standards. For example, the agreement between Israel and Facebook neutralized government efforts to pass formal legislation, but it also marked an increase in content removal and account suspensions among Palestinian users.⁵³ Neither Facebook nor Israeli officials provided details about what their cooperation agreement entailed, though both parties indicated a preference for Facebook's voluntary removal of content over formal state

⁵⁰ “[I]dealistic pronouncements about the redemptive power of democratic politics and democratic constitutionalism have become increasingly difficult to credit. But certain high-level constraints on institutional behavior—and in particular the principles of separation of powers, procedural due process, and public reason—have commanded widespread adherence in democratic societies and have limited arbitrary exercises of official power.” JULIE E. COHEN, *BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM* 219 (2019) (ebook).

⁵¹ *A Free People in Our Land: Israel as a Parliamentary Democracy*, ISR. MINISTRY OF FOREIGN AFF. (Apr. 1, 2005), <https://mfa.gov.il/mfa/aboutisrael/state/pages/a%20free%20people%20in%20our%20land-%20israel%20as%20a%20parliamentary%20democracy.aspx> [https://perma.cc/QL3B-ZQHE].

⁵² Mark Zuckerberg, *A Conversation on Free Expression with Mark Zuckerberg at Georgetown University* (Oct. 17, 2019).

⁵³ Bethan McKernan, *Facebook “Deliberately Targeting” Palestinian Accounts After Meeting with Israeli Government, Rights Groups Say*, INDEPENDENT (Oct. 24, 2016), <https://www.independent.co.uk/news/world/middle-east/israel-palestine-facebook-activist-journalist-arrests-censorship-accusations-incitement-a7377776.html> [https://perma.cc/AV37-NGND].

action.⁵⁴ Tensions came to a head when seven prominent Palestinian journalists found that Facebook suspended their personal accounts. Facebook claimed that the accounts were reported for violating the site's community standards and mistakenly suspended, but Palestinian outlets perceived the incident as related to Israel's recent push to combat incitement. The journalists noted that, in addition to their account suspensions, Facebook removed content from their news pages, including material that was non-political.⁵⁵ In response to public outcry, Facebook reinstated the journalists' profiles and apologized for the "error" without offering insight as to how such a mistake occurred.⁵⁶ This is significant because the platform's unilateral decision to remove content and suspend user profiles limited the journalists' ability to communicate with the public. Yet, by citing the Community Standards rather than local legal strictures to make these decisions, interested parties had no established means by which to interrogate Facebook's reason for removing the content any further.

Facebook's equivocal response about the journalists' profile suspension raised concerns among Palestinian Facebook users and activists about bias and censorship of their perspectives.⁵⁷ Absent evidence to the contrary, Palestinian advocates viewed these events as "a dangerous escalation" of Facebook's approach to content moderation.⁵⁸ The platform defines hate speech as "anything that directly attacks people based on what are known as their 'protected characteristics'—race, ethnicity, national origin, religious affiliation, sexual orientation, sex, gender, gender identity, or serious disability or disease."⁵⁹ This definition covers hate speech directed at both Palestinian and Israeli individuals; however, some journalists point to enforcement double standards as indicative of Facebook's bias against Palestinian users, claiming that "[c]alls by Israelis for the killing of Palestinians are commonplace on Facebook, and largely remain

⁵⁴ *Why Facebook and Google Are Complying with Israel to Delete Certain Content*, FORTUNE (Sept. 12, 2016), <http://fortune.com/2016/09/12/facebook-google-israel-social-media/> [<https://perma.cc/3QY9-YQ39>].

⁵⁵ Patel & Levinson-Walsman, *supra* note 30.

⁵⁶ Amar Toor, *Facebook Accused of Censoring Palestinian Journalists*, VERGE (Sept. 26, 2016), <https://www.theverge.com/2016/9/26/13055862/facebook-israel-palestinian-journalists-censorship> [<https://perma.cc/Y2YQ-AMD4>].

⁵⁷ Ylenia Gostoli, *Is Facebook Neutral on Palestine-Israel Conflict?*, AL JAZEERA (Sept. 26, 2016), <https://www.aljazeera.com/news/2016/09/facebook-neutral-palestine-israel-conflict-160921115752070.html> [<https://perma.cc/898Z-YJCC>].

⁵⁸ TAMLEH, *supra* note 6, at 1.

⁵⁹ Richard Allan, *Hard Questions: Who Should Decide What Is Hate Speech in an Online Global Community?*, FACEBOOK NEWSROOM (June 27, 2017), <https://newsroom.fb.com/news/2017/06/hard-questions-hate-speech/> [<https://perma.cc/S7LQ-NYPV>].

undisturbed.”⁶⁰ Palestinian users, meanwhile, experience disproportionate levels of content removal and find that violent rhetoric or hateful posts by Israelis against Palestinians frequently go undetected by content moderators.⁶¹

Facebook’s use of the Community Standards in these instances also demonstrates the ramifications of exercising soft power. In the context of regulating inciting and terroristic speech online, the distinction between state and private actors grew increasingly blurry after the Israeli government publicly exerted substantial political pressure on Facebook to remove content pursuant to Facebook’s own Community Standards rather than in terms of what is required by Israeli law.⁶² In so doing, it raised suspicions that the Community Standards functioned as a veil for state censorship.

⁶⁰ Glenn Greenwald, *Facebook Says It Is Deleting Accounts at the Direction of the U.S. and Israeli Governments*, INTERCEPT (Dec. 30, 2017), <https://theintercept.com/2017/12/30/facebook-says-it-is-deleting-accounts-at-the-direction-of-the-u-s-and-israeli-governments/> [<https://perma.cc/4JGP-3YXQ>]; see also *Is Palestinian-Israeli Violence Being Driven by Social Media?*, BBC NEWS (October 22, 2015), <https://www.bbc.com/news/world-middle-east-34513693> (stating that inciteful posts by Israelis increased after Israeli-Palestinian violence) [<https://perma.cc/5CGM-A9QM>]; Ofra Edelman, *Internet Incitement Against Arabs in Israel on the Rise*, HAARETZ (Oct. 13, 2015), <https://www.haaretz.com/.premium-anti-arab-internet-incitement-rising-1.5408041> (“Since the latest wave of violence began there has been a sharp rise in the number of statements inciting to violence against Arabs, and in the number of Facebook pages expressing extremist right-wing positions receiving ‘likes,’ according to experts on monitoring Internet discourse.”) [<https://perma.cc/P9RG-URKH>].

⁶¹ One report found that, in 2018, hate speech against Palestinians was published on social media platforms “every 66 seconds.” 7AMLEH, #ASHTAG PALESTINE 2018: AN OVERVIEW OF DIGITAL RIGHTS ABUSES OF PALESTINIANS 16 (2019), https://7amleh.org/wp-content/uploads/2019/03/Hashtag_Palestine_English_digital_pages.pdf [<https://perma.cc/X29E-R9KD>]. See also *The Index of Racism and Incitement in Israeli Social Media 2018: An Inciting Post Against Palestinians Every 66 Seconds*, 7AMLEH (Mar. 11, 2019), <https://7amleh.org/2019/03/11/the-index-of-racism-and-incitement-in-israeli-social-media-2018-an-inciting-post-against-palestinians-every-66-seconds/> [<https://perma.cc/C6MV-8YXN>]; Ruth Eglash, *An Arab, a Jew and a Facebook Post: How Similar Words Are Treated Differently*, WASH. POST (July 15, 2016), https://www.washingtonpost.com/world/middle_east/an-arab-a-jew-and-a-facebook-post-how-similar-words-are-treated-differently/2016/07/14/e346ef1c-47a4-11e6-8dac-0c6e4acc5b1_story.html?utm_term=.e48623dc1772 [<https://perma.cc/NM2M-X5A8>].

⁶² Toor, *supra* note 7.

Geopolitical indeterminacy tilts in favor of those who already have power and influence,⁶³ which, in this case, translates to compliance with Israel's definition of incitement. For example, Facebook received a request from the Israeli State Attorney's Office asking it to remove two pages that the Minister of Defense attributed to an unlawful association.⁶⁴ While Facebook determined that the pages themselves did not violate the Community Standards, the platform nevertheless decided to restrict access to both pages in Israel based on the request from the Israeli State Attorney's Office.⁶⁵ Here, geofencing within the region translates to restrictions on Palestinian access to public discourse within their communities. This case illustrates how Facebook's desire to comply with governmental requests for removal can result in expanding the legal reach and authority of state actors without firm grounding in law. In these circumstances, the content may well have been inciting or dangerous. However, the removed content could have simply been something the Israeli government found disagreeable. Facebook has strong incentives to remove posts when requested to avoid public blame or formal regulation by a government. Consequently, the coordination between government and digital platforms magnifies the power exercised by state actors while insulating it from public scrutiny.

B. Facebook's Opacity Eschews Democratic, Procedural Norms That Establish the Scope of Permissive Forms of Public Expression.

Procedural transparency of government actions serves an important role in democratic society, separate and distinct from substantive evaluations

⁶³ See Cohen, *supra* note 41, at 220. ("Network-and-standard-based legal-institutional arrangements connect protocol and policy directly to one another and eliminate separation between them. Within such arrangements, the point of mandated standardization is exactly to specify the kinds of flows that must, may, and may not travel via the network. The policy is the standard and vice versa. Power over one translates directly into power over the other. Under background conditions of vastly unequal geopolitical power, that equivalence sets up the two interlocking dynamics that produce policy hegemony . . . And policy hegemony is power that may be exercised without regard for the basic, high-level rule-of-law constraints that obtain in more traditional institutional settings."); See also Julia Angwin & Hannes Grassegger, *Facebook's Secret Censorship Rules Protect White Men from Hate Speech But Not Black Children*, PROPUBLICA (June 28, 2017), <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms> ("at least in some instances, the company's hate-speech rules tend[ed] to favor elites and governments over grassroots activists and racial minorities."). [<https://perma.cc/A9TJ-RP8B>].

⁶⁴ *Content Restrictions Based on Local Law*, FACEBOOK TRANSPARENCY (2019), <https://transparency.facebook.com/content-restrictions> (accessed Nov. 26, 2019) [<https://perma.cc/B3JN-XY4Q>].

⁶⁵ *Id.*

of permissible forms of expression and dissent. Put differently, inquiry into the procedural adequacy of state action is analytically distinguishable from debate over substantive line drawing between protected and unprotected speech. In contrast to Facebook's removal of content and account suspensions, the Israeli government took tangible action against Palestinian poet Dareen Tatour in response to a series of her social media posts by arresting and charging her for incitement of terrorism and support of terrorist organizations.⁶⁶ The charges relied on three posts, including a poem about the Knife Intifada, along with references to her opposition to the Israeli occupation. Her indictment contained an interpretation of her poem, specifically including the lines "I will not succumb to the 'peaceful solution' / Never lower my flags / Until I evict them from my land."⁶⁷ Although there was fervent dispute around how the poem's language should be interpreted and contextualized throughout the trial, Tatour was convicted and sentenced to several months in prison for inciting violence and supporting a terrorist organization based on her social media posts.⁶⁸

The government's decision to convict Tatour based on her social media posts sparked criticism and scrutiny from international literary groups, advocates, and individuals from around the world.⁶⁹ Critics of her conviction connected this case to an ongoing pattern of Israeli arrests of Palestinians for terrorism-related charges based on social media posts; others pointed out a double standard, asserting that Jewish Israeli poets that made similar, explicit calls for violence on social media do not face legal consequences.⁷⁰ Despite the many controversial aspects of this case, Tatour's experience highlights the value of transparency as a function of assigning responsibility and accountability, even in circumstances where public expression is punishable. This stands in contrast to instances where limitations on speech and the justification for removing it from the public square remain hidden.

⁶⁶ *Offline: Dareen Tartour*, ELECTRONIC FRONTIER FOUND. (2018), <https://www.eff.org/offline/dareen-tatour> [<https://perma.cc/5D3H-3892>].

⁶⁷ Noa Shpigel, *Israeli Arab Poet Dareen Tatour Gets Five-month Sentence for Incitement on Social Media*, HAARETZ (July 31, 2018), <https://www.haaretz.com/israel-news/israel-hands-palestinian-poet-dareen-tatour-five-month-prison-sentence-1.6335232> [<https://perma.cc/KU8W-G28W>].

⁶⁸ *Dareen Tatour: Israeli Arab poet convicted of incitement*, BBC NEWS (May 3, 2018), <https://www.bbc.com/news/world-middle-east-43990577> [<https://perma.cc/F2E2-U374>].

⁶⁹ Israel: End judicial proceedings and ensure Dareen Tatour's immediate release, PEN INT'L (July 2, 2018), <https://pen-international.org/news/israel-end-judicial-proceedings-and-ensure-dareen-tatours-immediate-release> [<https://perma.cc/3JF9-SVHH>]; *Artists all over the world express their solidarity by works of art based on the poem "Resist,"* FREE DAREEN, <https://freedareentatour.org/poems> [<https://perma.cc/A8MA-3BK7>].

⁷⁰ Electronic Frontier Found., *supra* note 66.

Importantly, the arrest, trial, and ensuing debate reveal the subjective nature of interpreting words like *incitement* and *terrorism*. While many people disagree about the scope of protection for expressions of dissent speech, hate speech, inciteful speech, and harassment in public discourse, a cornerstone of democracy is protecting and permitting expressions of legitimate dissent (though the line of legitimacy is often drawn in different places). The conclusion that Tatour's poem social media posts fit within the meaning of incitement and support terrorism turns on subjective judgments reflecting a set of cultural norms and priorities. To that end, the inquiry as to whether Tartour's speech appropriately fits within the scope of these charges depends on which narrative one chooses to accept. Moreover, the governing body that decides which expressive content satisfies the definitions of incitement and terrorism simultaneously has authority to determine the boundaries of permissible speech within which one can engage in the public sphere.

Throughout Tatour's case, the process of trial, adjudication, and conviction provided mechanisms requiring transparency around the charges that informed the government's restrictions on speech: prosecutors had to posit charges and support them with corresponding facts.⁷¹ This furnished the public with tools to identify the disputed content, interrogate the legitimacy of the Government's charges, and follow the process of adjudication. With this information, individuals could make judgments about the government's decisions as they related to their own political sensibilities and advocate around the issue. Transparency does not diminish the real and significant consequences that Tatour experienced for posting her poems. There is, however, an ability to trace which parties exercised power and the institutions where accountability, critique, and recourse might be directed. After the trial, advocacy around Tatour's conviction resulted in an early release from her sentence.⁷²

The procedural transparency highlighted in Tatour's case stands in contrast to the processes involved in content moderation decisions on Facebook's platform. Facebook's intervention in content removal and profile suspensions displaces conventional sources of transparency and process because its actions exist outside the formal exercise of law. While these decisions may be uncoordinated and ad hoc, the platform functions as a vehicle by which fundamental democratic rights can be exercised or suppressed. Consequently, content moderation decisions implicate important democratic functions and carry with them the capacity to reshape the dynamics of public

⁷¹ This analysis makes no evaluative judgments concerning the government's substantive interpretation and application of law in adjudicating these charges.

⁷² *Israel: Poet Dareen Tatour Released from Prison*, PEN INT'L (Sept. 20, 2018), <https://pen-international.org/news/israel-poet-dareen-tatour-released-from-prison> [<https://perma.cc/WN6S-JVQK>].

discourse. This creates a new mechanism through which Palestinian and Israeli authorities can vie for influence and an area where Facebook can allow political pressures and other soft influences to color the lens through which it interprets its own Community Standards.

C. Algorithmic Flagging and Artificial Intelligence (AI) Tools Heighten, Rather Than Mitigate, Unequal Levels of Suspicion and Surveillance.

Facebook uses a combination of algorithms and human oversight to moderate activity on its platform. The scale of content posted to Facebook requires the use of algorithms to organize and rank posts on individuals' newsfeeds, and content moderation algorithms are also used to flag and proactively remove problematic content before it is reported as an issue by a user.⁷³ Facebook uses these algorithms to identify, flag, and remove hate speech quickly and efficiently, yet the efficacy of these algorithms depends on the material used to train the algorithm, along with the scope of the algorithm's application and monitoring of its outputs. In the context of Israeli and Palestinian experiences on Facebook, algorithmic flagging and AI tools heighten—rather than mitigate—geopolitical tensions based on the limitations of machine learning and minimal oversight.

Bias can surface in many ways, and the nuances and challenges of content moderation relevant to Israeli and Palestinian users present difficult questions that algorithms may be fundamentally ill-suited to resolve. Studies on algorithmic fairness, accountability, and transparency suggest that mathematical formulas and algorithmic outputs may not always be appropriate tools to evaluate issues heavily dependent on social or historical context or questions of fairness.⁷⁴ In terms of both computer science and cultural sensitivity, a one-size-fits-all approach to content moderation creates a fundamentally flawed process by which to tackle thorny, context-specific issues in contested territories. And, at its core, Facebook's application of its flagging algorithms may not provide adequate sensitivity to minority perspectives or account for cultural nuances specific to contentious dynamics. The constitutive qualities of inciting speech cannot be distilled or determined in a vacuum because they vary based on sensitivities informed by a multitude

⁷³ FACEBOOK, FACEBOOK'S CIVIL RIGHTS AUDIT PROGRESS REPORT 7–8 (2019), https://fbnewsroomus.files.wordpress.com/2019/06/civilrightaudit_final.pdf [<https://perma.cc/C9LE-SJTF>]; Nathaniel Gleicher, *Removing Bad Actors from Facebook*, FACEBOOK NEWSROOM (June 26, 2018), <https://newsroom.fb.com/news/2018/06/removing-bad-actors-from-facebook/> [<https://perma.cc/8TYK-GHMU>].

⁷⁴ See generally Andrew D. Selbst et al., *Fairness and Abstraction in Sociotechnical Systems*, in PROCEEDINGS OF THE CONFERENCE ON FAIRNESS, TRANSPARENCY, AND ACCOUNTABILITY 2019 59 (forthcoming), <https://dl.acm.org/citation.cfm?id=3287598> [<https://perma.cc/7RBG-PMVG>].

of factors, including history, setting, and circumstance. Thus, universally applied algorithms are likely to produce erroneous outcomes in this area, either missing or mischaracterizing speech in an array of fact-specific circumstances.

Algorithmic bias contributes to the sense of unequal treatment and targeting that Palestinians report to experience on social media. For instance, Israeli law enforcement arrested a Palestinian man based on an error in Facebook's automatic translation program.⁷⁵ In that case, a construction worker in the West Bank posted a picture of himself alongside a bulldozer with the caption "يصبحهم" or "yusbihuhum," which translates as "good morning" in Arabic.⁷⁶ Facebook's algorithm translated this caption as saying "hurt them," in English or "attack them," in Hebrew.⁷⁷ This translation, coupled with the fact that bulldozers had previously been weaponized as vehicles for hit-and-run terrorist attacks,⁷⁸ resulted in the man's arrest for posting a picture to his private Facebook account.⁷⁹ Police officers detained the man after receiving notification about the post, though at no point before his arrest did "any Arabic-speaking officer read the actual post."⁸⁰ After several hours of questioning, officers realized the mistake and released him.

Here, bias permeates content moderation in a manner that reinforces existing power disparities in three ways. First, the translation algorithm used to regulate and flag content initiated punitive measures before the translation was checked for accuracy. Second, the heightened and suspicious surveilling of Palestinian users on Facebook increased incidents of flagging that can lead to arrest. Put differently, the algorithm serves as a source of confirmation bias for suspicious moderators, rather than a tool to objectively identify and respond to content inciting harm. Third, even after the man was released and

⁷⁵ Alex Hern, *Facebook Translates "Good Morning" into "Attack Them," Leading to Arrest*, GUARDIAN (Oct. 24, 2017), <https://www.theguardian.com/technology/2017/oct/24/facebook-palestine-israel-translates-good-morning-attack-them-arrest> [https://perma.cc/3T3X-WCYE].

⁷⁶ *Id.*

⁷⁷ *Id.*

⁷⁸ *Id.*

⁷⁹ Yuval Noah Harari, *Why Technology Favors Tyranny*, ATLANTIC (Oct. 2018), <https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/> [https://perma.cc/TYF5-VS5L].

⁸⁰ *Id.*

Facebook apologized for the translation error,⁸¹ the original post remained removed from the platform after the incident.⁸² This demonstrates how biased outcomes, even when proven erroneous, can further marginalize minority voices.

D. Secret Flows of Information Between Private and Public Actors for the Purposes of Surveillance and Arrest Thwart Procedural Oversight and Legal Guardrails that Are Normatively Desirable and Functionally Expected.

Facebook's coordination with law enforcement has the potential to magnify police power in targeted and politically significant ways. In the translation incident discussed above, it is notable that Facebook and Israeli law enforcement's actions leading up to the arrest were nebulously comingled. News coverage on this story stated that the police "were notified"⁸³ of the post, though it is not clear which actor can be assigned responsibility for what action. It is unclear, for example, whether Facebook flagged the post and reported the content to law enforcement based on the translation, which is a possibility discussed in the Community Standards, or if Israeli law enforcement used its own algorithms and flagging tools that rely on the platform's translation capabilities. All that is publicly known about Facebook's involvement with the arrest discussed above is that the translation algorithm's mistake catalyzed the arrest. These circumstances raise questions concerning both the frequency and freeness by which information flows between Facebook and state actors.

Content management and moderation are essential functions of Facebook's business. Yet, the nature and amount of data sharing or surveillance the platform does on behalf of states frustrates traditional allocations of accountability and trust. Adalah, the Legal Center for Arab Minority Rights in Israel, reported that, in 2016, "82 percent of those arrested for incitement-related offenses were Palestinian citizens, whereas only 18

⁸¹ Gizmodo received the following statement from an engineering manager at Facebook: "Unfortunately, our translation systems made an error last week that misinterpreted what this individual posted. Even though our translations are getting better each day, mistakes like these might happen from time to time and we've taken steps to address this particular issue. We apologize to him and his family for the mistake and the disruption this caused." Sidney Fussell, *Palestinian Man Arrested After Facebook Auto-Translates "Good Morning" as "Attack Them,"* GIZMODO (Oct. 23, 2017), <https://gizmodo.com/palestinian-man-arrested-after-facebook-auto-translates-1819782902> [<https://perma.cc/9QTB-DY67>].

⁸² Harari, *supra* note 79.

⁸³ Yotam Berger, *Israel Arrests Palestinian Because Facebook Translated "Good Morning" to "Attack Them,"* HAARETZ (Oct. 22, 2017), <https://www.haaretz.com/israel-news/palestinian-arrested-over-mistranslated-good-morning-facebook-post-1.5459427> [<https://perma.cc/RJY4-Y9A9>].

percent were Israeli Jewish citizens.”⁸⁴ Despite documented concerns from both sides of the conflict about hateful or inciting content circulated on Facebook, the distribution of arrests reveals an imbalance in either platform detection or legal enforcement.⁸⁵ Thus, from the perspective of Palestinians, Facebook not only imposes a double standard concerning the application of its own Community Standards, but it may also contribute to the disproportionate number of Palestinian arrests related to user activity on the platform.

Democratic societies broadly recognize that unrestricted government surveillance contravenes individual rights and liberties.⁸⁶ As such, the unknown arrangement concerning flows of information between Facebook and Israeli authorities proves significant in two ways: (1) by raising suspicions that Facebook is acting as an extension of law enforcement; and (2) by thwarting oversight or procedural safeguards to protect from oppressive forms of government surveillance. Because government surveillance programs often balance national security priorities with individual civil liberties, procedural safeguards and oversight operate as guardrails to check abusive exercise of authority. Statutes like the General Data Protection Regulation (“GDPR”),⁸⁷ in the European Union, and the Foreign Intelligence Surveillance Act (“FISA”),⁸⁸ in the United States, illustrate different constructions of procedural and legal protections that strive to balance governments national security pursuits with democratic norms of procedural due process. While

⁸⁴ *Adalah Fears Facebook's Online Incitement Deal with Israel Will Selectively Target Palestinian Citizens*, ADALAH: LEGAL CTR. ARAB MINORITY RTS. ISR. (Sept. 11, 2016), <https://www.adalah.org/en/content/view/8948> [<https://perma.cc/CA7Z-MG24>].

⁸⁵ Eglash, *supra* note 61.

⁸⁶ Cohen, *supra* note 41, at 191–99.

⁸⁷ Commission Regulation 2016/679, art. 23, 2016 O.J. (L 119) (providing a national security exception in a manner that “respects the essence of the fundamental rights and freedoms and is a necessary and proportionate measure in a democratic society. . .”).

⁸⁸ 50 U.S.C. §§ 1801–1885(c) (2018).

these statutes are imperfect and subject to robust debate,⁸⁹ they provide some threshold requirements and limitations on government surveillance, along with modes of recourse for individuals whose rights are violated under the law. By way of contrast, suspected informal or secret flows of information between Facebook and Israel frustrate notions of democratic legitimacy and accountability.

Due to these enabling factors, governance on Facebook becomes a proxy battle in which disputing narratives around Israeli and Palestinian activity emerge and collide. Palestinians view Israel's efforts to control the content shared on Facebook as an extension of physical occupation to the digital space;⁹⁰ meanwhile, Israelis view Facebook as a platform that magnifies the reach of terrorist actors, who encourage violence against their

⁸⁹ Karl Manheim & Lyric Kaplan, *Artificial Intelligence: Risks to Privacy and Democracy*, 21 YALE J. L. & TECH. 106, 168 (2019) (“[C]ontrol, transparency, and accountability are running themes throughout GDPR.”); Peter Swire & DeBrae Kennedy-Mayo, *How Both the EU and the U.S. Are “Stricter” Than Each Other for the Privacy of Government Requests for Information*, 66 EMORY L.J. 617, 619–20 (2017) (“Based on our study of both the EU and U.S. systems, we believe there are generally effective rule-of-law protections against excessive law enforcement surveillance in both the U.S. and EU Member States. We therefore conclude that these generally effective safeguards provide a promising basis for MLA reform, even where details of the systems differ and specific safeguards on one side do not have precise counterparts on the other.”); Laura K. Donohue, *The Case for Reforming Section 702 of U.S. Foreign Intelligence Surveillance Law*, COUNCIL FOREIGN REL. (June 26, 2017), <https://www.cfr.org/report/case-reforming-section-702-us-foreign-intelligence-surveillance-law> (“Section 702 is an important tool in the intelligence community’s arsenal. But the statute should be amended to bring it within constitutional bounds.”) [<https://perma.cc/5K77-BATX>]; L. Rush Atkinson, *The Fourth Amendment’s National Security Exception: Its History and Limits*, 66 VAND. L. REV. 1343, 1396–1405 (2013) (providing an account of the evolution of FISA, intelligence warrants, and warrantless surveillance, as well as a discussion of the procedural and substantive limitations of FISA revealed in litigation).

⁹⁰ Press release, 7amleh, #Palestine 2017 Report: Palestinian Online Content Targeted Through Mass Surveillance, Digital Occupation and Biased Content Moderation (April 3, 2018), <https://7amleh.org/2018/04/03/press-release-palestine-2017-report-palestinian-online-content-targeted-through-mass-surveillance-digital-occupation-and-biased-content-moderation/> (“[T]he report focuses on the ‘digital occupation’ of Palestinian social media through the use of algorithms, mass surveillance, and tens of military and Secret Service-affiliated Facebook pages utilized to hinder Palestinian online activism and infringe on freedom of speech.”) [<https://perma.cc/HF6L-2J2N>].

citizens.⁹¹ The informal relationships and vague criteria that influence users' experiences on Facebook fuel suspicion and distrust on both sides of these opposing narratives. While Facebook balks at the notion that it exercises greater leniency with Israeli users compared to their Palestinian counterparts,⁹² the platform's amorphous relationship with the government, algorithmic backboxes, and secrecy concerning information flows undermine trust in the platform's credibility in their claims of neutrality. Further, the lack of transparency around content moderation leaves users with few tools to contextualize anecdotal experiences, which creates fodder for speculation without any means to dispel of it.⁹³ These factors together foster the real and perceived double standards and biases that permeate the platform unchecked.

IV. OPPORTUNITIES: DEFINING, DELINEATING, AND CLARIFYING THE ROLE OF CONTENT MODERATION IN GLOBAL AFFAIRS

Today, Facebook's involvement and entanglement with state governments undermines public trust, weakens accountability, and frustrates traditional protections of democratic norms along with human rights. However, the current construction is neither inevitable nor beyond repair. Facebook can act as a better, more responsible steward in preserving the rights of its users by taking steps to grapple with its role in public life, by delineating its actions from those of state actors, and by adopting a human rights orientation toward content moderation decisions.

⁹¹ Associated Press Jerusalem, *Facebook and Israel to Work to Monitor Posts that Incite Violence*, GUARDIAN (September 12, 2016), <https://www.theguardian.com/technology/2016/sep/12/facebook-israel-monitor-posts-incite-violence-social-media>, ("Israel has argued that a wave of violence with the Palestinians over the past year has been fueled by incitement, much of it spread on social media sites. It has repeatedly said that Facebook should do more to monitor and control the content...") [<https://perma.cc/5AVY-VGYW>]; Maayan Jaffe-Hoffman, *Israelis Incite Against Arabs on Social Media Every 66 Seconds—Report*, JERUSALEM POST (May 22, 2019), <https://www.jpost.com/Israel-News/Israelis-incite-against-Arabs-on-social-media-every-66-seconds-report-590410> ("The Israeli government has in the past accused Arab citizens of Israel and Palestinians of using social networks to incite terrorism against Israelis.") [<https://perma.cc/2XHQ-FN7D>].

⁹² Ylenia Gostoli, *Palestinians Fight Facebook, YouTube Censorship*, AL JAZEERA (Jan. 20, 2018), <https://www.aljazeera.com/news/2018/01/palestinians-fight-facebook-youtube-censorship-180119095053943.html> [<https://perma.cc/79Z7-G6PH>].

⁹³ Speculation or anecdotal evidence could be dispelled with data to contextualize user experiences, but Facebook has declined to explain content moderation determinations in many instances that seem to tilt political power. See Olivia Solon, *Facebook Declines to Say Why It Deletes Certain Political Accounts, But Not Others*, GUARDIAN (Jan. 4, 2018), <https://www.theguardian.com/us-news/2018/jan/04/facebook-chechnya-ramzan-kadyrov-political-censorship> [<https://perma.cc/6PVG-Z6BW>].

A. Understanding Facebook's Power As Both Distinct From and In Relation to Government.

First, to understand the power dynamics at play, Facebook's agency and power must be disentangled from those of state actors. Legal scholars posit a number of theories used to untangle the complex power dynamics active between private platforms and state actors. The taxonomy developed by this scholarship creates a structure for understanding intermediaries' power in relation to government and law. Within the existing frameworks, there is not consensus around the distribution or balance of power between Facebook and state actors. At times, scholarly work either understates Facebook's agency as too meek or grants it too noble intentions. For instance, Professor Jack Balkin characterizes this dynamic as state co-option and expansion of regulatory reach through the shadow of private companies: in his view, "companies act as a private bureaucracy that implements the state's speech policies."⁹⁴ "Jawboning"⁹⁵ describes private companies' capitulation to state authority, where, in the absence of concrete law, platforms create opportunities for shadow exercises of government power. However, this framework oversimplifies the complex and competing power dynamics at play because it discounts the amount of leverage Facebook can use to reach consensus with sovereign nations.

Facebook is not a meek organization outmatched by the pressure of the bully pulpit. Kate Klonick's "New Governors" framework recognizes the elastic influence that digital intermediaries exercise in regulating speech independent of state authorities, asserting that, "platforms are both the architecture for publishing new speech and the architects of the institutional design that governs it."⁹⁶ While this framework is instructive in identifying the autonomy platforms exercise over the construction and regulation of speech in alignment with their own normative values and interests, it may offer too generous a read on platforms' internalization of, or commitment to, democratic norms.⁹⁷

⁹⁴ Jack M. Balkin, *Free Speech is a Triangle*, 118 COLUM. L. REV. 2011, 2030 (2018).

⁹⁵ Jack M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 U. CAL. DAVIS L. REV. 1149, 1177 (2018) ("Nation states have a range of different strategies to exert pressure. They can impose fines or criminal penalties. They can threaten prosecution. Or they can engage in jawboning—urging digital infrastructure operators to do the right thing and block, hinder, or take down content.").

⁹⁶ See Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1617 (2018). See generally Gillespie, *supra* note 12.

⁹⁷ Klonick, *supra* note 96, at 1603 ("They are private, self-regulating entities that are economically and normatively motivated to reflect the democratic culture and free speech expectations of their users.").

It is worth noting that Facebook's business interests center on profit maximization, which requires workable, constructive, and efficient relationships with both individual users and the countries in which Facebook operates.⁹⁸ Facebook as a corporate enterprise neither anticipated nor internalized the responsibilities inherent to its business operations and has remained reluctant to take ownership over its role in reshaping political and democratic norms.⁹⁹ As such, assuming its commitment to democratic norms as applied to individual users may assume too much of Facebook's established track record of accountability as an independent governing agent. It seems that Facebook's approach to regulating speech more accurately reflects uncoordinated, ad hoc decisions to advance its self-interest at a given moment in time, rather than a systematic, comprehensive approach to governance that accounts for the consequences flowing from its decisions.¹⁰⁰ While state regulation traditionally defined the scope of protections and permissibility of public expression, state authority runs up against—rather than above or below—that of platform intermediaries.

Facebook and its peer platforms are best understood as operating with power equivalent to that of state actors. Indeed, these private entities possess tremendous amount of economic and social capital equivalent to or surpassing that of discrete state governments. In fact, private platform companies perceive themselves as “on par with, not subordinate to, governments, including those governments that attempt to regulate them.”¹⁰¹ This is evident in the ways that platform companies like Facebook increasingly posture in a manner reflecting diplomatic relations between sovereign nations, rather than viewing themselves as answerable to any one jurisdiction. Professor Julie Cohen points to how these legal and institutional dynamics have shifted, pointing out that, “Facebook's privacy team travels the world meeting with government officials to determine how best to satisfy their concerns while

⁹⁸ See Julie E. Cohen, *Information Platforms and the Law*, 2 GEO. L. TECH. REV. 191, 192–193 (2018) (“[P]latform firms are also discrete legal entities with interests and agendas of their own. Platform firms also rely on law and legal institutions to advance their own self-interested goals.”).

⁹⁹ Sheera Frenkel et al., *Delay, Deny and Deflect: How Facebook's Leaders Fought Through Crisis*, N.Y. TIMES (Nov. 14, 2018), <https://www.nytimes.com/2018/11/14/technology/facebook-data-russia-election-racism.html?module=inline> [<https://perma.cc/768C-C5LN>]; Abby Ohlheiser, *Mark Zuckerberg Denies That Fake News on Facebook Influenced the Elections*, WASH. POST (Nov. 11, 2016), <https://www.washingtonpost.com/news/the-intersect/wp/2016/11/11/mark-zuckerberg-denies-that-fake-news-on-facebook-influenced-the-elections/> [<https://perma.cc/3JP3-NY7B>].

¹⁰⁰ See Daphne Keller, *Facebook Restricts Free Speech by Popular Demand*, ATLANTIC (September 22, 2019), <https://www.theatlantic.com/ideas/archive/2019/09/facebook-restricts-free-speech-popular-demand/598462/> [<https://perma.cc/RGQ8-SDY2>].

¹⁰¹ Kristen E. Eichensehr, *Digital Switzerlands*, 167 U. PA. L. REV. 665, 665 (2019).

continuing to advance Facebook's own interests, much as a secretary of state and his or her staff might do.”¹⁰² Building on this notion, the agreements and coordinated efforts platforms reach with governments represent a new balance in the exchange of power, which carries significance for those governed by these two regimes.

Michael Birnhack and Niva Elkin-Koren describe these secret flows of information between state and platform intermediaries as “the invisible handshake,” characterized by “collaboration between law enforcement agencies and the private sector, beyond the reach of judicial review and away from the critical eye of public opinion.”¹⁰³ This description explains the way that power has shifted from traditional accountability guardrails to those with authority. Facebook’s adoption of policies and practices that appeal to a particular state’s sensibilities exports certain value judgments and normative standards across borders. Still, the platform has struggled to reckon with the implications these choices carry for its users in context-specific scenarios, particularly in regard to individuals whose interests are not adequately represented by the state government engaged in coordination. Platform governance represents negotiations between lawmaking entities and private actors to advance some mutually beneficial end. Problems arise where this dynamic excludes consideration of the individual interests at stake—namely, those of the platform’s users—and leaves minority groups and marginalized communities particularly at risk.

B. Transparency Can Improve Facebook’s Accountability to Users by Creating Opportunities to Clarify and Correct.

Regardless of whether Facebook is justified in its cooperation with the government or in its interpretation of its Community Standards, the platform’s lack of transparency leads to confusion and the erosion of user trust. Facebook cannot both posture as a neutral arbiter and simultaneously agree to cooperate with one side of a disputing narrative without considering the broader, significant implications of its actions. Israelis and Palestinians agree that dangerous activity happens on Facebook and that affirmative steps are required to counter the proliferation of social media-incited violence. Facebook must take responsibility for delineating its actions from those of state actors to avoid its furtherance as a tool in geopolitical conflict. This requires Facebook to take dramatic steps in transparency, so that users are empowered to interrogate suspicions of bias and hold the platform accountable when it supports veiled prejudice or disparate treatment of vulnerable groups.

¹⁰² See Cohen, *supra* note 50, at 236.

¹⁰³ Michael D. Birnhack & Niva Elkin-Koren, *The Invisible Handshake: The Reemergence of the State in the Digital Environment*, 8 VA. J. L. & TECH. 6, ¶ 146 (2003).

Transparency is the first step toward forging pathways for accountability. Discussion around the platform's opacity in its content moderation practices indicates that "the biggest threat this private system of governance poses to democratic culture is the loss of a fair opportunity to participate, which is compounded by the system's lack of direct accountability to its users."¹⁰⁴ When Facebook uses its Community Standards as a means by which to enforce state preferences, and when it supplants procedural requirements to enforce the law, the platform's actions circumvent public accountability and skew public discourse in an anti-democratic fashion. Until recently, Facebook worked diligently to avoid formal regulation by pursuing informal, flexible agreements that facilitated regulatory capture and minimized the legal obligations that could have diminished its agency.¹⁰⁵ The resultant, ongoing lack of transparency may be used to serve the platform's business interests and state interests in accommodating ways at the expense of user accountability and traditional democratic restraints. Moreover, both individual users and state governments now struggle to locate tools by which intermediaries can be held accountable for their influence on the exercise of fundamental rights or denial of personal liberties.

Traditional mechanisms to advance corporate accountability do not stretch well when applied to platform governance and transnational rights. Some individuals have tried to use existing legal avenues available in the United States to seek recourse for harms that occurred in Israeli and Palestinian territories but had tenuous connections to conduct on Facebook.¹⁰⁶ These legal claims have not succeeded in court or created the friction necessary to incentivize different conduct on the part of Facebook. In the U.S., these tools conventionally include federal laws like the Alien Tort Claims Act (ATCA) and antiterrorism statutes, along with commonly accepted standards for Corporate Social Responsibility (CSR).¹⁰⁷ While ATCA and antiterrorism statutes provide a hook for corporate liability in some circumstances, both failed to gain traction with courts in the context of social media companies

¹⁰⁴ Klonick, *supra* note 96, at 1603.

¹⁰⁵ See e.g., Laura Kayali, *Facebook Embraces Regulation—Reluctantly*, POLITICO (Jan. 29, 2019), <https://www.politico.com/story/2019/01/29/facebook-reluctant-regulation-1130090> [<https://perma.cc/J672-8E7N>].

¹⁰⁶ See *infra* note 108.

¹⁰⁷ Practical Law Corporate & Securities, *Corporate Governance Standards: Overview*, THOMSON REUTERS PRACTICAL LAW (accessed May 12, 2019); *Expert Q&A on Trends in Corporate Social Responsibility*, THOMSON REUTERS PRACTICAL LAW (Sept. 19, 2013).

providing governments with the tools that facilitated the alleged harms.¹⁰⁸ Importantly, courts have expressly held that Facebook’s enforcement of its Community Standards “and implementation of networking algorithms for [the] benefit of all users” are protected from liability under the Communications Decency Act.¹⁰⁹ Finally, CSR norms provide instructive frameworks and guidelines for businesses operating in multinational settings. However, they neither create binding legal obligations nor necessarily speak to the business model of information platforms, which implicate a unique combination of business, government, and individual interests.

Consequently, accountability in the current landscape requires Facebook to take radical and proactive steps towards accountability and to move toward recognizing affirmative obligations and duties to its users. For content moderation practices, this calls for auditing its algorithms, collecting data on its internal practices, and providing content moderators with the tools, training, and support needed to perform their jobs. Facebook should not only collect information, but also make it publicly available. Presently, Facebook’s transparency reports cover only a fraction of its content moderation activity.¹¹⁰ Meaningful accountability to users requires empowering them with the information necessary to discern how their rights are accounted for when using the platform. Facebook has taken some initial, promising steps by publishing civil rights audit reports and releasing more transparency data than previously available.¹¹¹ Still, clearer protocols and accountability tools are needed to address concerns related to user experiences, particularly in conflicted

¹⁰⁸ See generally *Force v. Facebook*, 304 F. Supp. 3d 315 (E.D.N.Y. 2018) (dismissing a case that alleged Facebook materially supported terrorist activity by allowing Hamas and its supporters to use the social media platform to further their aims and finding that Facebook’s “maintenance of [an] internet platform” and algorithms applied to all users and were protected by the Communications Decency Act (“CDA”)); *Cohen v. Facebook*, 252 F. Supp. 3d 140 (E.D.N.Y. 2017) (finding that fear of future terrorist attacks based on online content was too speculative to satisfy standing requirements and that Facebook content moderation practices were covered by CDA immunity); *Corrie v. Caterpillar*, 403 F. Supp. 2d 1019 (W.D. Wash. 2005) (holding that company that sold the IDF bulldozers used to demolish Palestinian homes was not liable for aiding or abetting human rights violations).

¹⁰⁹ *Force*, 304 F. Supp. 3d at 331–332.

¹¹⁰ Facebook reports on content restrictions based on local law by country, with its report noting which removals were mandated by local law, as opposed to being mere violations of the platform’s community standards. Facebook does not identify content restrictions and removals by geographic region based on violation of the community standards. See *Content Restrictions*, FACEBOOK (2019), <https://transparency.facebook.com/content-restrictions> (accessed Oct. 27, 2019) [<https://perma.cc/S4UM-9QGV>].

¹¹¹ Facebook’s civil rights audit began in 2018. Up until that point, this information was not publicly available. See FACEBOOK, FACEBOOK’S CIVIL RIGHTS AUDIT–PROGRESS REPORT (June 30, 2019), https://fbnewsroomus.files.wordpress.com/2019/06/civilrightaudit_final.pdf [<https://perma.cc/C7CK-74QH>].

territories. To date, the platform has not provided concrete information as to how its content moderation practices change to account for competing narratives in geopolitically contested territories, nor how it might improve AI to account for these complex and nuanced issues.¹¹² Providing robust data to users would serve as a promising initial step to address Facebook's content moderation problems. While self-monitoring and self-governance regimes have significant limitations, they are necessary, albeit not sufficient, steps to confront this problem. Transparency without subsequent, thoughtful action cannot ameliorate the challenges of content moderation on global platforms.

C. A Human Rights Approach to Content Moderation Creates Space for Competing Narratives, Experiences, and Truths.

As a threshold matter, Facebook must recognize its responsibility as an influential force in international and geopolitical affairs, and then grapple with what those responsibilities demand in terms of conduct.¹¹³ Its current posture of neutrality undermines Facebook's ability to grapple with these questions. While Facebook vehemently defends its approach to content moderation as a neutral and objective process, this construction is a poor fit to tackle the thorny issues inherent to content moderation decisions.¹¹⁴ Here, it seems that Facebook confuses neutrality with its underlying normative values.¹¹⁵ Despite the evidence of bias that has surfaced in recent years, Facebook representatives still deny that Israeli and Palestinian users experience differential treatment on the platform because Facebook does not

¹¹² *Understanding Social Media and Conflict*, FACEBOOK NEWSROOM (June 20, 2019), <https://newsroom.fb.com/news/2019/06/social-media-and-conflict/> [<https://perma.cc/HE8R-JTB4>].

¹¹³ See Frenkel et al., *supra* note 99. In the wake of discovering Facebook's inadvertent facilitation of election meddling, humanitarian crises, and the proliferation of political propaganda, the platform remained reluctant to publicly engage on issues of responsibility and deflected criticism: "When researchers and activists in Myanmar, India, Germany and elsewhere warned that Facebook had become an instrument of government propaganda and ethnic cleansing, the company largely ignored them." *Id.*

¹¹⁴ See generally Anupam Chander & Vivek Krishnamurthy, *The Myth of Platform Neutrality*, 2 GEO. L. TECH. REV. 400 (2018).

¹¹⁵ In a recent speech at Georgetown University, Mark Zuckerberg outlined the values he believes Facebook stands for, stating, "I want to ensure the values of voice and free expression are enshrined deeply into how this company is governed." These normative values, in the abstract, do not demand a posture of neutrality. However, the instrumentalization of these normative values has translated into a posture of neutrality. Zuckerberg, *supra* note 52.

take a position on the issue.¹¹⁶ This statement reveals the problematic logic at work, as staking a position and suffusing bias are not the same thing. As Professor Kristen Eichensehr explains, the posture of neutrality “carries with it a risk of undue passivity, tending toward complicity.”¹¹⁷ Here, Facebook seems to use neutrality to shield itself from criticism regarding non-neutral outcomes. Facebook uses its posture of neutrality to deny its active influence on the nature of discourse. In so doing, it fails to acknowledge that “[t]he changes wrought by new technology do not benefit everyone equally.”¹¹⁸ As such, neutrality is not a productive position from which to unpack involvement in unintended outcomes; rather, this posture is self-defeating.

Additionally, Facebook’s Community Standards already stake non-neutral positions on matters like hate speech, terrorism, and incitement.¹¹⁹ At a fundamental level, limitations or rejection of certain kinds of expression reflect a balancing of values like civility, respect, and pluralism with individual rights to expression, participation, and access. While there are many legitimate positions that balance competing interests, grappling with these issues requires baseline recognition of the normative values underlying certain preferences.¹²⁰ Thus, Facebook should abandon its posture of neutrality and reorient its content moderation practices toward the protection of human rights.

Facebook should not be expected to resolve the Israeli-Palestinian conflict, but the company should be expected to moderate content in a manner consistent with the depth of trust and reliance it encourages users to place on the platform. This requires Facebook to grapple with competing truths and communities’ concerns that their speech, autonomy, and inherent dignity are

¹¹⁶ Ylenia Gostoli, *Palestinians Fight Facebook, YouTube Censorship*, AL JAZEERA (Jan. 20, 2018), <https://www.aljazeera.com/news/2018/01/palestinians-fight-facebook-youtube-censorship-180119095053943.html> (“[Facebook] engage[s] all over the world with governments, NGOs, academics. It doesn't mean we take a position.”) [<https://perma.cc/US73-QELC>].

¹¹⁷ Eichensehr, *supra* note 101, at 36.

¹¹⁸ Chander & Krishnamurthy, *supra* note 113, at 403; *see also* Zeynep Tufekci, *The Real Bias Built in at Facebook*, N.Y. TIMES (May 19, 2016), https://www.nytimes.com/2016/05/19/opinion/the-real-bias-built-in-at-facebook.html?rref=collection%2Fcolumn%2Fzeynep-tufekci&action=click&contentCollection=opinion®ion=stream&module=stream_unit&version=latest&contentPlacement=13&pgtype=collection (“Software giants would like us to believe their algorithms are objective and neutral, so they can avoid responsibility for their enormous power as gatekeepers while maintaining as large an audience as possible.”) [<https://perma.cc/RH7W-D9LM>].

¹¹⁹ Chander & Krishnamurthy, *supra* note 113, at 405–07.

¹²⁰ *Id.* at 405 (“[P]latforms are explicitly non-neutral with respect to certain issues specified in their community guidelines. These guidelines do not simply recapitulate the law, but rather set out a series of normative commitments.”).

vulnerable. While Facebook is currently in the process of developing an appeals process through its new Oversight Board,¹²¹ the bandwidth, priorities, agency, composition, and oversight of the Board will influence the degree to which it is able to address context-specific human rights concerns.¹²² In particular, Facebook should seek input and participation from human rights stakeholders to ensure that it facilitates fair review for individuals who may be marginalized or disproportionately discriminated against in the existing content moderation model.

As a matter related to but distinct from censorship concerns, Facebook must also grapple with respecting competing truths. Israelis and Palestinians have vastly different narratives about the geopolitical conflict and its motivations, history, and lived experiences. Instead of deferring to individual moderators or private self-governance mechanisms to decide the validity of opposing narratives, Facebook can pull from existing human rights guidelines and resources to inform content moderation decisions. The Global Network Initiative provides an instructive framework for the operationalization of platform governance in a manner that protects individual freedoms of expression and privacy.¹²³ Professors Deirdre Mulligan and Daniel Griffin explore the relationship between the respect for truth and Google search results, finding that a human rights framework can help platforms “avoid the slippery slope of arbitrating truthfulness” while still engaging in content moderation.¹²⁴ They draw from existing human rights reports, finding that, for content intermediaries, respect for the truth involves:

due diligence to identify, prevent, evaluate, mitigate and account for risks to the freedom of expression and privacy rights that are implicated by the company's products, services, activities and operations to assess actual and potential human rights impacts on individuals, integrating and acting upon the

¹²¹ Kate Cox, *Facebook Plans Launch of its Own “Supreme Court” for Handling Takedown Appeals*, ARS TECHNICA (Sept. 18, 2019), <https://arstechnica.com/tech-policy/2019/09/facebook-plans-launch-of-its-own-supreme-court-for-handling-takedown-appeals/> [<https://perma.cc/35ZX-25TQ>].

¹²² Concerns about the establishment of a Facebook Supreme Court are significant, but beyond the scope of this analysis.

¹²³ See *The GNI Principles*, GLOBAL NETWORK INITIATIVE (2019), <https://globalnetworkinitiative.org/gni-principles/> (“The duty of governments to respect, protect, promote and fulfill human rights is the foundation of this human rights framework. That duty includes ensuring that national laws, regulations and policies are consistent with international human rights laws and standards on freedom of expression and privacy.”) [<https://perma.cc/V9G7-8D34>].

¹²⁴ Deirdre K. Mulligan & Daniel S. Griffin, *Rescripting Search to Respect the Right to Truth*, 2 GEO. L. TECH. REV. 557, 574 (2018).

findings, tracking responses, and communicating how impacts are addressed . . . and where due diligence identifies circumstances when freedom of expression and privacy may be jeopardized or advanced[,] . . . employ human rights impact assessments and develop effective risk mitigation strategies as appropriate. . . .¹²⁵

While these steps may leave stakeholders on both sides unhappy with particular outcomes, by adopting an approach grounded in human rights, rather than in neutrality, Facebook can better account for each user's innate dignity and rights while navigating the challenges inherent to managing a platform in this setting.

V. CONCLUSION

As a private enterprise providing a global platform for public expression, Facebook plays a significant but convoluted role in modern political life. Over the past fifteen years, the platform has changed the way political power can be organized and exercised across borders. In its current construction, Facebook's enforcement of Community Standards and business practices function as a vector through which political power can be exerted, consolidated, and restricted in nontransparent ways. Facebook's entanglement in the Israeli-Palestinian conflict demonstrates the significant implications of platform governance along with its potential to reconfigure modes of democratic accountability. In navigating this landscape in a manner consistent with its expressed values, Facebook must grapple with the history and context of the geographic regions in which it operates. By adopting transparency, accountability, and human rights principles in its approach to content moderation, Facebook can disentangle its actions from those of state actors and rebuild trust and credibility with its users.

¹²⁵ *Id.* at 575–576 (internal quotations omitted).

RED LION BROADCASTING CO. V. FCC AND THE RISE OF SPEECH-ENHANCING REGULATIONS OF SOCIAL MEDIA PLATFORMS

Connor J. Suozzo*

CITE AS: 4 GEO. L. TECH. REV. 215 (2019)

TABLE OF CONTENTS

I. INTRODUCTION	215
II. BACKGROUND.....	218
A. Broadcasting and Print Media: <i>Red Lion</i> , <i>Tornillo</i> , and <i>Pacifica</i> ...	218
B. The Internet: <i>Reno</i> and <i>Packingham</i>	222
C. <i>Red Lion</i> Today and the Rise of Speech-Enhancing Regulations...	223
III. RED LION FOR SOCIAL MEDIA PLATFORMS	225
A. <i>Red Lion</i> Factors	225
1. <i>Historically Regulated Status</i>	226
2. <i>Scarcity Rationale</i>	229
3. <i>Reach and Interference</i>	232
B. <i>Pacifica</i> factors	233
IV. REGULATORY SOLUTIONS.....	234
A. Election transparency laws	235
B. Anti-bias bills.....	237
V. CONCLUSION.....	237

I. INTRODUCTION

Americans increasingly get their news through the Internet, and specifically, through social media platforms.¹ In fact, Americans get more of

* Georgetown Law, J.D. expected 2020. I am deeply grateful to Professor Erin Carroll for her help and guidance in developing this article. I would also like to thank the editors of the Georgetown Law Technology Review for their hard work and useful feedback.

¹ 47 U.S.C. § 230(a)(5) (2018) (“Increasingly Americans are relying on interactive media for a variety of political, educational, cultural, and entertainment services.”); Katerina Eva Matsa & Elisa Shearer, *News Use Across Social Media Platforms 2018*, PEW RES. CTR. (Sep. 10, 2018), <https://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/> [https://perma.cc/34JC-SUKM]; Amy Mitchell et al., *The Modern News Consumer*,

their news from social media than from print newspapers.² As the rate of this news consumption increases, two attributes of social media platforms become more readily apparent. First, these platforms play a filtering role, stemming from their status as *de facto* gatekeepers of an endless flow of information that includes news. Second, social media companies' incentive structure as technology platforms, rather than as media companies, does not align with the values of the traditional press.³ These two attributes combine to threaten the marketplace of ideas by facilitating the spread of disinformation⁴ or other harmful yet engagement-driving content, or by suppressing certain viewpoints. This ultimately leads to an increase in apathy and suspicion at best, or polarization and radicalization at worst.

Social media platforms⁵ provide users with a "modern public square" subject to First Amendment speech and press protections.⁶ Thus, any regulation of such platforms must satisfy constitutional restraints. Over the years, the Supreme Court has taken varying approaches to its review of regulations of speech, depending on the speech's medium. The Court views regulation of newspapers as speech-abridging,⁷ but comparable regulation of radio and television stations as speech-enhancing.⁸ The distinction between speech-enhancing regulations and speech-abridging ones is constitutionally significant due to the First Amendment's prohibition against any law "abridging the freedom of speech."⁹ Lawmakers can avoid constitutional obstacles even where their laws restrict some form of speech, as long as they

PEW RES. CTR. (Jul. 7, 2016), <https://www.journalism.org/2016/07/07/pathways-to-news/> [<https://perma.cc/L3S9-QD44>].

² Elisa Shearer, *Social Media Outpaces Print Newspapers in the U.S. as a News Source*, PEW RES. CTR. (Dec. 10, 2018), <https://www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-a-news-source/> [<https://perma.cc/E9RN-SE8L>].

³ See generally Kyle Langvardt, *Regulating Habit-Forming Technology*, 88 FORDHAM L. REV. 129 (2019).

⁴ See Nabihah Syed, *Real Talk About Fake News: Towards a Better Theory for Platform Governance*, 127 YALE L.J.F. 337, 339–341 (2017).

⁵ Rapidly evolving technology and the overlap of functionalities among social media and older services like email make it difficult to articulate a perfect definition of social media. Using the Merriam-Webster definition of social media as a guide, this note uses "social media" in reference to services like Facebook, Twitter, Instagram, and other websites where users interact and share expressive content with each other. See *Social Media*, MERRIAM-WEBSTER ONLINE DICTIONARY, <https://www.merriam-webster.com/dictionary/social%20media> (defining social media as "forms of electronic communication (such as websites for social networking and microblogging) through which users create online communities to share information, ideas, personal messages, and other content (such as videos)") [<https://perma.cc/KNX8-JBLB>].

⁶ *Packingham v. North Carolina*, 137 S. Ct. 1730, 1737 (2017).

⁷ See *Miami Herald Pub. Co. v. Tornillo*, 418 U.S. 241, 256–58 (1974).

⁸ See *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 375–76 (1969).

⁹ U.S. CONST. amend. I.

can convince courts that the restrictions ultimately have speech-enhancing effects.¹⁰ Noise ordinances are a common example of speech-enhancing regulations. Although they undoubtedly limit forms of expression, they nevertheless enhance speech by fostering an environment where other ideas will not be drowned out. Thus, litigators may choose to defend the constitutionality of speech regulations by painting them as speech-enhancing.

The Court has found that speech is particularly vulnerable to suppression by non-governmental forces in the medium of broadcasting, so it treats regulations of that medium leniently,¹¹ while it strictly scrutinizes regulations of speech in other media like printed news.¹² In the early years of the Internet, the Supreme Court adopted the strict scrutiny approach for regulations that affect the Internet as a whole,¹³ but the Court has not answered the question of how it will review regulations of social media platforms, where there is a unique threat to the marketplace of ideas.

This Note explores the question of whether, as a matter of constitutional law, regulations of social media platforms should be viewed as speech-enhancing or speech-abridging under *Red Lion Broadcasting Co. v. FCC*. The public's growing concern over social media, amplified by recent events like Russia's interference in the 2016 U.S. presidential election, have prompted legislators to propose or enact laws—specifically, electioneering and anti-bias laws—aimed at enhancing speech.¹⁴ These efforts indicate a general awareness that the marketplace of ideas is uniquely threatened on social media and that regulations affecting speech on this medium will tend to enhance it. In this note, Part II summarizes the competing approaches to First Amendment review of media regulations embodied by *Red Lion* and *Miami Herald Publishing Co. v. Tornillo* and then traces their progeny to the Internet era. Part II also highlights recently enacted and proposed legislation that targets social media and would likely implicate this body of case law. Part III analyzes the rationales underlying the *Red Lion* decision and argues that they also apply in the social media context. Specifically, the relevant characteristics of the broadcast medium, including its historically regulated status, limited airwave spectrum, long reach, intrusiveness, and unique accessibility to children, are comparable to the characteristics of social media. Part III argues that a comparison between the two media forms reveals that unregulated social media poses a greater threat to the marketplace of ideas than broadcasting did in 1971, the year *Red Lion* was decided. Finally, Part IV discusses the legal viability of specific proposed regulations under this framework, including

¹⁰ See *Red Lion*, 395 U.S. at 375.

¹¹ See *FCC v. Pacifica Found.*, 438 U.S. 726, 748 (1978).

¹² See *Miami Herald Pub. Co. v. Tornillo*, 418 U.S. 241, 256–58 (1974).

¹³ See *Reno v. Am. Civil Liberties Union*, 521 U.S. 844, 868–69 (1997).

¹⁴ See discussion *infra* Part I.C.

electioneering laws like the Honest Ads Act and anti-bias laws such as the Biased Algorithm Deterrence Act.

II. BACKGROUND

A. Broadcasting and Print Media: *Red Lion*, *Tornillo*, and *Pacifica*

The concept of the marketplace remains a useful and compelling analogy for the exchange of ideas online. Courts have long held that the First Amendment's protections of speech and press primarily function to foster a marketplace of ideas,¹⁵ though some commentators reject this theory as an inadequate description of information flow on social media.¹⁶ The Supreme Court upheld a broadcasting regulation as speech-enhancing in *Red Lion Broadcasting Co. v. FCC* because the marketplace of ideas was particularly endangered in that medium.¹⁷ Then, in *Tornillo*, the Court struck down a strikingly similar regulation that applied to newspapers, while firmly endorsing the same marketplace theory.¹⁸ Finally, the Court reaffirmed the *Red Lion* approach by upholding a restriction on broadcast speech in *FCC v. Pacifica Foundation*.¹⁹

The First Amendment's speech and press protections were included in the Constitution to end oppressive forms of censorship.²⁰ The Supreme Court later endorsed a theory of the amendment's protection as a means of fostering a marketplace of ideas.²¹ Under this theory, government suppression of bad ideas is ineffective; the best way to eliminate falsehood is to allow those bad ideas to compete with good ones, which will ultimately win out.²²

Media and technology lawyer Nabiha Syed argues that the marketplace theory fails to accurately describe online speech, as it "leaves both users and social media platforms ill-equipped to deal with rapidly evolving problems like fake news."²³ She recommends that a new, more realistic theory take its

¹⁵ See, e.g., *Citizens United v. Fed. Election Comm'n*, 558 U.S. 310, 354 (2010); *Turner Broad. Sys., Inc. v. FCC*, 512 U.S. 622, 641 (1994); *New York Times Co. v. Sullivan*, 376 U.S. 254, 269–272 (1964).

¹⁶ See Syed, *supra* note 4.

¹⁷ See *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 390 (1969) ("It is the purpose of the First Amendment to preserve an uninhibited marketplace of ideas in which truth will ultimately prevail, rather than to countenance monopolization of that market.").

¹⁸ See *Miami Herald Pub. Co. v. Tornillo*, 418 U.S. 241, 252 (1974).

¹⁹ See *FCC v. Pacifica Found.*, 438 U.S. 726, 748 (1978).

²⁰ See generally David A. Anderson, *The Origins of the Press Clause*, 30 UCLA L. REV. 455 (1983).

²¹ See *Abrams v. United States*, 250 U.S. 616, 630 (1919).

²² See *Red Lion*, 395 U.S. at 390; *Abrams*, 250 U.S. at 630.

²³ Syed, *supra* note 4.

place.²⁴ Syed is right to note that the systemic features of social media prevent the marketplace of ideas from functioning as it should,²⁵ but we need not be so quick to abandon what has been a central feature of First Amendment law for the past century.²⁶ Instead, these failings warrant a more lenient judicial attitude toward laws seeking to correct them.

Fifty years ago, the Supreme Court reviewed a Federal Communications Commission (FCC) regulation of the broadcasting medium. In that case, *Red Lion Broadcasting Co. v. FCC*, the Court upheld the “fairness doctrine,” a longstanding regulation requiring that a radio station allot equal time to all qualified candidates for a given public office and that targets of personal or group attacks must be offered a reasonable opportunity to respond through the radio station’s facilities.²⁷ After a Pennsylvania radio station carried a broadcast criticizing journalist and writer Fred J. Cook, Cook demanded free reply time, which the station refused. The FCC and the D.C. Circuit Court of Appeals held that the station was obligated to provide Cook reply time for free.²⁸

The Supreme Court unanimously affirmed these decisions,²⁹ holding that the fairness doctrine was not unconstitutional, but instead “enhance[d], rather than abridge[d], the freedoms of speech and press protected by the First Amendment.”³⁰ The Court reasoned that “differences in the characteristics of new media justify differences in the First Amendment standards applied to them.”³¹ The Court held that the government may permissibly regulate broadcasting equipment because of its potential to “snuff out the speech of others,” stemming from three characteristics of the “new media” of broadcasting.³² First, radio licenses are resources which “the Government has denied others the right to use.”³³ Second, there is a “scarcity of radio frequencies.”³⁴ Third, “the reach of radio signals is incomparably greater than

²⁴ See Syed, *supra* note 4.

²⁵ See Syed, *supra* note 4 (arguing that the necessity of online content filtering has a natural tendency to “feed insular ‘echo chambers,’” and that platforms’ profit incentives encourage the rapid dissemination of fake news because it is sensational and cheaply distributed).

²⁶ See *Abrams*, 250 U.S. at 630 (Holmes, J. Dissenting) (“the best test of truth is the power of the thought to get itself accepted in the competition of the market.”). *Abrams*, announced one hundred years ago in November 1919, was the first time the Supreme Court—albeit in Justice Holmes dissent—recognized the marketplace of ideas theory of free speech.

²⁷ See *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 373–75 (1969).

²⁸ *Id.* at 371–73.

²⁹ See *id.* at 401 (indicating that eight justices voted in favor of the FCC, while Justice Douglas did not participate in the opinion).

³⁰ *Id.* at 375.

³¹ *Id.* at 386.

³² *Id.* at 386–87.

³³ *Id.* at 391.

³⁴ *Id.* at 390.

the range of the human voice, and the problem of interference is a massive reality.”³⁵ Together, these characteristics justified government regulations intended to “preserve the marketplace of ideas in which truth will ultimately prevail, rather than to countenance monopolization of that market.”³⁶ The Court added that such regulations were consistent with First Amendment values because they prevented a small number of station owners from having “unfettered power” to limit use of their stations.³⁷ As the Court stated, “[t]here is no sanctuary in the First Amendment for unlimited private censorship operating in a medium open to all.”³⁸

Five years later, the Court rejected a similar regulation that applied to newspapers in *Miami Herald Publishing Company v. Tornillo*.³⁹ In that case, a Florida statute granted political candidates a right to equal space to reply to newspaper criticism.⁴⁰ The *Miami Herald* newspaper published editorials critical of a candidate running for a position in the Florida state legislature.⁴¹ The candidate, citing the Florida statute, demanded the newspaper print his replies, but the *Herald* refused.⁴² The Florida Supreme Court held that “free speech was enhanced and not abridged by the Florida right-of-reply statute.”⁴³ The U.S. Supreme Court reversed, rejecting the appellee’s arguments that the regulations protected the marketplace of ideas in a world where newspapers had become “noncompetitive and enormously powerful and influential in [their] capacity to manipulate popular opinion.”⁴⁴ Although both parties cited heavily to *Red Lion* in their briefs, the Supreme Court did not reference it at all when it unanimously held that the right-of-reply statute violated the First Amendment.⁴⁵ The Court opined, “[a] responsible press is an undoubtedly desirable goal, but press responsibility is not mandated by the Constitution, and, like many other virtues, it cannot be legislated.”⁴⁶

³⁵ *Id.* at 387–88.

³⁶ *Id.* at 390.

³⁷ *Id.* at 392.

³⁸ *Id.*

³⁹ 418 U.S. 241 (1974).

⁴⁰ *Id.* at 243.

⁴¹ *Id.*

⁴² *Id.* at 244.

⁴³ *Id.* at 245 (relying on *Red Lion*); *Tornillo v. Miami Herald Pub. Co.*, 287 So. 2d 78, 83–85 (Fla. 1973), rev’d, 418 U.S. 241 (1974).

⁴⁴ *Tornillo*, 418 U.S. at 249.

⁴⁵ See 418 U.S. 241; *Miami Herald Pub. Co. v. Tornillo*, 1974 WL 185859 (U.S.), 20–23 (U.S., 2004) (appellant’s brief); *Miami Herald Pub. Co. v. Tornillo*, 1974 WL 185860 (U.S.) (U.S., 2004) (appellee’s brief).

⁴⁶ *Tornillo*, 418 U.S. at 256.

Ironically, the *Tornillo* opinion partly relied on a kind of scarcity—though not the same scarcity that contributed to the result in *Red Lion*.⁴⁷ In the same paragraph that the Court came closest to referencing *Red Lion*—when it conceded that a newspaper is not “subject to the finite technological limitations of time that confront a *broadcaster*”⁴⁸ and is in that sense not as burdened by fairness requirements—the Court held that scarcity of print space made the right-of-reply statute too great of a burden on newspaper editors.⁴⁹ Ultimately, the Court found that such regulation of newspapers harmed speech by raising newspapers costs, discouraging the publishing of controversial articles, and “intru[ding] into the function of editors.”⁵⁰

Tornillo did not mark a shift in the Court’s philosophy of the First Amendment but rather solidified one approach reflecting suspicion of regulations in the context of print media. The Court’s alternative, *Red Lion* approach for broadcasting media remained intact. In 1978, the Supreme Court decided *FCC v. Pacifica Foundation*, holding that it was permissible under the First Amendment for the FCC to prohibit “indecent” language in broadcasting.⁵¹ A radio station broadcast a satiric monologue by George Carlin titled “Filthy Words,” which listed a number of “words you couldn’t say on the public . . . airwaves.”⁵² The FCC suit was prompted by a father’s complaint after he heard the broadcast while driving with his young child.⁵³ The Court reviewed the FCC’s determination that the monologue was indecent and therefore prohibited by statute, and it upheld the Commission’s decision, writing, “it is broadcasting that has received the most limited First Amendment protection.”⁵⁴ Among the “complex” reasons for distinguishing print media from broadcasting, the Court stated that the relevant ones here were the latter medium’s “uniquely pervasive presence in the lives of all Americans” and its “unique[] accessib[ility] to children.”⁵⁵

This history illustrates the context-specific nature of First Amendment protection. To preserve a vibrant marketplace of ideas, the Court adjusts its inquiry to address the unique dangers that might befall that marketplace in a

⁴⁷ *Id.* at 257 (“[I]t is not correct to say that, as an economic reality, newspaper can proceed to infinite expansion of its column space to accommodate the replies that a government agency determines or a statute commands the readers should have available.”).

⁴⁸ *Id.* at 256–57 (emphasis added).

⁴⁹ *See id.*

⁵⁰ Justice White addressed this speech-chilling argument in *Red Lion* as well, but he found that such an effect had not occurred and was not likely to occur in broadcasting. *See Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 393–94 (1969).

⁵¹ *See*, 438 U.S. 726, 738 (1978).

⁵² *Id.* at 729.

⁵³ *Id.* at 730.

⁵⁴ *Id.* at 748.

⁵⁵ *Id.* at 748–49.

given medium. Accordingly, the Court had to engage with these competing approaches when the Internet started to take off. The two cases discussed in the next section illustrate how the Court grappled with the rapidly evolving medium.

B. The Internet: *Reno* and *Packingham*

The FCC repealed the fairness doctrine in 1987.⁵⁶ Though *Red Lion* remains good law, courts have limited its scope to the broadcasting context. In the 1997 case *Reno v. American Civil Liberties Union*, the Court declined to extend the regulation-friendly approach of *Red Lion* to the Internet as a whole.⁵⁷ Although the broadcasting context made regulation of “indecent speech” permissible in *Pacifica* nineteen years earlier, the Court in *Reno* took a different approach to a statutory prohibition of “indecent communications” on the Internet.⁵⁸ It held that the factors justifying regulation of the broadcast media—the “history of extensive government regulation, the scarcity of available frequencies at its inception, and its ‘invasive’ nature,” are not present on the Internet.⁵⁹ In dismissing the application of *Red Lion*’s scarcity rationale, the Court wrote:

[U]nlike the conditions that prevailed when Congress first authorized regulation of the broadcast spectrum, the Internet can hardly be considered a “scarce” expressive commodity. It provides relatively unlimited, low-cost capacity for communication of all kinds Through the use of chat rooms, any person with a phone line can become a town crier with a voice that resonates farther than it could from any soapbox.⁶⁰

In 2017, the Court again considered regulation of Internet speech—specifically, speech by users on social media platforms.⁶¹ In that case, *Packingham v. North Carolina*, the Court struck down a North Carolina law making it a felony for a registered sex offender to post on social media websites like Facebook and Twitter.⁶² Writing for the majority, Justice Kennedy referred to social media as “the modern public square,” and wrote that today, it is clear that “cyberspace . . . and social media in particular” is the

⁵⁶ In re Complaint of Syracuse Peace Council Against Television Station WTVH Syracuse, N.Y., 2 FCC Rcd. 5043 (1987) (repealing the fairness doctrine).

⁵⁷ *Reno v. Am. Civil Liberties Union*, 521 U.S. 844, 868 (1997).

⁵⁸ *Id.* at 868–79.

⁵⁹ *Id.* at 868 (citations omitted).

⁶⁰ *Id.* at 870.

⁶¹ *Packingham v. North Carolina*, 137 S. Ct. 1730 (2017).

⁶² *Id.* at 1733, 1738.

“most important place[] . . . for the exchange of views.”⁶³ Kennedy urged the Court to “[e]xercise extreme caution before suggesting that the First Amendment provides scant protection for access to vast networks in that medium.”⁶⁴

Justice Alito concurred in the judgment but rejected Kennedy’s “undisciplined dicta” that “equate[s] the entirety of the Internet with public streets and parks.”⁶⁵ Alito added, “there are important differences between cyberspace and the physical world.”⁶⁶ He proceeded to list “a few that are relevant to Internet use by sex offenders,” including greater difficulty of parental monitoring in observing sex offenders as they approach children and ease of assuming a false identity due to the “unprecedented degree of anonymity” on the Internet.⁶⁷

Justice Alito also took issue with Justice Kennedy’s cautious approach to the review of Internet regulations.⁶⁸ Because the Internet poses a heightened danger to children, in Alito’s view, “caution” would warrant a more regulatory-friendly approach.⁶⁹

C. *Red Lion* Today and the Rise of Speech-Enhancing Regulations

After *Reno*, *Red Lion* has seldom been invoked in Internet cases. The few times litigators have raised the issue, they have not been successful.⁷⁰ In *Washington Post v. McManus*, for example, the U.S. District Court for the District of Maryland held that a state electioneering statute that required both social media websites and news websites to publish information about political advertisements and to submit related records for state inspection violated the First Amendment.⁷¹ Although the court recognized that “the [Supreme] Court has carved out special rules for broadcast media,” it relied on *Reno* to hold that the *Red Lion* rationale did not apply to the Internet.⁷²

Despite the Court’s refusal to extend *Red Lion* or adopt a similarly relaxed approach to Internet cases in general, the door remains open in the social media context. First, the Supreme Court has not decided how it will review speech-enhancing regulations of social media platforms, such as election disclosure laws or anti-bias laws. *Reno* did not foreclose the

⁶³ *Id.* at 1732, 1735.

⁶⁴ *Id.* at 1736.

⁶⁵ *Id.* at 1738 (Alito, J., concurring).

⁶⁶ *Id.* at 1743.

⁶⁷ *Id.* at 1743–44.

⁶⁸ *Id.* at 1744 (Alito, J., concurring).

⁶⁹ *See id.*

⁷⁰ *See, e.g., Washington Post v. McManus*, 355 F. Supp. 3d 272, 288 (D. Md. 2019).

⁷¹ *Id.*

⁷² *Id.*

possibility of a relaxed form of scrutiny in this context, as *Reno* considered a regulation affecting the Internet as a whole and not the narrower context of social media.⁷³ Moreover, both *Reno* and *Packingham* considered laws that abridged user speech, but they did not limit the speech of Internet “gatekeepers”—those who own and operate social media platforms. *Red Lion* was a case about the gatekeepers in broadcasting, and it was their status as gatekeepers that threatened the marketplace of ideas. The Supreme Court has not had the opportunity to address regulations that would enhance Internet user speech by placing limits on the content-moderating powers of Facebook, Twitter, and Google. It is therefore an open question whether *Red Lion* or a similar regulation-friendly approach applies to laws affecting the speech of gatekeepers of social media platforms.

And it is a question that will increasingly be asked. For most of the short history of the Internet, speech-affecting laws were designed to protect persons from harm, and they did not purport to enhance speech.⁷⁴ But recent events have raised the public’s consciousness as to dangers posed by social media platforms’ “unfettered power”⁷⁵ over what users see online, and those events have triggered a new dawn for speech-enhancing regulations. As Stanford Law School Professor Nathaniel Persily writes in reference to Twitter, Facebook, and Google, “[t]ry as they might not to be media companies, they nevertheless have power far in excess of that which legacy media institutions had in their heyday, let alone today.”⁷⁶ Legislatures have started to take an interest in these concerns: to combat perceived dangers to the online marketplace of ideas, the federal government and several states have proposed or passed laws that potentially limit platform speech in order to enhance speech overall, including election ad disclosure requirements and anti-bias statutes.⁷⁷ The constitutionality of these measures remains in question.

⁷³ *Reno v. Am. Civil Liberties Union*, 521 U.S. 844, 869 (1997).

⁷⁴ For example, the Communications Decency Act attempted to restrict indecent and obscene Internet content, and the recently enacted FOSTA-SESTA package targeted sex trafficking. *See* 47 U.S.C.A. § 223; Allow States and Victims to Fight Online Sex Trafficking Act of 2017, Pub. L. No. 115-164, 132 Stat. 1253 (2018).

⁷⁵ *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 392 (1969).

⁷⁶ Nathaniel Persily, *The 2016 U.S. Election: Can Democracy Survive the Internet?*, 28 J. DEMOCRACY 63, 74 (2017).

⁷⁷ Honest Ads Act, S. 1989, 115th Cong. (2017); *see also* Democracy Protection Act, 2018 N.Y. Laws 358–360 (codified at N.Y. ELEC. § 14-100, -106, -107, -126); 2018 Wash. Sess. Laws 2453–80; MD. CODE ANN., ELEC. LAW § 13-405 (West 2019).

III. RED LION FOR SOCIAL MEDIA PLATFORMS

Justice Kennedy cautioned the Court in *Packingham* not to declare any hard rules about how the Constitution applies to the Internet because the Internet is rapidly and constantly evolving.⁷⁸ Although the historical scar left by colonial-era Britain's prior restraints on *print* media—the regime that prompted the First Amendment's adoption—made the hands-off outcome in *Tornillo* almost inevitable,⁷⁹ online social media is new and different enough to justify consideration of a different approach. *Red Lion* demands that courts seriously consider how the factual contexts surrounding new media affect the marketplace of ideas and whether regulations addressing pressing concerns in those contexts will likely be speech-enhancing or speech-abridging.

For an attempt to extend *Red Lion* beyond the broadcasting context to succeed, litigators will have to show that *Red Lion*'s rationales—and the concerns underlying them—apply to the factual context of social media. Specifically, litigators will have to compare platforms with the three factors that distinguished broadcasting in *Red Lion*: (1) broadcasting's historically regulated status,⁸⁰ (2) spectrum scarcity,⁸¹ and (3) broadcasting's reach and potential for interference.⁸² Further, those litigators should consider the two other factors that the Court added in *Pacifica* to distinguish broadcasting: (1) its "invasiveness" and (2) its effect on vulnerable users like children.⁸³

A. *Red Lion* Factors

On a superficial level, social media platforms do not possess all of these characteristics on a one-to-one scale, but each of the *Red Lion* and *Pacifica* factors addressed underlying concerns that are implicated by social media platforms. Broadcasting's history as a regulated medium justified a relaxed First Amendment standard—not because licensure requirements had been in place long enough to get grandfathered into an exemption from First Amendment scrutiny, but because those requirements were necessary to cure a defective marketplace of ideas: Without an ordering of airwave usage, no

⁷⁸ *Packingham v. North Carolina*, 137 S. Ct. 1730, 1736 (2017).

⁷⁹ *See, e.g., New York Times Co. v. United States*, 403 U.S. 713, 719 (1971) (Black, J., concurring).

⁸⁰ *See Red Lion*, 395 U.S. at 388; *see also Reno v. Am. Civil Liberties Union*, 521 U.S. 844, 868 (1997).

⁸¹ *See Red Lion*, 395 U.S. at 387–88.

⁸² *See id.* at 388–89.

⁸³ *See FCC v. Pacifica Found.*, 438 U.S. 726, 763–64 (1978) (Brennan, J., concurring) (noting that "[b]oth the [majority] opinion . . . and the [concurrence] rely principally on two factors in reaching this conclusion: (1) the capacity of a radio broadcast to intrude into the unwilling listener's home, and (2) the presence of children in the listening audience").

one could intelligibly communicate by radio.⁸⁴ The second *Red Lion* factor, spectrum scarcity, necessitated a consolidation of power in the hands of relatively few actors and entrenched that power even when the spectrum became less scarce. Thus, regulations on broadcasting only served to enhance speech by transferring some of the broadcasters' power to those whose voices would otherwise not be heard. The third and final *Red Lion* factor, the far reach of airwaves and the potential to interfere with other speech over the radio, bears similarities to speech on social media. Posts online are potentially viewed and listened to by millions, and the gatekeepers' incentives cause them to moderate content in such a way that displaces certain speech.⁸⁵ The *Pacifica* factors are also implicated on social media: harmful online content can appear unexpectedly on one's feed, and social media is easily accessible to children.⁸⁶ That latter concern is compounded by the heightened vulnerability of children on the Internet, in contrast to physical locations where they are more easily supervised, as Justice Alito noted in his *Packingham* concurrence.⁸⁷ These "differences in the characteristics of new media" justify a different First Amendment standard.⁸⁸

1. *Historically Regulated Status*

The first rationale for permitting the fairness doctrine in *Red Lion* had to do with the historically regulated status of the broadcasting industry.⁸⁹ This was not a recognition of a historical exception, where longstanding acceptance of certain regulations opened the door to new ones of a similar nature. Instead, the historical regulatory regime in broadcasting created a status quo where the medium was controlled by a small number of powerful licensees.⁹⁰ Although social media platforms have by no means been regulated to the same extent as broadcasting, platforms' inherent characteristics give rise to an even more severe concentration of power.

The licensing regime in *Red Lion* was significant because it was necessary to increase speech by facilitating it in an orderly way through governmental grants. Without a licensing program, the Court stated, the airwaves would be chaotic because everyone would use them, and no one

⁸⁴ See *Red Lion*, 395 U.S. at 388.

⁸⁵ See discussion *infra* Part II.A.3.

⁸⁶ See discussion *infra* Part II.B.

⁸⁷ See *Packingham v. North Carolina*, 137 S. Ct. 1730, 1743 (2017) (Alito, J. concurring).

⁸⁸ *Red Lion*, 395 U.S. at 386.

⁸⁹ See *id.* at 399–400.

⁹⁰ See *id.*

would be heard.⁹¹ Without governmentally imposed order, there could be no marketplace of ideas on the air at all. Congress solved this problem by controlling who received licenses and who did not. Congress did not confer any speech rights, it merely conferred temporary⁹² exclusivity rights that made speech possible. It was therefore not a violation of anyone's freedom of speech for Congress or the FCC to transfer those exclusive rights to others. The FCC simply set up a trigger for when such a transfer would occur: whenever a broadcaster aired personal attacks.⁹³ Thus, when broadcasters exercised speech over the radio, the use of that amplification was a regulatory privilege not subject to the ordinary First Amendment safeguards.

It was not the fact of historical regulation that justified a relaxed standard in broadcasting, but the new status quo that regulation created.⁹⁴ After all, speech critical of the government was also historically regulated for a long period of time, yet the Sedition Act would never be tolerated today. The broadcasting licensure regime mattered in *Red Lion* because it created a distorted marketplace of ideas. Everyone can speak and be heard at the ordinary volume of the human voice, but the regulatory regime in broadcasting created a significant imbalance in who was allowed to speak. With the power of radio speech consolidated in the hands of so few, speech-enhancing mechanisms were necessary to reestablish the preconditions of a properly functioning marketplace of ideas.

The Internet has not come close to being regulated to the same extent as radio broadcasting, as the Court recognized in *Reno*. But present complaints about platforms' behavior echoes the complaints made by the FCC in *Red Lion*. Social media companies have come under fire for banning certain content and users.⁹⁵ Much of these companies' use of this control is laudable; for instance, Facebook recently announced that its algorithms will begin banning white supremacist messages.⁹⁶ Other platforms prohibit "hate speech" generally.⁹⁷ Still, these companies' power to control the content seen and heard

⁹¹ See *id.* at 388 ("It was . . . the chaos which ensued from permitting anyone to use any frequency at whatever power level he wished, which made necessary the enactment of the Radio Act of 1927 and the Communications Act of 1934.").

⁹² Broadcasting licenses expired in three years unless they were renewed. *Id.* at 394.

⁹³ *Id.* at 391.

⁹⁴ *Id.* at 388–89.

⁹⁵ Marissa Lang, *Blocked and Banned by Social Media: When is it Censorship?* S.F. CHRON. (Aug. 30, 2016), <https://www.sfchronicle.com/business/article/Blocked-and-banned-by-social-media-When-is-it-9193998.php?psid=4yAxS> [<https://perma.cc/9D2Y-LVSE>].

⁹⁶ Liam Stack, *Facebook Announces New Policy to Ban White Nationalist Content*, N.Y. TIMES (Mar. 27, 2019), <https://www.nytimes.com/2019/03/27/business/facebook-white-nationalist-supremacist.html/> [<https://perma.cc/YC5W-GYFX>].

⁹⁷ See, e.g., *Community Guidelines*, TUMBLR, <https://www.tumblr.com/abuse/hatespeech> (accessed Nov. 27, 2019) [<https://perma.cc/6FLD-YL2S>]; *Hateful Conduct Policy*, TWITTER,

by users, an increasing number of which use these websites as their primary source of news, is concerning. In *Red Lion*, an extensive regulatory regime was necessary to facilitate speech that otherwise would not exist, but it also created a problem of exclusive access by a small number of gatekeepers.⁹⁸ On social media, an extensive regulatory regime was not necessary to confer the exclusive power over content enjoyed by the medium's gatekeepers, which is even greater than it was in the broadcasting industry fifty years ago. Both the broadcasters then and the social media platforms now have the ability to purge certain perspectives from online communities and to amplify harmful and false information. And because their incentives and values diverge from those of the traditional news industry,⁹⁹ this danger has become a reality online. Professor Langvardt writes that the incentive for social media platforms to promote "engagement" has produced algorithms that "harm[] the quality of public discourse and deliberation."¹⁰⁰ In extreme cases, this has led to user radicalization.¹⁰¹ For example, one report found that YouTube "provides a breeding ground for far-right radicalization, where people interested in conservative and libertarian ideas are quickly exposed to white nationalist ones."¹⁰² Platform algorithms also tend to exclude perspectives that users disagree with, creating polarization.¹⁰³ Ultimately, this produces echo chambers where extremist groups can gain influence while having their ideas go unchallenged.¹⁰⁴ Therefore, the nature of online platforms likely poses an even greater danger to the marketplace of ideas than did the broadcasting industry's regulated status in *Red Lion*.

<https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (accessed Nov. 27, 2019) [<https://perma.cc/LX4P-YHZA>].

⁹⁸ *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 390 (1969).

⁹⁹ See Erin C. Carroll, *Platforms and the Fall of the Fourth Estate: Looking Beyond the First Amendment to Protect Watchdog Journalism*, MD. L. REV. (forthcoming 2020) (arguing that platform values and norms disincentivize watchdog reporting and compromise the press's ability to perform a core structural role).

¹⁰⁰ Kyle Langvardt, *Regulating Habit-Forming Technology*, 88 FORDHAM L. REV. 129 (2019).

¹⁰¹ *Id.* at 149 ("Many recommendation algorithms . . . have been shown to repeatedly send users along a 'radicalizing path.'").

¹⁰² Olivia Solon, *YouTube's "Alternative Influence Network" Breeds Right Wing Radicalization, Report Finds*, GUARDIAN (Sep. 18, 2018), <https://www.theguardian.com/media/2018/sep/18/report-youtubes-alternative-influence-network-breeds-rightwing-radicalisation> [<https://perma.cc/89YR-VZPV>].

¹⁰³ See, e.g., Michela Del Vicario et al., *Echo Chambers: Emotional Contagion and Group Polarization on Facebook*, SCI. REPS. (Dec. 1, 2016).

¹⁰⁴ See *id.*

2. *Scarcity Rationale*

The second rationale justifying the separation of broadcasting from other contexts in *Red Lion* is the “scarcity of radio frequencies.”¹⁰⁵ Because the physical realities of the electromagnetic spectrum prevent everyone from using the airwaves at once, if there is to be any orderly and productive use of the airwaves at all, it would necessarily be at the exclusion of the vast majority of would-be speakers.¹⁰⁶ When only so few actors have access to the means of expression, there exists a danger of some ideas being imposed on listeners while other ideas lack representation on the airwaves. As the Court emphasized in *Red Lion*, the First Amendment “right of the viewers and listeners . . . is paramount.”¹⁰⁷

It was not spectrum scarcity itself, but the result it produced—an endangered marketplace of ideas—that supported this leg of the Court’s reasoning. This is apparent from the Court’s answer to the argument that since the beginning of the broadcasting regulatory regime in 1929,¹⁰⁸ the airwave frequency spectrum has broadened significantly. The radio station in *Red Lion* argued that with the discovery of microwaves and other innovations, the frequency spectrum had increased and no longer justified the fairness doctrine.¹⁰⁹ The Court acknowledged that there were, in fact, unutilized “gaps” in the spectrum, but the Court also observed that the effects of spectrum scarcity remained because the broadcasting industry was dominated by a small number of organizations.¹¹⁰ Even though it was “technologically possible” for new entrants to begin broadcasting, that possibility could not by itself render the fairness doctrine unconstitutional.¹¹¹ Established broadcasters enjoyed a “substantial advantage” due to “[l]ong experience in broadcasting, confirmed habits of listeners and viewers, network affiliation, and other advantages.”¹¹² Existing outlets had unprecedented power over what was said and heard on the radio, and that justified the continuance of a pro-regulation First Amendment approach. The problem originally created by spectrum scarcity was thus perpetuated by the entrenchment of the gatekeepers, even though the spectrum had become less scarce.

The Internet is certainly not scarce, as *Reno* noted, but a small number of gatekeepers enjoy a dominance over social media platforms resembling the

¹⁰⁵ *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 390 (1969).

¹⁰⁶ *See id.* at 388.

¹⁰⁷ *Id.*

¹⁰⁸ Radio Act of 1927, Pub. L. N. 69-632, 47 U.S.C. §§ 81-83 (repealed 1934).

¹⁰⁹ *Red Lion*, 395 U.S. at 396–97.

¹¹⁰ *See id.* at 400.

¹¹¹ *Id.*

¹¹² *Id.*

dominance of broadcasting companies. It is, in fact, more severe and consolidated than the broadcasting industry ever was.¹¹³ Facebook, Google, and Twitter have the power to control what is viewed on their platforms, as well as what is posted. These companies are increasingly policing content on their websites, which Justice Kennedy identified as “the most important places . . . for the exchange of views.”¹¹⁴ Although the Internet does not possess a comparable physical scarcity comparable to the broadcasting industry in the 1930s, it shares the de facto gatekeeper scarcity that existed in broadcasting when *Red Lion* was decided.

Entrenchment of the existing platforms seems permanent given the barriers to enter into the social media platform market. What makes a social media platform desirable for users is, first and foremost, that it has already captured other users.¹¹⁵ Users are attracted to Facebook because their friends are on Facebook; they read Twitter posts because influential celebrities and public figures post on Twitter. A new platform would have to displace these online giants to succeed.¹¹⁶ For this reason, many, including Senator Elizabeth Warren and Facebook co-founder Chris Hughes, have called for Facebook to be broken up.¹¹⁷ The concentration of power in the social media context, as with broadcasting, threatens the natural competition of ideas. As Justice White wrote, “[i]t is the purpose of the First Amendment to preserve an uninhibited marketplace of ideas in which truth will ultimately prevail, rather than to

¹¹³ J. Clement, *Leading Social Media Websites in the United States in August 2019, Based on Share of Visits*, STATISTA, (Sept. 9, 2019), <https://www.statista.com/statistics/265773/market-share-of-the-most-popular-social-media-websites-in-the-us/> [<https://perma.cc/YV68-KU37>].

¹¹⁴ *Packingham v. North Carolina*, 137 S. Ct. 1730, 1735 (2017).

¹¹⁵ See Dipayan Ghosh, *Don't Break up Facebook—Treat it Like a Utility*, HARV. BUS. REV. (May 30, 2019), <https://hbr.org/2019/05/dont-break-up-facebook-treat-it-like-a-utility> [<https://perma.cc/7JA2-Q5HE>].

¹¹⁶ For example, in 2014, Ello tried to become a Facebook alternative, promising not to sell ads or to “sell data about you to third parties.” Ed Cumming, *Ello—and Goodbye to the New Facebook?*, GUARDIAN (Oct. 5, 2014), <https://www.theguardian.com/media/2014/oct/05/ello-and-goodbye-facebook-competitor-social-networking/> [<https://perma.cc/NH5A-PCB9>]. The company failed because it could not reach a “critical mass of users”; even if Facebook users were frustrated by its data tracking and advertising, they would not leave Facebook because “that’s where everyone is, and that’s the point of a social network.” Nick Srnicek, *We Need to Nationalise Google, Facebook and Amazon. Here’s Why*, GUARDIAN (Aug. 30, 2017), <https://www.theguardian.com/commentisfree/2017/aug/30/nationalise-google-facebook-amazon-data-monopoly-platform-public-interest/> [<https://perma.cc/S2CD-EQGD>].

¹¹⁷ Steven Overly, *Facebook Co-Founder Calls on the Government to Break Up the Company*, POLITICO (May 9, 2019), <https://www.politico.com/story/2019/05/09/facebook-cofounder-criticism-1313365/> [<https://perma.cc/KAH5-XFVK>]; Jason Abbruzzese, *Elizabeth Warren Calls to Break Up Facebook, Google, and Amazon*, NBC NEWS (Mar. 8, 2019), <https://www.nbcnews.com/tech/tech-news/elizabeth-warren-calls-break-facebook-google-amazon-n980911/> [<https://perma.cc/P6VX-DCAG>].

countenance monopolization of that market, whether it be by the Government itself or a private licensee.”¹¹⁸

The Court discussed the scarcity factor in *Reno*, dismissing *Red Lion* as inapplicable to the Internet because there was no spectrum scarcity; rather, there was an abundance.¹¹⁹ The Court opined,

[T]he Internet can hardly be considered a “scarce” expressive commodity. It provides relatively unlimited, low-cost capacity for communication of all kinds. . . . This dynamic, multifaceted category of communication includes not only traditional print and news services, but also audio, video, and still images, as well as interactive, real-time dialogue.¹²⁰

But the Internet’s information overload—its spectrum abundance, so to speak—contributes to the same ills that endangered speech in broadcasting in *Red Lion*. There, spectrum scarcity necessitated ordering to ensure that broadcasters used their airwaves responsibly.¹²¹ In the social media context, the abundance of information online forces content hosts to perform a filtering function, thus thrusting upon them a similar responsibility. One free speech advocate, who testified at a House Judiciary Committee hearing examining the filtering practices of social media, stated in his written testimony that “[c]ontent moderation is an inevitable part of the Internet. Website operators will always have to make judgments about what content to take down and what to leave up, monitor and moderate objectionable content, promote effective counter-speech, educate their users, and generally create healthy, positive and dynamic online communities.”¹²² Just as those in control of the scarce airwaves in *Red Lion* had a fiduciary-type responsibility in presenting content to their wide audiences, those in control of the over-abundance of content online have a responsibility to filter that content fairly.¹²³

¹¹⁸ *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 390 (1969).

¹¹⁹ *See Reno v. Am. Civil Liberties Union*, 521 U.S. 844, 870 (1997).

¹²⁰ *Id.*

¹²¹ *See Red Lion*, 395 U.S. at 389–92.

¹²² *An Internet ‘Fairness Doctrine’ Would Stifle Free Speech, Protect Tech Giants*, TECHFREEDOM (Apr. 25, 2018), <http://techfreedom.org/Internet-fairness-doctrine-stifle-free-speech-protect-tech-giants/> [<https://perma.cc/7ERL-DAAT>].

¹²³ *See Red Lion*, 395 U.S. at 389; Jonathan Zittrain, *How to Exercise the Power You Didn’t Ask For*, HARV. BUS. REV. (Sept. 19, 2018), <https://hbr.org/2018/09/how-to-exercise-the-power-you-didnt-ask-for> (discussing online platforms as “information fiduciaries”) [<https://perma.cc/ZH8U-7DTP>].

3. *Reach and Interference*

The third leg of *Red Lion*'s reasoning was that "the reach of radio signals is incomparably greater than the range of the human voice and the problem of interference is a massive reality."¹²⁴ Half the world could be speaking at the same time while the other half listened, the Court said, and even with so many voices employed at once, speech could flourish because the human voice is simply not loud enough to drown out other voices.¹²⁵ Radio broadcasting, on the other hand, reaches far enough that users of the same frequency drown each other out, even when they are far away.¹²⁶ This problem of reach and interference justified not only the licensing regime that divvied out exclusive privileges to use the airwaves, but also the fairness regulations that aimed to improve the marketplace of ideas by increasing representation of differing ideas. This factor builds on the first two because it exacerbates the threat posed by a consolidated industry.

Content on social media has a similarly wide reach as radio broadcasts. For instance, a single Tweet from President Trump on March 28, 2019 was retweeted 15,153 times and liked 54,032 times in seven hours.¹²⁷ A single YouTube video can have millions, or even billions, of views.¹²⁸ Although the reach of online media rivals that of broadcasting, the phenomenon of direct interference recognized in *Red Lion* is admittedly absent. One YouTube video will not drown out another, preventing it from being seen by audiences. Nevertheless, new media with a powerful reach has the potential to undermine the marketplace of ideas by displacing speech with other online expression. This is because the same organizations that dispense information provide a filtering service that prefers some information over others and that insulates radical posts from sources that disagree.¹²⁹

While a variety of ideas exist on the Internet, like ideas are packaged with like, and sensationalist news is overemphasized.¹³⁰ This is akin to the interference in *Red Lion*, but with a key difference: In *Red Lion*, if multiple

¹²⁴ *Red Lion*, 395 U.S. at 387–88. See also *FCC v. Pacifica Found.*, 438 U.S. 726, 748 (1978) (“[T]he broadcast media have established a uniquely pervasive presence in the lives of all Americans.”).

¹²⁵ See *Red Lion*, 395 U.S. at 387–88.

¹²⁶ See *id.*

¹²⁷ Donald Trump (@realDonaldTrump), TWITTER (Mar. 28, 2019, 2:41 PM), <https://twitter.com/realdonaldtrump/status/1111352552615043078> [<https://perma.cc/7Z8M-V7L9>].

¹²⁸ J. Clement, *Most Popular YouTube Videos Based on Total Global Views as of August 2019 (in Billions)*, STATISTA, (Aug. 19, 2019), <https://www.statista.com/statistics/249396/top-youtube-videos-views/> [<https://perma.cc/6KKB-3E8U>].

¹²⁹ See Vicario et al., *supra* note 103.

¹³⁰ See *id.*

radio speakers used the same airwaves, none of them would be heard and the marketplace of ideas would cease to exist. By contrast, when interference through filtering and displacement occurs on Internet platforms, the marketplace of ideas is instead distorted. Truth does not win over falsehood because false ideas are disseminated along with congruent ideas that match the same theme, so that their falsity frequently remains unexposed. This displacement, coupled with the far-ranging reach of online posts, threatens informed communication.

B. *Pacifica* factors

In addition to historic regulation, spectrum scarcity, and reach and interference, the *Pacifica* Court recognized two other factors that influenced how broadcasting should be treated under the First Amendment. As described in a later case, “[t]he *Pacifica* opinion . . . relied on the ‘unique’ attributes of broadcasting, noting that broadcasting . . . can intrude on the privacy of the home without prior warning as to program content, and is ‘uniquely accessible to children, even those too young to read.’”¹³¹ Broadcasting has the potential to be invasive because listeners might “turn[] on a radio and be[] taken by surprise by an indecent message.”¹³² This invasiveness was held not to apply to the Internet in *Reno* because “[c]ommunications over the Internet do not ‘invade’ an individual’s home or appear on one’s computer screen unbidden. Users seldom encounter content ‘by accident.’”¹³³

But *Reno* was decided in 1997 when social media did not exist.¹³⁴ Since then, social media has become a massive part of our lives, and video makes up a significant portion of its content. Facebook generates an average of eight billion video views per day, with five hundred million people viewing videos on Facebook daily.¹³⁵ Video content has at least as great a potential for invasiveness as radio content did in *Pacifica*; instead of merely hearing an offensive scene, video can surprise viewers with offensive imagery as well. And the sheer volume of content creators working under the cloak of anonymity exacerbates such invasiveness in a disturbing way. Earlier this

¹³¹ *Sable Commc’ns of California, Inc. v. FCC*, 492 U.S. 115, 127 (1989) (quoting *FCC v. Pacifica Found.*, 438 U.S. 726, 748–49 (1978)).

¹³² *Id.* at 128.

¹³³ *Reno v. Am. Civil Liberties Union*, 521 U.S. 844, 869 (1997).

¹³⁴ The first recognizable social media website, Six Degrees, was created later that year in 1997. Drew Hendricks, *Complete History of Social Media: Then and Now*, SMALL BUS. TRENDS (May 8, 2013), <https://smallbiztrends.com/2013/05/the-complete-history-of-social-media-infographic.html/> [<https://perma.cc/4FWY-X452>].

¹³⁵ Derek Doeing, *55+ Powerful Video Marketing Statistics for 2019*, G2CROWD.COM, (Aug. 9, 2018), <https://learn.g2crowd.com/video-marketing-statistics/> [<https://perma.cc/8XBJ-5NE3>].

year, a mother discovered a YouTube video on a children's website that contained instructions for how to self-harm, spliced in between clips of a popular Nintendo game.¹³⁶ Even though such videos violate the content policies of YouTube and other online platforms, they can receive tens of millions of views before they are taken down.¹³⁷ The ability to share these videos on social media thus creates the possibility that unexpected and unwanted content could invade a user's experience.

The other factor relied on in *Pacifica* was the accessibility of broadcasting to children.¹³⁸ Concern for children on the Internet was also addressed in Justice Alito's concurrence in *Packingham*.¹³⁹ Echoing the language of the *Red Lion* opinion, Justice Alito observed that "there are important differences between cyberspace and the physical world."¹⁴⁰ Among these was the heightened vulnerability of minors in cyberspace as compared to physical locations.¹⁴¹ Justice Alito warned that, "if the entirety of the internet or even just 'social media sites' are the 21st century equivalent of public streets and parks, then States may have little ability to restrict the sites that may be visited by even the most dangerous sex offenders."¹⁴² In accordance with Justice Alito's concern, many commentators have spoken out about the dangers that could befall children on social media and on the Internet as a whole.¹⁴³ Unsurprisingly, this has been a central focus of federal legislation¹⁴⁴ and constitutional review.¹⁴⁵ Because these *Pacifica* factors, in addition to the *Red Lion* factors, are implicated by social media, a different standard of constitutional review for First Amendment issues is justified.

IV. REGULATORY SOLUTIONS

Red Lion represents the Court's willingness to uphold laws affecting speech in new contexts, where the marketplace of ideas does not function

¹³⁶ Sophie Lewis, *Horrified Mom Discovers Suicide Instructions in Video on YouTube and YouTube Kids*, CBS NEWS (Feb. 23, 2019), <https://www.cbsnews.com/news/youtube-kids-inappropriate-horrified-mom-discovers-suicide-instructions-in-video-on-youtube-and-youtube-kids/> [<https://perma.cc/3NPQ-D2WL>].

¹³⁷ *The Great Momo Panic*, GIMLET, (Mar. 14, 2019), <https://gimletmedia.com/shows/reply-all/j4h6jd/138-the-great-momo-panic> [<https://perma.cc/688S-U28N>].

¹³⁸ See *FCC v. Pacifica Found.*, 438 U.S. 726, 749 (1978).

¹³⁹ See *Packingham v. North Carolina*, 137 S. Ct. 1730, 1739 (2017) (Alito, J., concurring).

¹⁴⁰ See *id.* at 1743 Alito, J., concurring).

¹⁴¹ See *id.*

¹⁴² *Id.*

¹⁴³ See Jodi Whitaker & Brad J. Bushman, *Online Dangers: Keeping Children and Adolescents Safe*, 66 WASH. & LEE L. REV. 1053 (2009).

¹⁴⁴ See 47 U.S.C. § 223 (2018); Allow States and Victims to Fight Online Sex Trafficking Act of 2017, Pub. L. No. 115-164, 132 Stat. 1253.

¹⁴⁵ See *Packingham*, 137 S. Ct. 1730; *Reno v. Am. Civil Liberties Union*, 521 U.S. 844 (1997).

properly on its own.¹⁴⁶ If social media is such a context, then the Court is more likely to uphold legislative solutions aimed at enhancing speech, whereas the Court would likely strike down comparable laws on First Amendment grounds in more traditional contexts, like print media. This Part outlines some of the possible legislative and administrative solutions that have been discussed, proposed, or enacted. Specifically, recent election transparency and anti-bias legislation represent two kinds of attempts to enhance speech on social media platforms by placing limitations on platform gatekeepers. These laws would become significantly more viable as a constitutional matter if a *Red Lion* approach is applied to the social media context. Because these efforts represent attempts to repair a broken or distorted marketplace of ideas brought to light either by Russian election interference or gatekeeper abuse of control, they are likely to revitalize the *Red Lion* debate in the courts.

A. Election transparency laws

In response to massive Russian interference with the 2016 U.S. presidential election,¹⁴⁷ both federal and state legislatures have proposed or passed election transparency laws. Advertisements on the radio and television—media contexts that receive more limited First Amendment protection—are currently subject to disclosure requirements.¹⁴⁸ Recent bills aim to extend those requirements to the Internet. If social media platforms are treated similarly to radio and television under a *Red Lion* theory, then these extensions will pass constitutional muster. But if a *Tornillo*-type standard applies, then these laws could potentially be struck down.

Several bills addressing online electioneering transparency have been passed at the state level.¹⁴⁹ In Maryland, for instance, the Online Electioneering Transparency and Accountability Act banned the use of foreign money to pay for political ads.¹⁵⁰ While some in the tech community had approved the Act,¹⁵¹ *The Washington Post* and other newspapers sued the state, arguing that, among other things, the Act compelled speech in violation

¹⁴⁶ See discussion *supra* Part I.A.

¹⁴⁷ S. SELECT COMM. INTELLIGENCE, 116TH CONG., REP. ON RUSSIAN ACTIVE MEASURES, CAMPAIGNS, AND INTERFERENCE IN THE 2016 U.S. ELECTION (2019).

¹⁴⁸ See, e.g., 11 C.F.R. § 110.11 (2019).

¹⁴⁹ See Democracy Protection Act, 2018 N.Y. Sess. Laws ch. 59, pt. JJJ (McKinney) (codified at N.Y. Elec. § 14-100, -106, -107, -126); 2018 Wash. Legis. Serv. ch. 304 (West); MD. CODE ANN., ELEC. LAW § 13-405 (West 2018).

¹⁵⁰ See MD. CODE ANN., ELEC. LAW § 13-405 (West 2018); Edward Ericson, Jr., *Newspapers Sue Over Electioneering Disclosure Law*, COURTHOUSE NEWS SERV. (Aug. 17, 2018), <https://www.courthousenews.com/newspapers-sue-over-electioneering-disclosure-law/> [https://perma.cc/A4JD-TYXU].

¹⁵¹ Ericson, *supra* note 150 (“Facebook told the newspaper that it was on board with the law.”).

of *Tornillo*.¹⁵² A district court reviewing the Act agreed with the plaintiffs and granted a preliminary injunction preventing enforcement of certain provisions.¹⁵³ However, the court declined to follow *Red Lion*, citing *Reno*.¹⁵⁴

At the federal level, a proposed electioneering bill called the Honest Ads Act would require online platforms like Facebook and Google to disclose information about the groups of people targeted by political ads, what those ads cost, and who ran them.¹⁵⁵ These companies would also have to make “reasonable efforts” to ensure that foreign nationals did not fund political ads.¹⁵⁶ The disclosure requirements would cover ads made about “national legislative issue[s] of public importance,” even though this scope raises constitutional concerns.¹⁵⁷ The government’s constitutional license to impose disclosure requirements on broadcasters is, of course, rooted in *Red Lion*’s relaxed standards for broadcasting.¹⁵⁸ Although the Honest Ads Act has not yet been passed, the Federal Election Commission has already initiated a proposed rulemaking process to require more disclaimers for Internet political ads.¹⁵⁹

The election disclosure bills and proposed regulations are likely to raise the *Red Lion* question in the courts. While these requirements are permissible in the broadcasting context, Russian interference in the 2016 election made it clear that social media poses a special threat to the democratic process. Advocates for these measures should seek an expansion of *Red Lion* to the Internet or argue that social media platforms present a new kind of space deserving a regulation-friendly approach for similar reasons to those given in *Red Lion*.

¹⁵² *Washington Post v. McManus*, 355 F. Supp. 3d 272 (D. Md. 2019).

¹⁵³ *See id.*

¹⁵⁴ *Id.* at 288 (“This rationale, the Court has made clear, is particular to broadcast communications and does not apply to cable transmissions or material posted on the Internet.”).

¹⁵⁵ Honest Ads Act, S. 198, 115th Cong. § 8 (2017).

¹⁵⁶ *Id.* at § 9.

¹⁵⁷ In *Buckley v. Valeo*, the D.C. Circuit rejected a part of a statute to the extent that it could be read to impose mandatory disclosure of ads on “issues of public importance.” 519 F.2d 821, 870 (D.C. Cir. 1975), *aff’d in part, rev’d in part*, 424 U.S. 1 (1976), and *modified*, 532 F.2d 187 (D.C. Cir. 1976). Although the case went to the Supreme Court, this issue was not discussed on appeal. *See* 424 U.S. 1 (1976).

¹⁵⁸ *See Well-Intentioned ‘Honest Ads’ Bill Raises Serious Free Speech Concerns*, TECHFREEDOM (Oct. 19, 2017), <http://techfreedom.org/well-intentioned-honest-ads-bill-raises-serious-free-speech-concerns/> (arguing the Honest Ads Act would be unconstitutional because the Internet is a different context from broadcasting) [<https://perma.cc/GH2B-872W>].

¹⁵⁹ *See FEC Public Hearing on Internet Disclaimers* FEC RECORD: REGULATIONS (July 18, 2018), <https://www.fec.gov/updates/public-hearing-internet-disclaimers-2018/> (discussing public hearings resulting after the notice proposed rulemaking) [<https://perma.cc/TT2U-QKP5>].

B. Anti-bias bills

Another form of speech-enhancing legislation of social media platforms includes proposals that target perceived bias on those platforms. Many lawmakers and commentators—primarily those on the right—have spoken out against bias in social media, both in filtering algorithms and in content-policing decision-making.¹⁶⁰ House bill H.R. 492, the “Biased Algorithm Deterrence Act,” proposes that social media platforms showing political bias be deprived of statutory protections from liability.¹⁶¹ Even though they have been withdrawn, similar bills have been presented at the state level.¹⁶²

These proposals, which somewhat resemble the fairness doctrine—ironically, a policy reviled by many conservatives—will undoubtedly raise speech concerns. Indeed, when Ted Cruz accused Facebook of suppressing conservative speech and raised the idea of limiting CDA protections to “neutral platforms,” speech advocate Berin Szóka’s response referred to these proposals as a “Fairness Doctrine for the Internet.”¹⁶³ In order for these laws to survive judicial scrutiny, an expansion of the *Red Lion* approach to the context of social media would seem necessary.

V. CONCLUSION

Justice Kennedy warned the judiciary in *Packingham* to be cautious in its approach to the rapidly and unpredictably evolving Internet.¹⁶⁴ But we do not know what this caution should look like; is it more cautious to take a regulation-friendly approach to Internet cases or a regulation-hostile approach? That all depends on which approach best preserves a well-functioning online marketplace of ideas. Given the recent use of media manipulation as a means of destabilizing our democracy and the structural realities giving tech giants vast reach and control, the more cautious approach would seem to be the regulation-friendly one. Although the rationales of *Red*

¹⁶⁰ See, e.g., Margaret Harding McGill, et al., *Trump Pushes Government Action Against “Terrible Bias” at Social Media Summit*, POLITICO (Jul. 11, 2019), <https://www.politico.com/story/2019/07/11/anti-tech-bill-sponsor-white-house-social-media-summit-1586902> [<https://perma.cc/DUK5-MMHN>]; Cecelia Kang & Sheera Frenkel, *Republicans Accuse Twitter of Bias Against Conservatives*, N. Y. TIMES (Sept. 5, 2018), <https://www.nytimes.com/2018/09/05/technology/lawmakers-facebook-twitter-foreign-influence-hearing.html> [<https://perma.cc/TYP8-UZLU>].

¹⁶¹ Biased Algorithm Deterrence Act, 116th Cong. (2019).

¹⁶² Stop Social Media Censorship Act, H.B. 1028, 92nd Gen. Assemb., Reg. Sess. (Ark. 2019); Stop Social Media Censorship Act, S.B. 1722, 2019 Leg., Reg. Sess. (Fla. 2019).

¹⁶³ See TECHFREEDOM, *supra* note 122.

¹⁶⁴ *Packingham v. North Carolina*, 137 S. Ct. 1730, 1736 (2017).

Lion do not directly map onto the attributes of social media platforms, the concerns underlying those rationales are implicated even more than they were by the broadcasting media fifty years ago. Given the Court's recognition that "differences in the characteristics of new media justify differences in the First Amendment standards applied to them,"¹⁶⁵ supporters of regulatory solutions to social media-related issues should look to *Red Lion* as the constitutional ground for their efforts.

¹⁶⁵ *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 386 (1969).

STUDENT WRITING COMPETITION

During the 2018–19 academic year, the *Georgetown Law Technology Review* was pleased to conduct its first annual student writing competition. This year's topic invited entrants to address a legal or public policy question relating to artificial intelligence, machine learning, the use of data analytics or algorithmic decision-making. The competition was open to any student at an ABA-accredited law school, and GLTR received over fifty entries. These entries were judged by a panel consisting of law professors, members of GLTR, and professionals in the law and technology policy field. Thank you to all of our judges for helping select these pieces.

Our top three, cash prize winners are:

- First Place: Lauren Renaud with her piece *Will You Believe It When You See It?: How and Why the Press Should Prepare for Deepfakes*.
- Second Place: Thomas Belcastro with his paper *How the Business Judgment Rule Should Apply to Artificial Intelligence Devices Serving as Members of a Corporate Board*.
- Third Place: Theodore Bruckbauer with his piece *CFIUS and A.I.: Defending National Security While Allowing Foreign Investment*.

Congratulations to all!

For those interested in competing in this year's competition, submissions are due on May 31, 2020. Entrants are invited to submit papers exploring the ways in which emerging technologies and services interact with or challenge existing civil rights and consumer protection laws. Suggested topics include: questions relating to the adequacy of existing laws, administrative structures and processes to protect consumers and civil rights; proposals for updating legal and administrative frameworks to increase their relevance for emerging technologies; questions relating to privacy, bias, discrimination, or disparate impact on marginalized communities; the potential for technology to improve access, equity and accountability in the justice system; or subject-matter-specific legal issues arising from various applications of a technology.

Please visit <https://www.georgetowntech.org/writingcompetition> for more information.

Last and certainly not least thanks to the generous contributions of the BSA: The Software Alliance for supporting these competitions.

WILL YOU BELIEVE IT WHEN YOU SEE IT? HOW AND WHY THE PRESS SHOULD PREPARE FOR DEEPPAKES

Lauren Renaud*

CITE AS: 4 GEO. L. TECH. REV. 241 (2019)

TABLE OF CONTENTS

I. INTRODUCTION	241
II. BACKGROUND.....	243
A. Technical Details	243
B. Current Capabilities & Potential Uses	245
C. Existing Verification Methods Must Be Supplemented	247
III. CONSEQUENCES OF FAILING TO ADDRESS THE THREAT.....	249
A. Government Responses.....	249
1. <i>Criminal Prohibitions of Deepfakes</i>	250
2. <i>Amend Section 230 Immunity</i>	253
B. Loss in Public Trust	255
IV. SOLUTIONS FOR THE PRESS TO PURSUE.....	257
A. Technology-based Solutions.....	258
1. <i>Digital Authentication</i>	258
2. <i>Digital Signatures</i>	259
3. <i>AI Watermarks</i>	260
B. Education, Dialogue, and Collaboration.....	260
V. CONCLUSION.....	261

I. INTRODUCTION

On April 23, 2013, the Associated Press Twitter account tweeted “Breaking: Two Explosions in the White House and Barack Obama is

* Georgetown Law, J.D. 2019. Lauren Renaud is an attorney with the U.S. Department of Justice. The views expressed in this paper are those of the author and do not necessarily represent the views of the Department of Justice or the United States. The author would like to thank her parents, Laura and Jerry Renaud, for their encouragement and support and Professor Erin Carroll for her insightful feedback on the topic and substance of this paper.

injured.”¹ While the claim was false and the result of a hack, its effects were very real: in three minutes the Dow Jones Industrial Average plummeted \$136 billion in market value.² The market recovered after the White House and Associated Press scrambled to refute the claim, but what if the tweet had been equally alarming yet far less rebuttable? A tweet falsely warning of an impending attack on the White House accompanied by a realistic video of terrorists planning for it would lead to a panic that could not be stopped by a single retraction tweet. The technology to fabricate such a video already exists, and future advances will render forgeries undetectable to the human eye.

Government officials and scholars have begun raising the alarm about the danger posed by these forgeries, commonly known as deepfakes, but few have focused on their implications for the press.³ As deepfakes become more realistic and more widely used, the press will play a unique role due to its ability to give credence to and broadcast a deepfake. The press is also uniquely vulnerable to deepfakes because the news industry relies on the trust of its viewers. Because future deepfakes may affect viewers’ perceptions of truth, that trust may be impacted.

The press has traditionally adopted practices to verify source information and media. However, as will be explained, these current practices must be supplemented to adequately address deepfakes. If the press does not increase efforts to prepare for deepfakes, the government will fill the void with measures that are more restrictive and less desirable to the press. Three already proposed government measures would criminalize certain uses of deepfakes and potentially impose criminal liability on journalists who publish them.⁴ Additionally, if deepfakes are published as legitimate content, news

¹ Max Fisher, *Syrian Hackers Claim AP Hack That Tipped Stock Market By \$136 Billion. Is It Terrorism?*, WASH. POST (Apr. 23, 2013), https://www.washingtonpost.com/news/worldviews/wp/2013/04/23/syrian-hackers-claim-ap-hack-that-tipped-stock-market-by-136-billion-is-it-terrorism/?utm_term=.4737b70139d1 [https://perma.cc/SC5Q-EV9Y].

² Fisher, *supra* note 1; Shawn Langlois, *This Day In History: Hacked AP Tweet About White House Explosions Triggers Panic* (Apr. 23, 2018, 2:08 PM), <https://www.marketwatch.com/story/this-day-in-history-hacked-ap-tweet-about-white-house-explosions-triggers-panic-2018-04-23> [perma.cc/3M68-R4YT].

³ See Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CAL. L. REV. 1753, 1784–85 (2019); Press Release, Senator Mark Rubio, VIDEO: At Intelligence Committee Hearing, Rubio Raises Threat Chinese Telecommunications Firms Pose to U.S. National Security (May 15, 2018), <https://www.rubio.senate.gov/public/index.cfm?p=Press-Releases&id=B913F422-DC4F-4F19-A664-D9CE70559F87> [https://perma.cc/9GLA-DQ4S]; *Worldwide Threat Assessment of the US Intelligence Community: Hearing before the S. Select Comm. on Intelligence*, 116th Cong. 66–67 (2019) (statement of Senator Angus King, Member, S. Select Comm. on Intelligence) [hereinafter *Worldwide Threat Hearing*].

⁴ See *infra* Section III.A.1.

organizations will experience a loss of public trust in news—a trust that is fundamental to the news industry’s ethos and business model. To avoid such government measures and loss of public trust, the press should begin educating journalists about current deepfake capabilities and experiment with counter-deepfake technology.

II. BACKGROUND

Deepfakes have already begun to impact the press, and their impact will increase as the technology improves and diffuses. This Part will briefly explain deepfake technology, provide an overview of how the technology is being used to impact the press, and illustrate why current digital verification techniques will not be sufficient to address deepfakes.⁵ A brief note on terminology: this paper uses “the press” and “news organizations” to collectively refer to news-producing entities—*e.g.*, newspapers, TV news, bloggers—and not *conduits* for news, most notably social media platforms.⁶

A. Technical Details

What sets deepfakes apart from previous forgery technology is not just that they are a more realistic product but also how they are created. Current deepfake technology uses artificial intelligence (AI), specifically neural networks, to produce images, audio, and videos (audiovisuals) far more realistically and quickly than a human could create tinkering on Photoshop. Neural networks are complex systems of interconnected processing nodes loosely modeled after the human brain.⁷ Neural networks learn to recognize

⁵ This paper uses the term “deepfakes” to refer only to artificial intelligence-assisted alteration or generation of images, audio, or video. Therefore, merely duplicating and splicing frames in a video, such as the recent Jim Acosta-White House Intern video, does not constitute a deepfake. See Drew Harwell, *White House Shares Doctored Video To Support Punishment Of Journalist Jim Acosta*, WASH. POST (Nov. 8, 2018), https://www.washingtonpost.com/technology/2018/11/08/white-house-shares-doctored-video-support-punishment-journalist-jim-acosta/?utm_term=.f2e2623891f2 [<https://perma.cc/8NVG-NTP8>].

⁶ Though social media companies play an important role in the news cycle, entities serving merely as a conduit for news will be impacted by deepfakes differently than news-producing entities and are thus outside the scope of this paper.

⁷ See Larry Hardesty, *Explained: Neural Networks*, MIT NEWS (Apr. 14, 2017), <http://news.mit.edu/2017/explained-neural-networks-deep-learning-0414> [<https://perma.cc/PS2U-G7TM>].

patterns from data; generally, the more training data fed into a network, the more accurate the model will be.⁸

The use of neural networks is not a new phenomenon, nor is it exclusive to deepfakes. Facebook's image tagging service, for example, uses neural networks to recognize users' faces in images uploaded to the site.⁹ The increasing sophistication of these techniques, however, has allowed networks to be trained beyond merely recognizing patterns to identify content; networks can now work in reverse and use patterns to alter or even generate content, with increasing realism.¹⁰ Though machine learning has not yet achieved true photorealism—a perfect, undetectable fake—the generative adversarial network (GAN) approach is pushing deepfakes towards that goal. The GAN approach involves two neural networks: a generative network which runs the desired model, and a discriminator network that checks the work of the generative network in real-time by assessing the degree to which the generative network succeeded.¹¹ The discriminator network feeds its results to the generative network which uses the feedback to improve its own output.¹² GANs, in a sense, train themselves and self-perpetuate the arms race between deepfake technology and counter-deepfake technology. The GAN approach has a high potential to improve the accuracy of generated images and eventually generate highly realistic videos.¹³

⁸ Worldwide Threat Hearing, *supra* note 3, at 82 (statement of Def. Intelligence Agency Dir. Lt. Gen. Robert Ashley) (“How do you get deep fakes that are really, really good? Lots of data—that’s how you train your algorithms.”); Chesney & Citron, *supra* note 3, at 1759; Will Knight, *Real or Fake? AI Is Making It Very Hard to Know*, MIT TECH. REVIEW (May 1, 2017), <https://www.technologyreview.com/s/604270/real-or-fake-ai-is-making-it-very-hard-to-know/> [<https://perma.cc/MX6E-QW7Y>]; Hardesty, *supra* note 7; Cade Metz & Keith Collins, *How an A.I. ‘Cat-and-Mouse Game’ Generates Believable Fake Photos*, N.Y. TIMES (Jan. 2, 2018), <https://www.nytimes.com/interactive/2018/01/02/technology/ai-generated-photos.html> [<https://perma.cc/S3TC-WKNW>].

⁹ Lily Hay Newman, *Facebook Can Even ID You in Photos Where Your Face Isn’t Showing*, SLATE (June 23, 2015, 1:54 PM), <https://slate.com/technology/2015/06/facebooks-new-machine-vision-algorithm-can-identify-people-without-their-faces.html> [<https://perma.cc/8EGB-HEDG>].

¹⁰ Metz & Collins, *supra* note 8.

¹¹ Chesney & Citron, *supra* note 3, at 1760; Ian J. Goodfellow, et. al., *Generative Adversarial Networks*, 1–2 (arXiv:1406.2661v1, 2014), <https://arxiv.org/pdf/1406.2661.pdf> [<https://perma.cc/CHX4-H4MG>].

¹² Chesney & Citron, *supra* note 3, at 1760; Goodfellow, *supra* note 11 at 1–2.

¹³ The GAN approach has already seen limited success in generating videos. See Carl Vondrick, et. al., GENERATING VIDEOS WITH SCENE DYNAMICS (2016), <http://www.cs.columbia.edu/~vondrick/tinyvideo/> (website created in conjunction for the paper, *Generating Videos with Scene Dynamics*, submitted to the 29th Conference on Neural Information Processing Systems) [<https://perma.cc/XW6J-R5XP>].

B. Current Capabilities & Potential Uses

As noted, there are no publicly known instances of a perfectly undetectable deepfake, but increasingly convincing audiovisual examples have been produced. Importantly, not all deepfakes are malicious—many are educational or further the arts or other professions.¹⁴ That said, perhaps the most widespread and obscene use of deepfakes is deepfake pornography, popularized in part by a Reddit forum, r/deepfakes, that circulated a tool allowing users to superimpose a celebrity's (or anyone's) face onto pornography videos.¹⁵ Of course, there is now an app for that, several in fact, which use AI to superimpose faces in videos, pornography or otherwise.¹⁶ These browser-based apps make it very easy to produce deepfakes, though the results vary in quality.¹⁷

If web-based browser apps represent the low end of the sophistication spectrum, university researchers, technology companies, and governments represent the upper end. These entities are likely to have more processing power and better technology, as well as time and access to caches of high-quality photos. Researchers at the University of Washington, for example, created a tool that alters videos to change the speech of the video's speaker.¹⁸ Meanwhile, technology companies have successfully altered audio clips of politicians.¹⁹ Little is publicly known about the United States government's or foreign governments' capabilities to create deepfakes. In 2015, the United States spent roughly \$1.1 billion on unclassified artificial intelligence research and development, but it is unknown if any of those funds were spent

¹⁴ Chesney & Citron, *supra* note 3, at 1769–71.

¹⁵ Chesney & Citron, *supra* note 3, at 1763; Jaime Dunaway, *Reddit (Finally) Bans Deepfake Communities, but Face-Swapping Porn Isn't Going Anywhere*, SLATE (Feb. 8, 2018, 4:27 PM), <https://slate.com/technology/2018/02/reddit-finally-bans-deepfake-communities-but-face-swapping-porn-isnt-going-anywhere.html> [<https://perma.cc/U7F8-H3WQ>].

¹⁶ See, e.g., DeepFakesApp, DEEPFAKESAPP <https://deepfakesapp.online> (advertising that “DFs may be used to create fake celebrity pornographic videos or revenge porn” on the website's home page) (accessed Oct. 30, 2019) [<https://perma.cc/7T5P-9CMQ>]; FakeApp, MALAVIDA, <https://www.malavida.com/en/soft/fakeapp/#gref> (accessed (Oct. 30, 2019) [<https://perma.cc/TW2W-QGLU>].

¹⁷ A New York Times reporter, for example, was able to create a semi-realistic deepfake video of himself superimposed onto Chris Pratt's body using 1,841 photos of himself. Adam Dodge, et. al., *Using Fake Video Technology to Perpetrate Intimate Partner Abuse*, WITHOUT MY CONSENT 5 (2018), <https://withoutmyconsent.org/perch/resources/2018-04-25deepfakedomesticviolenceadvisory.pdf> [<https://perma.cc/T2ND-NT3W>].

¹⁸ Chesney & Citron, *supra* note 3, at 1760.

¹⁹ *Id.* at 1761.

researching deepfake-related technology.²⁰ However, it is known that the Department of Defense's Defense Advanced Research Project Agency is currently supporting the development of counter-deepfake technology.²¹

There are many potential ways deepfake technology could impact or is impacting journalism. News organizations, for example, have begun to experiment with deepfakes. BuzzFeed's CEO coordinated with director Jordan Peele to create and release what it called a PSA video on deepfakes that altered video and audio of President Barack Obama.²² Meanwhile, a television news agency in China created an English-speaking "AI-anchor"—a computer-generated news anchor with the "facial expressions and actions of a real person."²³ Although the automated tone of the AI-anchor gives its true nature away, its visuals are impressive and its developers contend the AI-anchor will become more realistic over time.²⁴

As the introduction suggested, deepfakes could be used to create fake breaking news which may be unwittingly or intentionally broadcasted.²⁵ Strategically released deepfakes could amplify their effects. For example, a deepfake video of a police altercation might seek to inflame tensions after a real-life police altercation or, as Senator Marco Rubio of Florida has noted, a

²⁰ GREG ALLEN & TANIEL CHAN, BELFER CTR., ARTIFICIAL INTELLIGENCE AND NATIONAL SECURITY 52 (2017), <https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf> [<https://perma.cc/UM3Z-VDTV>].

²¹ Matt Turek, *Media Forensics (MediFor)*, DARPA, <https://www.darpa.mil/program/media-forensics> [<https://perma.cc/KM33-KM8L>]; Will Knight, *The Defense Department Has Produced The First Tools For Catching Deepfakes*, MIT TECH. REV. (Aug. 7, 2018), <https://www.technologyreview.com/s/611726/the-defense-department-has-produced-the-first-tools-for-catching-deepfakes/> [<https://perma.cc/BDT7-PJVD>].

²² David Mack, *This PSA About Fake News from Barack Obama Is Not What It Appears*, BUZZFEED (Apr. 17, 2018, 11:26 AM), <https://www.buzzfeednews.com/article/davidmack/obama-fake-news-jordan-peepe-psa-video-buzzfeed> [<https://perma.cc/67FX-8A64>]. The video showed President Obama speaking about the dangers of deepfakes then transitioned to a side-by-side view of President Obama and Peele speaking the same words. In reality, President Obama never said any of things heard in the video; prior audio and video of him had been altered to make it realistically seem as if he was speaking those words. *Id.*

²³ Merrit Kennedy, *AI News Anchor Makes Debut In China*, NPR (Nov. 9, 2018, 6:11 PM), <https://www.npr.org/2018/11/09/666239216/ai-news-anchor-makes-debut-in-china> [<https://perma.cc/QF4V-74ZZ>].

²⁴ Taylor Telford, *These News Anchors Are Professional and Efficient. They're Also Not Human*, WASH. POST (Nov. 9, 2018, 11:40 AM), https://www.washingtonpost.com/business/2018/11/09/these-news-anchors-are-professional-efficient-theyre-also-not-human/?utm_term=.8ff2bd548b32 [<https://perma.cc/3D4Z-SE6T>]. Currently, the AI-anchor is trained with live broadcasting videos and social media and requires "only 10 minutes of data to effectively mimic a person's voice." *Id.*

²⁵ See *supra* Part I. There is already one known instance of a TV news station broadcasting a doctored video that it represented as true. See *infra*, Section III.B.

deepfake image of a political candidate engaging in compromising behavior could be timed to the days prior to an election.²⁶ Referring to the United States 2020 presidential election and beyond, then-Director of National Intelligence Daniel Coats stated “[a]dversaries and strategic competitors probably will attempt to use deepfakes or similar machine-learning technologies to create convincing—but false—image, audio, and video files to augment influence campaigns directed against the United States and our allies and partners.”²⁷ The topical relevance of certain deepfakes would shorten the timeframe a news organization has in which to verify the media at issue.

Additionally, deepfake technology can be used to target, undermine, or impersonate journalists. Sadly, there has already been at least one instance of this. As a result of her reporting, Indian investigative journalist Rana Ayyub endured hateful commentary online which escalated when a pornographic video with her face superimposed onto another woman went viral.²⁸ Despite technical confirmation that the video was a fake, the video spread via social media.²⁹ Deepfakes could also be used to undermine journalists by placing them in compromising situations, such as accepting a bribe or colluding with a politician before a political debate. Deepfakes can also be used to impersonate journalists. Recently, a fake Twitter account impersonated a “Senior Journalist at Bloomberg” and used what is likely an AI-generated image as its Twitter profile picture.³⁰

C. Existing Verification Methods Must Be Supplemented

The press has long verified source-provided information and audiovisuals (hereinafter collectively referred to as user-generated content) to ensure its veracity before publication. Indeed, the Society of Professional Journalists’ Code of Ethics (Code) instructs “Journalists should . . . [v]erify information before releasing it.”³¹ The Society notes that the Code is merely a

²⁶ 164 CONG. REC. S5010 (daily ed. July 17, 2018) (statement of Senator Rubio).

²⁷ Worldwide Threat Hearing, *supra* note 3, at 17 (statement of Dir. Nat’l Intelligence Daniel Coats).

²⁸ Rana Ayyub, *In India, Journalists Face Slut-Shaming and Rape Threats*, N.Y. TIMES, (May 22, 2018), <https://www.nytimes.com/2018/05/22/opinion/india-journalists-slut-shaming-rape.html> [<https://perma.cc/R5EK-GK2M>]; *see also* Danielle Citron, *Sexual Privacy*, 128 YALE L.J. 1870, 1922 (2019).

²⁹ Ayyub, *supra* note 28.

³⁰ Glenn Fleishman, *How to Spot the Realistic Fake People Creeping Into Your Timelines*, FAST CO. (Apr. 30, 2019), <https://www.fastcompany.com/90332538/how-to-spot-the-creepy-fake-faces-who-may-be-lurking-in-your-timelines-deepfakes> [<https://perma.cc/4URY-SWG7>]; Sean O’Kane (@sokane1), TWITTER (Mar. 27, 2019, 2:54 PM), <https://twitter.com/sokane1/status/1111023838467362816> [<https://perma.cc/L6J5-DVLA>].

³¹ *SPJ Code of Ethics*, SOC’Y OF PROF. JOURNALISTS (Sept. 6, 2014, 4:49 PM), <https://www.spj.org/ethicscode.asp> [<https://perma.cc/K3WS-DAD6>].

guide, not “a set of rules” or legally enforceable,³² but its principles—especially regarding verification—are foundational to the journalism field and critical to its success.³³ The Code’s instruction to verify user-generated content are reflected in the ethics guidelines of many news entities.³⁴

The Code directs journalists to various resources on verification, including the *Verification Handbook: An Ultimate Guideline on Digital Age Sourcing for Emergency Coverage* (Handbook).³⁵ The Handbook specifically addresses how to verify images and video, but the principles and techniques advocated focus primarily on contextualizing user-generated content and will not be sufficient to verify future deepfakes. The Handbook, for example, suggests videos be verified via a five-step process which involves identifying the source, investigating the source, identifying or confirming the video’s location and date, and ensuring the video depicts what it says it depicts.³⁶ Smart and well-resourced creators, however, could anticipate steps journalists will take to verify and avoid common pitfalls such as using audiovisuals already on the Internet and giving away unwanted clues in the background. AI-generated content is literally tailor-made and, while it is currently visually or forensically detectable, it is constantly being improved. For this reason, current practices alone will be insufficient.

³² *Id.* The Society of Professional Journalists has nearly 250 chapters and 6,000 total members. *About SPJ*, SOC’Y OF PROF. JOURNALISTS, <https://www.spj.org/aboutspj.asp> [<https://perma.cc/K7FC-FMDC>]; *Chapters*, SOC’Y OF PROF. JOURNALISTS, <https://www.spj.org/chapters.asp> [<https://perma.cc/KH3A-4LMG>]; *See SPJ Code of Ethics*, *supra* note 31.

³³ *See generally* BILL KOVACH & TOM ROSENSTIEL, *ELEMENTS OF JOURNALISM* (3rd ed., 2013); *see also infra* Section III.B;

³⁴ *See, e.g., Guidelines on Integrity*, N.Y. TIMES (Sept. 25, 2008), <https://www.nytimes.com/editorial-standards/guidelines-on-integrity.html> (“it is imperative that The Times and its staff maintain the highest possible standards to insure that we do nothing that might erode readers’ faith and confidence in our news columns . . . Images in our pages, in the paper or on the Web, that purport to depict reality must be genuine in every way.”) [<https://perma.cc/TJS5-YYF6>]; *Journalistic Integrity*, WARNER MEDIA GRP. (Nov. 07, 2016), <https://www.warnermediagroup.com/company/corporate-responsibility/telling-the-worlds-stories/journalistic-integrity> [<https://perma.cc/A64P-VNC3>]; *Los Angeles Times Ethics Guidelines*, L.A. TIMES (June 16, 2014, 12:00 AM), <https://www.latimes.com/la-times-ethics-guidelines-story.html> (“Photographs and graphics must inform, not mislead. Any attempt to confuse readers or misrepresent visual information is prohibited.”) [<https://perma.cc/SHQ5-K9QC>].

³⁵ *See*, EUROPEAN JOURNALISM CENTRE, *VERIFICATION HANDBOOK: AN ULTIMATE GUIDELINE ON DIGITAL AGE SOURCING FOR EMERGENCY COVERAGE* (Craig Silverman, ed. 2016), <https://verificationhandbook.com/downloads/verification.handbook.pdf> [<https://perma.cc/XEZ6-U2CH>].

³⁶ *Id.* at 47–53.

III. CONSEQUENCES OF FAILING TO ADDRESS THE THREAT

To date, the use of deepfake technology in ways that impact journalism is largely anecdotal. The technology will improve and diffuse, however, and the news industry should not wait until it does to adapt its practices and seriously address deepfakes. This Part highlights several potential consequences that may befall the news industry should it fail to appreciate the threat that deepfakes present.

A. Government Responses

Deepfakes already have Congress's and the intelligence community's attention. Members of Congress have repeatedly warned of the threat deepfakes present and of the government's need to be a part of the solution.³⁷ Other members have sought to ensure the intelligence community has the legal authority and funding it needs to address deepfakes.³⁸ The Director of National Intelligence, FBI Director, Defense Intelligence Agency Director, and National Geospatial Agency Director have all expressed their concern about the threat deepfakes pose.³⁹ This is all to underscore that the Executive and Legislative branches are starting to think about deepfakes, and their concern and motivation to seek solutions will only grow as deepfake technology

³⁷ Worldwide Threat Hearing, *supra* note 3, at 82 (statement of Senator Ben Sasse, Member, S. Select Comm. on Intelligence) (“The asymmetric exposure we have where the barrier to entry for deep fakes technology is so low now – lots of entities short of nation-state actors are going to be able to produce this material and again destabilize not just American public trust but markets very rapidly.”)

³⁸ 164 CONG. REC. S5010 (daily ed. July 17, 2018) (statement of Sen. Rubio); Press Release, Senator Maggie Hassan, Senator Hassan Presses Counterterrorism Official on Size of ISIS, Urges FBI Director to Crack Down on “Deepfakes” (Oct. 11, 2018), <https://www.hassan.senate.gov/news/in-the-news/senator-hassan-presses-counterterrorism-official-on-size-of-isis-urges-fbi-director-to-crack-down-on-deepfakes> [<https://perma.cc/5274-HLGA>]; Letter from Members of Congress Adam Schiff, Stephanie Murphy, & Carlos Cabelo to Daniel Coats, Dir. Of Nat'l Intelligence (Sept. 13, 2018), https://murphy.house.gov/uploadedfiles/2018-09_odni_deep_fakes_letter.pdf [<https://perma.cc/VZY7-5M2Z>].

³⁹ Press Release, Senator Maggie Hassan, *supra* note 38; OFF. OF THE DIR. OF NAT'L INTELLIGENCE., THE AIM INITIATIVE 1 (document undated), <https://www.dni.gov/files/ODNI/documents/AIM-Strategy.pdf> [<https://perma.cc/GV8U-27LV>]; WWTH Director Ashley; Worldwide Threat Hearing, *supra* note 3, at 71 (NGA Director Robert Cardillo) (“As [deepfake] technology advances – and it will – I do worry about, as a community that needs to seek the truth and then speak the truth, in a world in which we can't agree on what is true our job becomes much more difficult.”); *id.* at 82 (statement of Def. Intelligence Agency Dir. Lt. Gen. Robert Ashley) (“Our challenge is how do you build the algorithm to identify the anomaly? Because every deep fake has a flaw, or at least now they do.”).

improves. What follows is an analysis of two suggested solutions which may in turn impact journalism.

1. *Criminal Prohibitions of Deepfakes*

Bills banning some form of deepfakes have been proposed in both the California legislature and the United States Congress. As originally proposed, the California bill would make it a misdemeanor to “willfully distribut[e] a deceptive recording that the person knows, or *reasonably should have known*, is a deceptive recording” that is likely to deceive a viewer or “defame, slander or embarrass the subject of the recording.”⁴⁰ The statute does not define “reasonably should have known,” and it is unclear what the standard of care would be. If a journalist follows the verification steps detailed in Part II yet fails to realize a video is a deepfake and publishes it, is she liable? If one hundred journalists believe a video to be real but readily-available software would flag it as a fake, does their failure to use digital verification mean they reasonably should have known it was a fake? This unknown should trouble journalists.

The Malicious Deep Fake Prohibition Act of 2018, introduced in the Senate by Senator Ben Sasse of Nebraska, would make it unlawful to distribute an audiovisual with “[1] actual knowledge that the audiovisual record is a deepfake; and [2] the intent that the distribution of the audiovisual record would facilitate criminal or tortious conduct.”⁴¹ The bill’s prohibition is

⁴⁰ A.B. 602, 2019 Leg., Reg. Sess (Cal. 2019), https://leginfo.legislature.ca.gov/faces/billVersionsCompareClient.xhtml?bill_id=201920200AB602&cversion=20190AB60299INT [<https://perma.cc/C7ND-EUYA>] (emphasis added). The bill has since been amended and was signed into law on October 3, 2019. The final version does not impose criminal liability; instead, it provides a private right of action for victims of nonconsensual deepfake pornography, except material that has “newsworthy value.” Its original form is nevertheless discussed because the original proposal may be used as a model by other states in the future. All subsequent references to the California bill are to its original proposed form. *See id.* California has also passed a bill intended to prevent deepfakes from influencing elections. The bill prohibits persons and entities from distributing a deepfake involving an election candidate within 60 days of an election, unless the deepfake includes a disclaimer identifying the content as manipulated. The bill requires actual malice and the intent to injure a candidate’s reputation or deceive a voter. Candidate victims can seek injunctive or equitable relief, as well as monetary damages. A.B. 730, 2019 Leg, Reg. Sess., (Cal. 2019), https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB730 [<https://perma.cc/DD6N-XFR8>].

⁴¹ Malicious Deep Fake Prohibition Act of 2018, S. 3085, 115th Cong. § 2(a) (2018), <https://www.congress.gov/115/bills/s3805/BILLS-115s3805is.pdf> [<https://perma.cc/9MTP-76BW>]. The bill has since expired, but Senator Sasse reportedly intends to reintroduce it. *The Newest Front in the Deepfakes War: Does Congress Need to Step In?*, COUNTABLE (Jan. 31, 2019),

largely superfluous as using a deepfake to facilitate a crime is already a crime.⁴² The Act's prohibition should nevertheless be concerning to journalists because it prohibits the distribution of deepfakes to facilitate any tortious conduct—mirroring broad language in the Computer Fraud and Abuse Act (CFAA).⁴³

Regarding the CFAA's broad invocation of tortious acts, Professor Orin Kerr commented that “[f]ederal prosecutors have been very creative in coming up with such crimes and torts.”⁴⁴ Victoria Baranetsky has noted that corporations have successfully used the CFAA to challenge web scraping by “treating the site’s terms of service as a contract and prohibiting the act therein.”⁴⁵ If platforms and websites were to ban deepfakes—or certain uses of deepfakes, as some have already done⁴⁶—in its terms of service, knowingly posting a news story containing a deepfake on those platforms may violate the Act’s prohibition under the terms of service-contract theory. Since the Act does not require malicious intent, educational or parody deepfakes may be swept into the prohibition.⁴⁷

For example, if YouTube were to prohibit deepfakes in its terms of service, BuzzFeed would potentially be criminally liable for having posted its highly-viewed deepfake PSA depicting President Obama alerting the public to the dangers of deepfakes.⁴⁸ Or, borrowing on an extreme example used by Professor Kerr, BuzzFeed would potentially be committing a felony in violation of the Act if instead of posting the deepfake on YouTube, it had released the video at a party that had music “so loud that the party is tortious under state law, as it’s a private nuisance.”⁴⁹ Since the loud party’s purpose is

<https://www.countable.us/articles/20740-newest-front-deepfakes-war-does-congress-need-step> [<https://perma.cc/C4WC-ELMB>].

⁴² Blackmailing an individual, for example, by threatening to release a deepfake would already violate the criminal prohibition on blackmail. *See* 18 U.S.C. § 873 (2018).

⁴³ *See* 18 U.S.C. § 1030(c)(2)(B)(ii).

⁴⁴ Orin Kerr, *Should Congress Pass A "Deep Fakes" Law?*, VOLOKH CONSPIRACY (Jan. 31, 2019, 6:05 PM), <https://reason.com/volokh/2019/01/31/should-congress-pass-a-deep-fakes-law> [<https://perma.cc/6JJX-KJYW>].

⁴⁵ D. Victoria Baranetsky, *Data Journalism and the Law*, TOW CENTER FOR DIGITAL JOURNALISM (Sept. 19, 2018), https://www.cjr.org/tow_center_reports/data-journalism-and-the-law.php#newsgathering [<https://perma.cc/W423-NECB>].

⁴⁶ *Reddit Finally Bans Deepfake Communities but Face Swapping Porn Isn't Going Anywhere*, SLATE (Feb. 8, 2018, 4:27 PM), <https://slate.com/technology/2018/02/reddit-finally-bans-deepfake-communities-but-face-swapping-porn-isnt-going-anywhere.html> [<https://perma.cc/6QUL-AL7S>].

⁴⁷ The First Amendment may protect these uses and will be discussed below.

⁴⁸ *See* BuzzFeedVideo, *You Won't Believe What Obama Says In This Video!*, YOUTUBE (Apr. 17, 2018), <https://www.youtube.com/watch?v=cQ54GDm1eL0> [<https://perma.cc/EX33-D3WJ>].

⁴⁹ Kerr, *supra* note 44.

to unveil the deepfake, BuzzFeed's distribution of the deepfake at the party would be facilitating tortious conduct. The likelihood that prosecutors would elect to prosecute BuzzFeed for the party under the Act is slim, as are the odds that YouTube will categorically ban deepfakes on its platform,⁵⁰ but these scenarios illustrate the broad and potentially unintended effects criminal prohibitions of deepfakes might have.

Additionally, the DEEPFAKES Accountability Act, introduced in the House of Representatives in June 2019, would require deepfake producers to include easily visible disclaimers or watermarks to identify the content as altered.⁵¹ Any person who knowingly fails to comply with the identification requirements would be subject to a fine and/or up to five years in prison if one of four conditions are met: (1) the deepfake depicts sexual acts and is produced with intent to humiliate; (2) the person intended to cause violence or interfere with an election; (3) the noncompliance occurred in the course of criminal conduct relating to fraud or identity theft; or (4) the person is a foreign power or agent of a foreign power engaging in unlawful activity, to include interfering in an election.⁵²

It is important to note that the Malicious Deep Fake Prohibition Act expressly exempts "activity protected by the First Amendment;"⁵³ however, all three legislative proposals could nevertheless potentially chill speech and conduct. The bills, if enacted, would likely face First Amendment challenges. The Senate bill might withstand scrutiny given that it is tied to already unlawful criminal or tortious conduct and expressly exempts First Amendment activity. It is difficult to imagine any of the proposals being upheld if a deepfake involved any category of speech that receives strict scrutiny review (e.g., political speech).⁵⁴ For these reasons, their constitutionality is dubious in many situations in which journalists might find themselves. Even still, journalists without access to significant legal support may be afraid to take actions that might contravene the law.

Finally, although the two proposed laws have not yet garnered wide support, that could change in the event of a serious deepfake incident—especially if it occurred during the 2020 election or otherwise impacted national security. A recent Belfer Center report on AI concluded "[t]he bigger and more visible the impacts of AI become (and we argue the impacts are

⁵⁰ Under current law, YouTube and other platforms are unlikely to categorically ban deepfakes but this may change if Communications Decency Act Section 230 immunity is amended. *See infra* Section III.A.2.

⁵¹ Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019, H.R. 3230, 116th Cong. (2019).

⁵² *Id.* § 2.

⁵³ Malicious Deep Fake Prohibition Act of 2018, S. 3085, 115th Cong. § 2(a) (2018).

⁵⁴ *See Chesney & Citron, supra* note 3, at 1803–04.

likely to be increasingly large and obvious over time) the more policymakers will feel justified in making extreme departures from existing policy.”⁵⁵ Put simply, “[r]adical technology change begets radical government policy ideas.”⁵⁶

2. *Amend Section 230 Immunity*

Section 230 of the Communications Decency Act shields “interactive computer services” from liability resulting from content hosted or shared on its service.⁵⁷ At the time of its passage, the prevailing thought was that the possibility of liability would inhibit the growth of the Internet and would disincentivize companies from taking measures to remove obscene content from their services.⁵⁸ Twenty-two years later, as Professors Danielle Citron and Bobby Chesney have explained, “Section 230 has evolved into a kind of super-immunity” with the result that “platforms have no liability-based reason to take down illicit material, and. . . victims have no legal leverage to insist otherwise.”⁵⁹

Though Section 230 immunity is primarily thought to protect platforms such as Facebook and Yelp, it can protect news organizations in important ways. News sites, for example, enjoy immunity for comments posted to online articles by third-parties.⁶⁰ News organizations also enjoy secondary benefits: platforms facilitate free expression, and amending Section 230 immunity may cause platforms to over-moderate content and bar legitimate speech.⁶¹ Taken to the extreme, the litigation risk might lead to platforms removing news articles containing audiovisuals it fears or mistakenly detects are fabricated and potentially unlawful. For these reasons, it is worth briefly exploring why and how Section 230 immunity might be amended to address deepfakes.⁶²

⁵⁵ ALLEN & CHAN, *supra* note 20, at 49.

⁵⁶ *Id.* at 3.

⁵⁷ 47 U.S.C. § 230 (2018).

⁵⁸ Chesney & Citron, *supra* note 3, at 1796–97.

⁵⁹ *Id.* at 1798.

⁶⁰ *Republication in The Internet Age*, REPORTERS COMM. FOR FREEDOM OF THE PRESS, <https://www.rcfp.org/journals/news-media-and-law-summer-2014/republication-internet-age/> (accessed Oct. 30, 2019) [<https://perma.cc/6XK3-EXDE>].

⁶¹ *See infra* notes 71–72 and accompanying text.

⁶² This paper does not take a position on whether Section 230 should be amended; rather, it assumes the possibility and highlights the potential impact reform might have on journalists. For a more detailed analysis on the merits and mechanics of Section 230 immunity, see Danielle Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 *FORDHAM L. REV.* 2 (2017); Citron, *supra* note 28.

Amending Section 230 immunity could be beneficial in addressing deepfakes because it would restore a legal incentive to remove harmful deepfakes (i.e., ones that facilitate illegal or tortious conduct). Senator Mark Warner has highlighted the necessity of this type of legal response noting that platforms are the “least-cost avoiders” in addressing the harm.⁶³ Although strong market incentives exist to remove harmful deepfakes, these can be insufficient.⁶⁴

Several amendments to Section 230 have already been proposed. One proposal would amend Section 230 such that platforms would be immune only if “they could show that their response to unlawful uses of their services was reasonable.”⁶⁵ This “reasonableness” standard is meant to mitigate potential concerns that small providers would be unduly burdened because what is “reasonable” for an emerging platform would be different from what is reasonable for a larger platform like Twitter.⁶⁶ Moreover, the standard would account for emerging counter-deepfake technologies. If counter-deepfake technologies improve and become widely available, it would become increasingly unreasonable for companies not to employ them.⁶⁷ Another proposal would similarly seek to revoke immunity only from “bad” actors with an exemption that is slightly more friendly to platforms.⁶⁸ Under this proposal, a platform would enjoy immunity unless it “knowingly and intentionally [left] up unambiguously unlawful content that clearly creates a serious harm to others.”⁶⁹ Another proposal, even more platform-friendly, would provide for immunity unless the platform “intentionally solicit[ed] or induce[d] illegality or unlawful content.”⁷⁰

If Section 230 was amended in a way that incentivized platforms to act on deepfakes, news organizations that publish defamatory (or otherwise harmful) deepfakes could potentially benefit because the platforms’ efforts to counter fakes would likely mitigate the harm caused by the defamatory deepfake. This benefit, however, would be marginal and only help those organizations that publish harmful deepfakes. In contrast, potential effects to

⁶³ Mark R. Warner, Potential Policy Proposals for Regulation of Social Media and Technology Firms (July 30, 2018), https://www.warner.senate.gov/public/_cache/files/d/3/d32c2f17-cc76-4e11-8aa9-897eb3c90d16/65A7C5D983F899DAAE5AA21F57BAD944.social-media-regulation-proposals.pdf (draft white paper) [<https://perma.cc/UK2F-RQPL>].

⁶⁴ For example, one victim of a deepfake pornography video attempted to get the video removed from a website and became the victim of a sextortion attempt. Citron, *supra* note 28 at 40.

⁶⁵ Citron & Wittes, *supra* note 62, at 419.

⁶⁶ *Id.*

⁶⁷ *See id.*

⁶⁸ Citron, *supra* note 28, at 63.

⁶⁹ *Id.*

⁷⁰ *Id.*

free speech and online discourse would affect all news organizations. Legal liability can lead to overzealous removal efforts because “a platform’s easiest and cheapest course is to take accusations at face value” and remove the content.⁷¹ The Fourth Circuit has articulated this concern: “Liability upon notice has a chilling effect on the freedom of Internet speech” because of platforms’ “natural incentive simply to remove messages upon notification.”⁷²

It is possible that criminalization of deepfakes and Section 230 immunity reform could occur simultaneously. This would be the worst-case scenario for news organizations because legal liability for harmful deepfakes would likely lead many platforms to ban at least some forms of deepfakes. As previously discussed, the Malicious Deep Fake Prohibition Act combined with more restrictive terms of service could increase liability for news organizations.⁷³

Finally, it is important to note that the government has many options other than the two explored here. The government could restrict access to deepfake-related technology via export controls,⁷⁴ mandate the use of digital signatures or AI-watermarks,⁷⁵ or invest in education to increase the next generation’s ability to analyze the credibility of Internet content.⁷⁶ These options, however, are less directly applicable to the press and thus were not explored in this analysis.

B. Loss in Public Trust

Public trust in news media is already on the decline. According to a Gallup/Knight Institute survey, sixty-nine percent of American adults say their

⁷¹ DAPHNE KELLER, INTERNET PLATFORMS: OBSERVATIONS ON SPEECH, DANGER, AND MONEY 5 (2018), https://www.hoover.org/sites/default/files/research/docs/keller_webreadypdf_final.pdf [<https://perma.cc/N9TN-VS5R>]. One study that looked at removal in response to copyright claims found that most smaller and lower-profile platforms removed content even when uncertain about the strength of the copyright claim. *Id.*

⁷² *Zeran v. AOL*, 129 F.3d 327, 333 (4th Cir. 1997).

⁷³ See *supra* Section III.A.1.

⁷⁴ ALLEN & CHAN, *supra* note at 20, at 29–30.

⁷⁵ See *infra* Section IV.A.2.

⁷⁶ Recent studies have shown that adults and minors are generally unable to visually spot fakes. One researcher explained that “[m]ore often than not, people think that the real images are fake and that things that are fake are real. And their confidence is very high. So people are both ignorant and confident, which is the worst combination.” Tiffanie Wen, *The Hidden Signs That Can Reveal a Fake Photo*, BBC (June 30, 2017), <http://www.bbc.com/future/story/20170629-the-hidden-signs-that-can-reveal-if-a-photo-is-fake> [<https://perma.cc/8KS2-CC8M>].

trust in news media has declined over the past decade.⁷⁷ Perhaps the most important finding relevant to deepfakes is that two-thirds of respondents indicated inaccuracy was a reason for their mistrust.⁷⁸ Any loss of trust should be concerning to the news industry because consumer trust is a business necessity.⁷⁹ The publication of deepfakes—or even the perception that deepfakes are being published—is likely to exacerbate concerns about the inaccuracy of news. Two characteristics of today’s climate will help drive this development.

The first characteristic is the increasing tendency to group news organizations together and reinforce monolithic stereotypes (think “mainstream media,” “conservative media,” and “fake news.”) This stereotyping is particularly concerning in the advent of deepfakes because one news organization’s blunder may have wider confidence ramifications. NBC News Political Director Chuck Todd recently articulated this effect: “When I make a mistake, it’s going to have an impact on Jake Tapper at CNN, it’s going to have an impact on Chris Wallace at Fox . . . any error in the ‘mainstream media,’ we all pay the price.”⁸⁰

A recent incident involving doctored video illustrates this effect. In January 2019, a Seattle TV news station broadcasted a doctored video of President Donald Trump’s Oval Office address regarding the United States’ southern border.⁸¹ In the video, the coloring was saturated to make the president appear orange, his head was enlarged, and his facial expressions

⁷⁷ Knight Foundation, *Indicators of News Media Trust*, KNIGHT FOUND. (Sept. 11, 2018), <https://www.knightfoundation.org/reports/indicators-of-news-media-trust> [<https://perma.cc/KDD9-YX34>].

⁷⁸ The open-ended question was: “Thinking now about some of the news media organizations you DO NOT trust, what are some of the reasons why you DO NOT trust those news organizations?” While “two-thirds mentioned accuracy-related reasons at least once,” three-quarters of respondents mentioned bias. Interestingly, “Republicans, Democrats, and independents are about equally likely to bring up inaccuracy as a reason they distrust certain news organizations.” *Id.*

⁷⁹ See *A New Understanding: What Makes People Trust and Rely on News*, AM. PRESS INST. (Apr. 17, 2016), <https://www.americanpressinstitute.org/publications/reports/survey-research/trust-news/single-page/> [<https://perma.cc/N7FH-H5CJ>]; *Guidelines on Integrity*, N.Y. TIMES (Sept. 25, 2008), <https://www.nytimes.com/editorial-standards/guidelines-on-integrity.html> [<https://perma.cc/9E9Z-HAWQ>].

⁸⁰ Jake Sheridan, *Chuck Todd on Journalism’s Long Road Back to Win Public Trust*, DUKE TODAY (Jan. 15, 2019), <https://today.duke.edu/2019/01/chuck-todd-journalisms-long-road-back-win-public-trust> [<https://perma.cc/R6KK-G4CP>].

⁸¹ Kyle Swenson, *A Seattle TV Station Aired Doctored Footage Of Trump’s Oval Office Speech. The Employee Has Been Fired*, WASH. POST (Jan. 11, 2019), https://www.washingtonpost.com/nation/2019/01/11/seattle-tv-station-aired-doctored-footage-trumps-oval-office-speech-employee-has-been-fired/?utm_term=.573f89517b1a [<https://perma.cc/SAV3-AH9Q>].

were altered such that he was sticking out his tongue.⁸² Viewers took to social media to speculate on whether the video had been doctored until the station confirmed it had aired a doctored version and that the single editor responsible for doing so had been fired.⁸³ Despite the station's quick response, many Internet users were outraged and, notably, some of their comments were quick to generalize, blaming "the media" writ large: "they wonder why we all think of the Media and News as fake!";⁸⁴ "the news media is out of control!"; "the news media are despicably corrupt."⁸⁵

The second characteristic relates to the spread of falsities on social media. A recent study assessing false news on Twitter between 2006 to 2017 found that "falsehood diffused significantly farther, faster, deeper, and more broadly than truth in all categories of information."⁸⁶ The study's findings further underscore the heightened responsibility journalists have in the digital age where, though all news spreads quickly, false news spreads even faster. Given this reality, if a deepfake is published and then retracted, it may be more difficult for news of the correction to reach the original readers. Additionally, attempts to undermine journalists would spread rapidly, as journalist Rana Ayyub unfortunately experienced.⁸⁷

IV. SOLUTIONS FOR THE PRESS TO PURSUE

Unfortunately, there is not a single solution to address deepfakes and this paper does not, and could not, proclaim to have the ultimate preparedness plan. While this Part provides a few technology-based solutions, the most

⁸² *Id.*

⁸³ *See id.*; *Was video of President Trump's Tuesday address doctored?*, REDDIT (2019), https://www.reddit.com/r/The_Donald/comments/aejp7a/was_video_of_president_trumps_tuesday_address/ [perma.cc link unavailable].

⁸⁴ @RSutter, Twitter (Jan. 9, 2019, 11:55 PM), <https://twitter.com/RSutter/status/1083225828651978752> [https://perma.cc/Z7BE-YDY6].

⁸⁵ Todd Herman, *Q13 FOX Editor Fired Over Doctored Trump Address Video*, 770 KKTH (Jan. 10, 2019 at 1:11 PM), <https://mynorthwest.com/1237906/was-video-of-president-trumps-tuesday-address-doctored/?show=comments#comment-4278103291> [https://perma.cc/SQ8J-TSZ9].

⁸⁶ Soroush Vosoughi et al., *The Spread Of True And False News Online*, 359 SCI. 1146 (2018), <http://science.sciencemag.org/content/359/6380/1146> [https://perma.cc/65VG-TW5C]; Till Daldrup & Francesco Marconi, *How The Wall Street Journal Is Preparing Its Journalists To Detect Deepfakes*, NIEMAN LAB (Nov. 15, 2018, 8:48 AM), <http://www.niemanlab.org/2018/11/how-the-wall-street-journal-is-preparing-its-journalists-to-detect-deepfakes/> ("False stories were 70 percent more likely to be retweeted than the truth and reached 1,500 people six times more quickly than accurate articles.") [https://perma.cc/G494-ENPC].

⁸⁷ *See Ayyub, supra* note 28.

crucial step journalists can take right now is engaging in dialogue, raising awareness, and collaborating within the industry.

A. Technology-based Solutions

There are three promising technological solutions that could aid verification: digital authentication software, digital signatures, and watermarking. Each of these will be explored in turn.

1. *Digital Authentication*

Digital forensics are becoming an increasingly accessible option journalists can use to verify audiovisuals. Current services offer a variety of different tools that use AI, human verifiers, or both. These tools take the form of web-based tools or on-demand verification where a user uploads a specific audiovisual to verify. AI Foundation's Reality Defender software, for example, will soon release a free Google Chrome extension that will scan media that the user encounters while browsing on Chrome and flag media that is manipulated or generated using AI or other means.⁸⁸ Another company, Truepic, offers both on-demand and software solutions.⁸⁹ Users can drag and drop an image into the Truepic Insight web panel where results on its authenticity will appear instantly.⁹⁰ Alternatively, the API version automates the process allowing thousands of images to be quickly verified.⁹¹ In addition to using high-tech verification techniques to flag AI-generated or manipulated content, Truepic's technology also automates many current contextual verification techniques.⁹² For example, it can identify whether an image is being repurposed by flagging whether it is an original image (which a journalist might do via Reverse Google Image search), examine metadata to see if it has been manipulated, and flag location spoofing. Finally, while the

⁸⁸ *Defend Reality*, AI FOUND., <http://www.aifoundation.com/responsibility> (accessed Oct. 30, 2019) [<https://perma.cc/6DZF-RGWF>].

⁸⁹ *Our Technology*, TRUEPIC, <https://truepic.com/technology/> (accessed Oct. 30, 2019) [<https://perma.cc/PFL9-QEJG>].

⁹⁰ *Id.*

⁹¹ *Id.*

⁹² *Id.*

previous solutions have all been services journalists can personally use, the digital verification process can also be outsourced to other entities.⁹³

One challenge to using digital verification is ensuring that flagged media has been manipulated in a significant way.⁹⁴ Merely opening a photo in Photoshop can alter metadata in a way that software can detect.⁹⁵ For this reason, journalists should not automatically discount an audiovisual flagged by software and remember that the technology is an aid, not a final arbiter.

2. Digital Signatures

Another promising technology solution is digital signatures. Digital signatures “enable a party to sign a digital object in a way that proves he or she was the one who signed it.”⁹⁶ A digitally-signed audiovisual would allow viewers to confirm the audiovisual is authentic, was signed by a particular device/person, and has not been modified.⁹⁷ Digital signatures can be embedded on devices or in an app. Cannon and Nikon have both implemented the idea in several versions of their cameras, but no smartphone companies are known to have adopted it yet.⁹⁸ Serelay, a UK-based company, has an iOS and Android-compatible app that creates digital signatures at the moment of capture when a photo or video is taken using the app. Content taken via the Serelay app can then later be verified either using an on-demand platform or

⁹³ For example, Storyful, a “social media intelligence and news agency,” offers user-generated content verification services and has partnered with news organizations such as the *Wall Street Journal* to verify images. About, STORYFUL, <https://storyful.com/about/> (accessed Nov. 30, 2019) [<https://perma.cc/5Q2E-P6VB>]; Eamonn Kennedy & Keelan Byrne, *Introducing A New and Improved Storyful Wire*, STORYFUL (Sept. 12, 2018), <https://storyful.com/blog/introducing-a-new-and-improved-storyful-newswire/> [<https://perma.cc/R2VR-BTM8>]; see Daldrup & Marconi, *supra* note 86.

⁹⁴ *Defend Reality*, *supra* note 88.

⁹⁵ Martin Harran et al., *A Method For Verifying Integrity & Authenticating Digital Media*, 14 SCI. DIRECT, 145, 150 (2018), <https://www.sciencedirect.com/science/article/pii/S2210832717300753> [<https://perma.cc/CK8S-ASLT>].

⁹⁶ Herb Lin, *The Danger of Deep Fakes: Responding to Bobby Chesney and Danielle Citron*, LAWFARE BLOG (Feb. 27, 2018, 7:00 AM), <https://www.lawfareblog.com/danger-deep-fakes-responding-bobby-chesney-and-danielle-citron> [<https://perma.cc/KWL7-9B4D>].

⁹⁷ *Id.* A digital signature is distinct from metadata because a signature would remain intact unless the audiovisual was altered whereas metadata can change merely from being viewed in certain applications. See Harran, *supra* note 95.

⁹⁸ *Originality Verification Function | OSK-E3*, CANON, <http://web.canon.jp/imaging/osk/osk-e3/verifies/index.html> (accessed Oct. 30, 2019) [<https://perma.cc/6HU4-U4JJ>]; *Image Authentication Software*, NIKON, https://imaging.nikon.com/lineup/software/img_auth/index.htm (accessed Oct. 30, 2019) [<https://perma.cc/442C-3F7Y>]; Lin, *supra* note 96.

an API service similar to Truepic's. If content is flagged as manipulated, the area of manipulation will be highlighted.⁹⁹

3. *AI Watermarks*

There are many responsible uses for AI-generated audiovisuals (educational use, artistic use, etc.), but, given their increasing realism, they can go undetected which could undermine their intended use. AI watermarks can be used to responsibly mark AI-generated audiovisuals so that viewers are aware the content has been computer-generated. The AI Foundation has taken this approach and “partner[ed] with content creators to establish and use an ‘Honest AI watermark’ to clearly identify and call out AI-generated text images, audio, and video.”¹⁰⁰ While malicious actors are unlikely to use AI watermarks, their increased use would, at a minimum, make it easier for journalists and others to quickly establish marked audiovisuals are AI-generated.

B. Education, Dialogue, and Collaboration

While the tools outlined above show promise in aiding digital verification for journalists and Internet users, they cannot supplant current methods. Instead, the tools should be used in addition to contextual verification techniques to most effectively identify deepfakes. As put by *Wall Street Journal* (WSJ) Chief Technology Officer Rajiv Pant: “The way to combat deepfakes is to augment humans with artificial intelligence tools.” The most appropriate AI tool and the optimal level of AI assistance will vary according to the audiovisual at issue, how it is being used, and the person or organization that seeks to use it. For this reason, it is imperative that news organizations and journalists begin to examine how deepfakes might affect their organization or products and make individualized assessments about how to best prepare for them. Too few organizations have begun this process.

The WSJ is an outlier and established the WSJ Media Forensics Committee, an “internal deepfakes task force.”¹⁰¹ The Committee is studying current and emerging deepfake technology, evaluating verification practices, and educating its newsroom about deepfakes via training seminars and newsroom guides.¹⁰² Reuters has also publicly acknowledged that it is

⁹⁹ For a demonstration of what browsing Twitter would look like while using Serelay, see *Verified Twitter Feed*, SERELAY, <http://www.verifiedtwitterfeed.com> [<https://perma.cc/X7JN-TZKV>]. For a demonstration of the on-demand service, see *Demo*, SERELAY, <https://www.serelay.com/our-products/media/> [<https://perma.cc/9WCB-4Y6J>].

¹⁰⁰ *Defend Reality*, *supra* note 88.

¹⁰¹ Daldrup & Marconi, *supra* note 86.

¹⁰² *Id.*

preparing for deepfakes.¹⁰³ Hazel Baker, head of user-generated content news-gathering for Reuters, commented that “[t]here’s not a slew of deepfakes on my desk, but I don’t want to wait till there are.”¹⁰⁴ Reuters has doubled the number of staff verifying video content (from six to twelve) and even worked with a specialist production company to create a deepfake video to test its user-generated content team.¹⁰⁵ Those who were aware the video was manipulated in some manner identified the inconsistencies, while those who were not aware “noticed something was off in the audio but struggled to define it.”¹⁰⁶

News organizations should follow the lead of the WSJ and Reuters and begin thinking about this issue and exploring solutions. If news organizations are planning internally, they should consider publicly identifying their findings and processes to help others prepare because, if the deepfake threat materializes as this author and many others expect, self-preparedness will only go so far. A weakest-link mentality is needed: If the collective news industry is not prepared, each news entity will be affected by the resulting loss in public confidence to the industry, by government action, or both. One scholar has even suggested that a well-prepared news industry may benefit from the advent of deepfakes. He argues that “[d]ire as the case may be, it could offer a great comeback opportunity for mainstream media. As the public learns that it can no longer trust what it sees online, few intermediaries are better placed to function as trusted validators and assessors of mediated reality than professionally trained journalists with access to advanced forensics tools.”¹⁰⁷ Benefiting from deepfakes, however, is a lofty goal for an industry that has only just begun to acknowledge them. The news industry currently has a window of opportunity while deepfake technology is relatively nascent and imperfect, but it will quickly close as the technology improves and its wielders become more creative.

V. CONCLUSION

The government measures previously outlined are not inevitable. They, or more extreme measures, are only likely to be implemented if a legal solution seems necessary. Moreover, though the public’s perception of

¹⁰³ See Lucinda Southern, *How Reuters Is Training Reporters To Spot “Deepfakes,”* DIGIDAY (Mar. 26, 2019), <https://digiday.com/media/reuters-created-a-deepfake-video-to-train-its-journalists-against-fake-news/> [<https://perma.cc/C6EF-83G5>].

¹⁰⁴ *Id.*

¹⁰⁵ *Id.*

¹⁰⁶ *Id.*

¹⁰⁷ Nicholas Diakopoulos, *Reporting In A Machine Reality: Deepfakes, Misinformation, and What Journalists Can Do About Them*, TOW CENTER FOR DIGITAL JOURNALISM (May 15, 2018), https://www.cjr.org/tow_center/reporting-machine-reality-deepfakes-diakopoulos-journalism.php [<https://perma.cc/SQ45-AYLL>].

audiovisuals, including those in the news, is likely to change, the news industry can mitigate this by showing it is thinking about these issues and by being prepared when more and better deepfakes circulate. News organizations should prepare by educating their journalists about current deepfake capabilities and assessing which technology-based solutions could best supplement their digital verification practices. We are not yet at an inflection point, but it will not serve the news industry well to wait until we are.

GETTING ON BOARD WITH ROBOTS: HOW THE BUSINESS JUDGMENT RULE SHOULD APPLY TO ARTIFICIAL INTELLIGENCE DEVICES SERVING AS MEMBERS OF A CORPORATE BOARD

Thomas Belcastro*

CITE AS: 4 GEO. L. TECH. REV. 263 (2019)

TABLE OF CONTENTS

I. INTRODUCTION	264
II. THE BUSINESS JUDGMENT RULE.....	265
A. Fiduciary Duties and the Board of Directors	265
B. Policy Underlying the Business Judgment Rule.....	266
C. The Business Judgment Rule in United States Law	267
D. Summary	269
III. ARTIFICIAL INTELLIGENCE DEVICES.....	270
A. Advancements in Technology.....	270
B. Artificial Intelligence Devices in Industry and Businesses Today .	271
C. How Artificial Intelligence Devices Make Business Judgments	273
D. Summary	274
IV. THE FUTURE OF THE BUSINESS JUDGMENT RULE	275
A. Legal Application of the Business Judgment Rule to Artificial Intelligence.....	275
B. Policy Considerations	276
V. CONCLUSION.....	278

* Cornell Law School, J.D. 2019. I would like to extend my deepest gratitude to my wife, Linnaea, for her unfailing love and support for my educational and professional pursuits, and my children, Harley, Ethel and Luna for their love and patience as I worked on this note. I would also like to thank the editors of the Georgetown Law Technology Review for their valuable feedback and contributions. Thanks also to Professor Ed Walters who encouraged me to submit this piece to the GLTR Student Writing Competition.

I. INTRODUCTION

In 1742, the Lord Chancellor of England noted that:

[Directors] are most properly agents to those who employ them in this trust, and who empower them to direct and superintend the affairs of the corporation. In this respect they may be guilty of acts of commission or omission, of malfeasance or nonfeasance. Now where acts are executed within their authority, . . . though attended with bad consequences, it will be very difficult to determine that these are breaches of trust. For it is by no means just in a judge, after bad consequences have arisen from such executions of their power, to say that they foresaw at the time what must necessarily happen; and therefore, were guilty of a breach of trust.¹

This recognition that directors should not be held personally liable for honest business decisions that result in negative consequences has continued to be a bedrock of corporate law.² In modern times, this and other policy goals have been secured by the implementation of the business judgment rule.³ Under the business judgment rule, board decisions are not second guessed by courts so long as they are made on an informed basis, in good faith, and in the honest belief that the actions taken are in the best interest of the company.⁴ This approach has served its purpose well up to now, but one recently introduced wrinkle has yet to be addressed.

In 2014, Deep Knowledge Ventures, a Hong Kong based venture capital fund focused on life sciences, became the first company to appoint an artificial intelligence entity to its board of directors.⁵ The computer algorithm, called Vital, has a vote and is treated “as a member of [the] board with observer

¹ *Charitable Corp. v. Sutton*, 26 Eng. Rep. 642 (1742) (citations omitted).

² Bernard S. Sharfman, *The Importance of the Business Judgement Rule*, 14 N.Y.U. J. L. & Bus. 27, 33–34 (2017).

³ *Id.* at 33–37.

⁴ *Aronson v. Lewis*, 473 A.2d 805 (Del. 1984).

⁵ *Algorithm Appointed Board Director*, BBC NEWS (May 16, 2014), <https://www.bbc.com/news/technology> [<https://perma.cc/V42X-9FHN>]; Will Pugh, *Why Not Appoint an Algorithm to Your Corporate Board*, SLATE (March 24, 2019, 8:00 AM), <https://slate.com/technology/2019/03/artificial-intelligence-corporate-board-algorithm.html> [<https://perma.cc/MY3U-S9LU>]; Ellie Zolfaghari, *Would You Take Orders from a ROBOT? An Artificial Intelligence becomes the World's First Company Director*, DAILYMAIL (May 19, 2014, 5:50 PM), <https://www.dailymail.co.uk/sciencetech/article-2632920/Would-orders-ROBOT-Artificial-intelligence-world-s-company-director-Japan.html> [<https://perma.cc/SWW4-VR6Z>].

status.”⁶ Importantly, the human members of the board “agreed that [they] would not make positive investment decisions without corroboration by Vital.”⁷ This is not only a new frontier for corporate governance, but also creates a unique situation for the application of the traditional business judgment rule, which until now has only had to deal with human directors.

This Essay proceeds in three parts. Part II introduces the business judgment rule and its underlying policy. Part III discusses the current advancements in artificial intelligence, machine learning, algorithms, and robotics (hereinafter collectively “artificial intelligence devices”) that have implications for understanding the role of the business judgment rule in the future. Part IV argues that, though the business judgment rule could be as easily applied to artificial intelligence devices as their human director counterparts, the policy needs underlying the rule are not applicable to such devices and the protections of the rule should not apply to them.

II. THE BUSINESS JUDGMENT RULE

A. Fiduciary Duties and the Board of Directors

Ultimately, the business judgment rule is the result of corporate law’s attempt to deal with an agency problem. Agency problems arise when the interests of a principal depend on actions taken by their agent.⁸ One of the ways the law addresses the multitude of problems raised by principal-agent relationships is the imposition of fiduciary duties on agents (called “fiduciaries”). The two primary duties of fiduciaries are the duty of loyalty and the duty of care.⁹ The duty of loyalty requires agents to act for the interests of the principal rather than for their own interests.¹⁰ More importantly for the purposes of the analysis in this paper, the duty of care requires agents making

⁶ Nicky Burrige, *Artificial Intelligence Gets a Seat in the Boardroom*, NIKKEI ASIAN REV. (May 10, 2017, 10:52 PM), <https://asia.nikkei.com/Business/Companies/Artificial-intelligence-gets-a-seat-in-the-boardroom> [<https://perma.cc/3ARE-FFGP>]. Vital is an acronym for “Validating Investment Tool for Advancing Life Sciences.” See Press Release, Deep Knowledge Ventures, Deep Knowledge Venture’s Appoints Intelligent Investment Analysis Software VITAL as Board Member (May 13, 2014), <http://www.prweb.com/releases/2014/05/prweb11847458.htm> [<https://perma.cc/LDS3-LH7L>].

⁷ Burrige, *supra* note 6.

⁸ See John Armour, Henry Hansmann & Reinier Kraakman, *Agency Problems, Legal Strategies and Enforcement 2–5* (John M. Olin Ctr. for L., Econ., & Bus., Discussion Paper No. 644, 2009), http://www.law.harvard.edu/programs/olin_center/papers/pdf/Kraakman_644.pdf [<https://perma.cc/6ZA8-2UET>].

⁹ *Fiduciary Duty*, LEGAL INFO. INST., https://www.law.cornell.edu/wex/fiduciary_duty (accessed Oct. 30, 2019) [<https://perma.cc/DW4Q-NXB3>].

¹⁰ *Id.*

decisions in their capacity as fiduciaries to act in the same manner that a reasonably prudent person would under the circumstances.¹¹ Failure to abide by either of these duties subjects a fiduciary to personal liability for damages to the principal.¹² After all, the principals supply the capital, and the agents, if they are not liable for losing that capital, only stand to benefit from excessive risk-taking. In the corporate setting, shareholders are principals and directors serve as their agents.¹³

B. Policy Underlying the Business Judgment Rule

Generally speaking, directors are elected by shareholders to run the business and increase its value.¹⁴ Indeed, shareholder wealth maximization is the default legal obligation of a board of directors.¹⁵ However, this legal obligation often presents a conundrum for the board. On the one hand, the board is required not only to *increase* but to *maximize* the value of the company for the shareholders. On the other hand, the board owes a fiduciary duty of care to shareholders to act in the same manner as a reasonably prudent person under the circumstances or be held personally liable for their decisions.¹⁶ Often, to maximize shareholder value, the board must take certain risks such as selling the company or releasing a new product. When those risks result in large returns, the directors are often lauded but, when the results are negative, the shareholders can attempt to hold the board personally liable for damages.¹⁷ Thus, directors face liability whether they fail to maximize shareholder value by inaction or action that proves unfruitful.

Due to this conundrum, the individuals most qualified to serve as directors, without any sort of protection from personal liability, would likely refuse to take a seat on the board.¹⁸ Additionally, courts would find themselves

¹¹ See *Duty of Care*, LEGAL INFO. INST., https://www.law.cornell.edu/wex/duty_of_care (accessed Oct. 30, 2019) [<https://perma.cc/SZ9Q-RJDD>].

¹² See *Fiduciary Duty*, *supra* note 9.

¹³ See Dalia T. Mitchell, *Status Bound: The Twentieth Century Evolution of Directors' Liability*, 5 N.Y.U. J. L. & BUS. 63, 65 (2009). See also Chad Langager, *Who is Responsible for Shareholders Interests?*, INVESTOPEDIA (Mar. 8, 2018), <https://www.investopedia.com/ask/answers/05/shareholderinterest.asp> [<https://perma.cc/D6XR-DBUT>].

¹⁴ See Langager, *supra* note 13.

¹⁵ See Sharfman, *supra* note 2, at 31, 56–68.

¹⁶ See Bernard S. Sharfman, *The Enduring Legacy of Smith v. Van Gorkom*, 33 DEL. J. CORP. L. 287, 288–89 (2008).

¹⁷ See Lindsay C. Llewellyn, *Breaking Down the Business Judgment Rule*, 14 COM. & BUS. LITIG. 16 (2013) (“In bringing shareholder derivative suits, shareholders seek to impose liability on corporate directors for failing to carry out their corporate duties in accordance with this standard of care”).

¹⁸ See Sharfman, *supra* note 16, at 301–02.

mired in litigation for damages resulting from either the board's inaction or action. This would result in requiring judges to essentially assume the role of board members to determine *ex post* what the correct course of action would have been for what should have been a routine business decision. That situation would not only be cumbersome for the courts but would also suffer heavily from hindsight bias.¹⁹ In light of these concerns, a shield for business decisions had to be developed in order to protect directors from personal liability for merely doing their jobs.

C. The Business Judgment Rule in United States Law

One of the most important developments in United States corporate law is the business judgment rule.²⁰ The business judgment rule is a standard of review that shields directors from personal liability when they take actions that negatively affect corporations, resulting in shareholder lawsuits alleging violations of the duty of care. Without the business judgment rule, courts are obliged to perform a "fairness review" ("entire fairness" in Delaware).²¹ A fairness review is a court's "most onerous" standard of review.²² Under this standard, the defendant bears the burden of proving that the transaction satisfied the requirements of "fair dealing and fair price."²³ The clearest articulation of the business judgment rule is found in *Aronson v. Lewis*, a Delaware Supreme Court case, which states that:

[i]t is a presumption that in making a business decision the directors of a corporation acted on an informed basis, in good faith, and in the honest belief that the action taken was in the best interest of the company. Absent an abuse of discretion, that judgment will be respected by the courts. The burden is on the party challenging the decision to establish facts rebutting the presumption.²⁴

The court went on to make "informed basis" a requirement by stating:

to invoke the rule's protection directors have a duty to inform themselves, prior to making a business decision, of all material information reasonably available to them. Having become so

¹⁹ See Hal R. Arkes & Cindy A. Schipani, *Medical Malpractice v. The Business Judgement Rule: Differences in Hindsight Bias*, 73 OR. L. REV. 587, 588–89 (1994).

²⁰ See Sharfman, *supra* note 2, at 28–29.

²¹ *See id.*

²² *Id.* at 40.

²³ *Id.*

²⁴ *Aronson v. Lewis*, 473 A.2d 805, 812 (Del. 1984).

informed, they must then act with requisite care in the discharge of their duties. While the Delaware cases use a variety of terms to describe the applicable standard of care, our analysis satisfies us that under the business judgment rule director liability is predicated upon concepts of gross negligence.²⁵

This requirement opened the door for the most widely recognized business judgment rule case, and arguably the most significant case to come out of the Delaware Supreme Court, *Smith v. Van Gorkom*.²⁶

In *Van Gorkom*, the board of directors of TransUnion Co. approved the sale of the company after a single two-hour meeting.²⁷ The transaction resulted in lost value and, while many believed that the court would still apply the business judgment rule, the court surprised many observers by holding that all of the directors were personally liable for the resulting monetary damages.²⁸ Reaffirming gross negligence as the standard by which to determine whether a business decision was “informed” as required by *Aronson*, the court held that the directors “were grossly negligent in approving the ‘sale’ of the Company upon two hours’ consideration, without prior notice, and without the exigency of a crisis or emergency.”²⁹ *Van Gorkom* was such a deviation from the common approach to the business judgment rule that it prompted the Delaware legislature to enact new laws to effectively let businesses opt out of its holding.³⁰ *Van Gorkom*’s progeny, however, reveals that the Delaware Supreme Court intends to remain an outlier in regards to business judgment rule jurisprudence.³¹

While the approach may differ slightly from state to state, the idea that courts will not second guess business judgments is repeated throughout state court opinions. Indeed, what makes Delaware an outlier is that other states are more likely to afford greater deference to boards. For example, in 1968 the Illinois Appellate Court found that a shareholder had no cause of action where the board of directors of the Chicago Cubs decided against putting lights in the stadium at Wrigley field to increase attendance and revenue by scheduling

²⁵ *Id.*

²⁶ *Smith v. Van Gorkom*, 488 A.2d 858 (Del. 1985).

²⁷ *Id.* at 869.

²⁸ See Henry N. Butler, *Smith v. Van Gorkom, Jurisdictional Competition, and the Role of Random Mutations in the Evolution of Corporate Law*, 45 WASHBURN L.J. 267, 282 (2006) (“*Van Gorkom* was a major and seemingly random change in the corporate law of fiduciary duties. It certainly took corporate law observers and commentators by surprise”).

²⁹ *Van Gorkom*, 488 A.2d at 874.

³⁰ DEL. CODE ANN. TIT. 8, § 102(b)(7) (2019).

³¹ See, e.g., *Emerald Partners v. Berlin*, 787 A.2d 85 (Del. 2001); *Cede & Co. v. Technicolor*, 634 A.2d 345 (Del. 1993).

games at night—despite the fact that nearly every other Major League Baseball team had been successfully adopting the approach since 1935.³² The court held that:

[d]irectors are elected for their business capabilities and judgment and the courts cannot require them to forgo their judgment because of the decisions of directors of other companies. Courts may not decide these questions in the absence of a clear showing of dereliction of duty. . . and a mere failure to ‘follow the crowd’ is not such a dereliction.³³

Similarly, in a case where the board of directors was found to have breached its duty to minority shareholders, the Ohio Supreme Court took the time to reaffirm the idea that “[t]he [business judgment] rule is a rebuttable presumption that directors are better equipped than the courts to make business judgments . . . A party challenging a board of directors’ decision bears the burden of rebutting the presumption that the decision was a proper exercise of the business judgment of the board.”³⁴ Some states, like California, have even gone so far as to codify the common law business judgment rule in statutory law.³⁵

D. Summary

The fiduciary duties of directors can conflict with their need to maximize shareholder wealth. To mitigate this conundrum, the business judgment rule creates a rebuttable presumption of good faith in favor of the board, shielding its members from personal liability. This not only encourages directors to take board positions but also to apply the appropriate amount of risk-taking necessary to maximize shareholder wealth. Without the intervention of the business judgment rule, a much more onerous fairness standard would apply. When the business judgment rule does apply, the burden is on plaintiffs to show that directors acted in gross negligence, bad faith, or had a conflict of interest to overcome the presumption.³⁶ While Delaware courts are more likely to find gross negligence than other states, the idea underlying that approach still applies—directors must have some

³² See *Shlensky v. Wrigley*, 237 N.E.2d 776, 781 (Ill. App. Ct. 1968).

³³ *Id.*

³⁴ *Gries Sports Enter., Inc. v. Cleveland Browns Football Co.*, 496 N.E.2d 959, 963 (Ohio 1986).

³⁵ CAL. CORP. CODE § 7231 (West 2019).

³⁶ See, e.g., *Aronson v. Lewis*, 473 A.2d 805 (Del. 1984); *Smith v. Van Gorkom*, 488 A.2d 858 (Del. 1985).

informed basis for their decision in order to receive the protection of the business judgment rule.

III. ARTIFICIAL INTELLIGENCE DEVICES

A. Advancements in Technology

One need only look to the Industrial Revolution or the invention of the internal combustion engine to see what effect new technologies can have on industries, businesses, and markets. Indeed, many have made the claim that the current unprecedented advancements in artificial intelligence devices constitute a sort of “Fourth Industrial Revolution” or “Industry 4.0.”³⁷ As Ryan Calo³⁸ aptly pointed out in 2015, “[i]t is becoming increasingly obvious that advances in robotics will come to characterize the next several decades” and that “[t]he same government and hobbyists that developed the Internet, and the handful of private companies that have come to characterize it, have begun a significant shift toward robotics and artificial intelligence.”³⁹ This looming change has been expedited by the availability of low-cost robotics-adjacent technologies, like Microsoft’s Kinect, and behemoth companies, like Google parent company Alphabet, which are spending hundreds of billions of dollars racing to take a share of the market for these technologies.⁴⁰

In describing which attributes of these developing technologies will be most relevant to the law, Calo suggests that “[e]mbodiment, emergence, and social valence—alone, and especially in combination—turn out to be relevant to an extraordinarily wide variety of legal contexts. . . .”⁴¹ He goes on to describe embodiment as relating to the nature of artificial intelligence devices occupying physical space in the world,⁴² emergence as the ability to learn and act in a way that is unpredictable,⁴³ and social valence as how people perceive

³⁷ See, e.g., Bernard Marr, *The 4th Industrial Revolution Is Here-Are You Ready?*, FORBES (Aug. 13, 2018, 12:26 AM), <https://www.forbes.com/sites/bernardmarr/2018/08/13/the-4th-industrial-revolution-is-here-are-you-ready/> [<https://perma.cc/UM2H-6DDT>].

³⁸ Ryan Calo is an associate professor at the University of Washington School of Law and co-chair of the American Bar Association Committee on Robotics and Artificial Intelligence. He is recognized nationally as a leading expert in the cross section of emerging technology in law. See e.g. Ctr. Internet & Society, *Ryan Calo*, CYBERLAW.STANFORD.EDU, <http://cyberlaw.stanford.edu/about/people/ryan-calo> (accessed Dec. 21, 2019) [<https://perma.cc/3DWY-LL3W>].

³⁹ Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513, 526, 562 (2015).

⁴⁰ See *id.*; see also Cade Metz, *Good News: A.I. Is Getting Cheaper. That’s Also Bad News*, N.Y. TIMES, (Feb. 21, 2018), <https://www.nytimes.com/2018/02/20/technology/artificial-intelligence-risks.html> [<http://perma.cc/6YZU-7D77>].

⁴¹ Calo, *supra* note 39, at 532.

⁴² See *id.* at 532–37.

⁴³ See *id.* at 538–45.

AI devices—noting that the advancements “*feel* different to us, more like living agents.”⁴⁴

While embodiment may not be as applicable to robots or artificial intelligence devices acting as members on a board as it is to workers in a factory, the emergence and social valence aspects of these technologies will likely play a role in the future application of the business judgment rule to artificial intelligence devices. Emergent behavior, originally implemented in military applications,⁴⁵ has already been observed in the financial sector and even caused a minor crash, now known as the “Flash Crash of 2010,”⁴⁶ resulting in the disappearance of one trillion dollars in less than thirty minutes.⁴⁷ It is easy to see how such a mistake emanating from an artificial intelligence director would be grounds for a shareholder derivative suit. It also shows why social valence is important. After all, in order for the shareholders to bring a suit against directors, they would have to *feel* that the directors are agents. Thus, emergence and social valence will play a key role in understanding how to apply the business judgment rule to decisions made by artificial intelligence directors.⁴⁸

B. Artificial Intelligence Devices in Industry and Businesses Today

Elon Musk, the founder and CEO of Tesla, Inc., has famously called his mostly automated “Gigafactory” a “machine that builds the machine.”⁴⁹ Tesla is one of the many companies invested in the development of fully autonomous self-driving cars—this technology is not only poised to fundamentally change how individuals will travel in personal automobiles, but could also change the operation of ride hailing services and, with the release of autonomous semitrucks, the trucking industry. With the transportation and trucking industry comprising a large sector of the U.S. economy (in terms of employment) it is not hard to see how artificial intelligence will be extremely disruptive to industry as a whole, especially if factory work continues to be automated as well.⁵⁰

⁴⁴ See *id.* at 532, 545–59.

⁴⁵ See *id.* at 538.

⁴⁶ See Tom C.W. Lin, *The New Investor*, 60 UCLA L. REV. 678, 704 (2013).

⁴⁷ *Id.*

⁴⁸ See *infra* Part IV.

⁴⁹ Sean O’Kane, *Tesla Will Live and Die by the Gigafactory*, VERGE (Nov. 30, 2018, 10:01 AM), <https://www.theverge.com/transportation/2018/11/30/18118451/tesla-gigafactory-nevada-video-elon-musk-jobs-model-3> [<https://perma.cc/9YLY-N5A4>].

⁵⁰ See, e.g., *id.*; Nick Wingfield, *As Amazon Pushes Forward With Robots, Workers Find New Roles*, N.Y. TIMES (Sept. 10, 2017), <https://www.nytimes.com/2017/09/10/technology/amazon-robots-workers.html> [<https://perma.cc/TJT6-AVZC>].

In the finance industry, artificial intelligence devices are already being used to effectively manage exchange traded funds (ETFs) and execute high frequency trades (HFTs) on an unprecedented scale.⁵¹ Artificial intelligence supercomputers, like IBM's Watson, best known for its performance on the trivia game show "Jeopardy!," have even been able to beat the market.⁵² However, the use of artificial intelligence has not always gone well on Wall Street. In 2010 and again in 2015, artificial intelligence managed ETFs caused "flash crashes" resulting in the loss of trillions of dollars for investors, though blame was pinned on human individuals.⁵³

As far as the boardroom is concerned, every state in the United States currently limits the availability of board positions to "natural persons," thus excluding artificial intelligence devices.⁵⁴ However, several foreign jurisdictions, including the U.K., Cayman Islands, and Hong Kong have less stringent limitations on who or what can serve as a member of the board.⁵⁵ While Vital in Hong Kong is currently the only artificial intelligence device with a board seat, the World Economic Forum released a 2015 report where nearly half of the 800 IT executives surveyed expected additional artificial intelligence devices to be on corporate boards by 2025.⁵⁶ Were this to occur, it would be a tipping point ushering in a new era of artificially intelligent board members.

It is not hard to imagine how this could play out. A state like California, which is already well established as the technology capital of the United States, could begin to see artificial intelligence directors succeeding as directors on the boards of foreign companies and loosen its "natural persons"

⁵¹ See Paul J. Lim, *The First Ever Fund Managed by a Robot Is Here. So Far It's Beating the Market*, MONEY (Oct. 25, 2017), <http://money.com/money/4993744/robot-mutual-fund-beating-stock-market/> [<https://perma.cc/7CRC-SF86>].

⁵² *Id.*; Clive Thompson, *What is I.B.M.'s Watson?*, N.Y. TIMES (June 20, 2010), <https://www.nytimes.com/2010/07/04/magazine/04Letters-t.html> [<https://perma.cc/N2MB-L4WU>].

⁵³ Mark Melin, *Here's What Actually Caused the 2010 "Flash Crash"*, BUS. INSIDER (Jan. 30, 2016, 10:57 AM), <https://www.businessinsider.com/what-actually-caused-2010-flash-crash-2016-1> [<https://perma.cc/ZUF6-LKKQ>].

⁵⁴ See, e.g., DEL. CODE ANN. TIT. 8, § 141(b) (2019); Stephen M. Bainbridge, *Corporate Directors in the United Kingdom*, 59 WM. & MARY L. REV. ONLINE 65, 67 n. 3 (2017–2018).

⁵⁵ See *id.*; The Directors Registration and Licensing Law (Law 10/2014) § 2 (Cayman Is.) (defining "corporate director"); Companies Ordinance, (2014) Cap. 622, 1§ 2(1) (H.K.) (defining a director as including "any person occupying the position of director (by whatever name called)"); Companies Ordinance, (2014) Cap. 622, § 457 (H.K.) (requiring at least one director who is a natural person).

⁵⁶ GLOBAL AGENDA COUNCIL ON THE FUTURE OF SOFTWARE & SOCIETY, WORLD ECONOMIC FORUM, DEEP SHIFT TECHNOLOGY TIPPING POINTS AND SOCIETAL IMPACT 4 (2015), http://www3.weforum.org/docs/WEF_GAC15_Technological_Tipping_Points_report_2015.pdf [<https://perma.cc/3DCF-PWQQ>].

requirement to adapt. Then a state like Delaware, in an attempt to retain its title as winner of the “race to the bottom” for corporate friendly laws, would have to amend its current laws in order to avoid losing out on the registration fees obtained from businesses determined to reap the benefits of artificial intelligence serving as board members (i.e. lower costs and higher efficiency). Delaware has already taken steps toward attracting tech companies by adopting rules addressing blockchain technologies.⁵⁷ In doing so, Delaware has demonstrated its willingness to adapt its laws to new and innovative technologies in the business sector.

C. How Artificial Intelligence Devices Make Business Judgments

In order to understand how the business judgment rule will apply to artificial intelligence devices, it is necessary to understand how they make business judgments. Unfortunately, doing so runs into one of the biggest problems facing wide scale adoption of artificial intelligence: it is difficult for an artificial intelligence device to explain why it reaches the conclusions that it does. Generally speaking, large amounts of data are fed to the artificial intelligence system.⁵⁸ This data is sorted both quantitatively and qualitatively using machine learning processes like natural language processing.⁵⁹ The data is evaluated for matter such as sentiment, semantic roles, and concepts.⁶⁰ Then, after running through an algorithm, the artificial intelligence machine supplies a judgment. This lack of transparency is commonly referred to as the “black box of AI.”⁶¹ To address this issue, companies like IBM are pouring great amounts of resources into creating programs and processes that allow artificial

⁵⁷ Press Release, Delaware Office of the Governor, Governor Markell Launches Delaware Blockchain Initiative (May 2, 2016), <https://www.prnewswire.com/news-releases/governor-markell-launches-delaware-blockchain-initiative-300260672.html> [<https://perma.cc/C5DM-HT5D>]; see Wonnie Song, *Bullish on Blockchain: Examining Delaware’s Approach to Distributed Ledger Technology in Corporate Governance Law and Beyond*, 8 HARV. BUS. L. REV. ONLINE 9, 17–18 (2017–2018).

⁵⁸ See generally *Recommend with Confidence*, IBM, <https://www.ibm.com/watson/advantages/recommend> [<https://perma.cc/2LDF-VL6A>].

⁵⁹ See generally *Watson Discovery*, IBM, <https://www.ibm.com/watson/services/discovery/> (accessed Oct. 30, 2019) [<https://perma.cc/K7XJ-QNA7>].

⁶⁰ See *id.*; Frank Zhao, *Natural Language Processing – Part I: Primer*, S&P GLOBAL MARKET INTELLIGENCE, (Sept. 2017), <https://www.spglobal.com/marketintelligence/en/documents/sp-global-market-intelligence-nlp-primer-september-2018.pdf> [<https://perma.cc/C7YX-4893>].

⁶¹ Tasmin Lockwood, *Artificial Intelligence Can Now Explain Its Own Decision Making*, MEDIUM (Sept. 23, 2018), <https://medium.com/datadriveninvestor/artificial-intelligence-can-now-explain-its-own-decision-making> [<https://perma.cc/2YRD-BM9D>].

intelligence devices to explain how they reach decisions in ways that humans can understand.⁶²

The lack of an artificial intelligence device's ability to explain how it comes to conclusions is a major hindrance to humanity's ability to trust such devices. As it currently stands, business judgments by artificial intelligence devices will likely be called into question quite frequently by human shareholders. This is commonly cited as one of the primary reasons artificial intelligence is not ready to take on the role of a director.⁶³ But on closer inspection, this is not much different than how humans make decisions. In a study done by the University of Boston and the University of California Berkeley, researchers attempted to understand the black box of artificial intelligence and to make it understandable for humans.⁶⁴ In explaining why it is difficult to cross this understanding barrier, Kate Saenko, one of the lead researchers on the project, noted "[i]t's the same reason that we don't understand how people think," that is "[y]ou could ask me why I wore this shirt today and I could come up with some rationalization, but who knows how my thinking really works? I don't know what my brain process was really like."⁶⁵ Perhaps it is not quite as important to understand exactly how artificial intelligence devices make decisions: as with humans, we can just look to the information that was relied upon and the circumstances under which the decision was made.

D. Summary

The fourth industrial revolution is upon us. Advancements in machine learning, natural language processing, and artificial intelligence have all but ensured that artificial intelligence devices will eventually serve on future

⁶² See, e.g., Ben Dickson, *IBM, Harvard Develop Tool to Tackle Black Box Problem in AI Translation*, VENTUREBEAT (Nov. 1, 2019 at 2:15 PM), <https://venturebeat.com/2018/11/01/ibm-harvard-develop-tool-to-tackle-black-box-problem-in-ai-translation> [<https://perma.cc/6UJV-R3P9>].

⁶³ See, e.g., Pugh, *supra* note 5; *Will AI Board Members Run the Companies of the Future?*, BRINK (June 14, 2018), <https://www.brinknews.com/will-ai-board-members-run-the-companies-of-the-future/> [<https://perma.cc/RN6B-XUMZ>]. See also, Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J. L. & TECH. 889 (2018); Will Knight, *The Dark Secret at the Heart of AI*, MIT TECH. REV. (Apr. 11, 2017), <https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/> [<https://perma.cc/D4TH-VU2W>].

⁶⁴ Ronghang Hu et al., *Explainable Neural Computation via Stack Neural Module Networks*, (arXiv:1807.08556, July 23, 2018), <https://arxiv.org.proxy.library.cornell.edu/pdf/1807.08556.pdf> [<https://perma.cc/L4JT-HZTD>].

⁶⁵ Art Janhke, *To Make Sense of A.I. Decisions, 'Peek Under the Hood'*, FUTURITY (Oct. 8, 2018), <https://www.futurity.org/artificial-intelligence-decision-making> [<https://perma.cc/9VJU-USFW>].

boards of directors. The emergent nature of this technology will surely lead artificial intelligence directors to make decisions that no one can predict, which will be likely to have both positive and negative consequences. The inability of artificial intelligence devices to explain their rationale at all, let alone in a way that humans will understand, plays negatively to the social valence relationship with shareholders. However, this is not unlike the current role that directors play. Even if human directors give satisfactory explanations for why they reach their decisions, there is no way to truly understand their thought processes. In some ways, this is exactly the shortcoming that the business judgment rule is meant to mitigate.

IV. THE FUTURE OF THE BUSINESS JUDGMENT RULE

A. Legal Application of the Business Judgment Rule to Artificial Intelligence

Application of the business judgment rule to artificial intelligence devices serving as directors is relatively straightforward. First, there must be a questioned business decision. For example, take an AI device called AID (for Artificially Intelligent Director) which serves on the board of XYZ Corp. After being fed data provided by accountants, lawyers, and investment bankers, AID decides that XYZ Corp. should merge with another company, ABC Inc. Unfortunately, the new venture does not pan out and the shares lose value. As a result of this, a shareholder brings a derivative suit against AID. The court will first take stock of AID's other interests, that is: does AID have a financial interest in ABC Inc? If not, then the court will look to see if proper procedure was followed. Here, the information relied upon was obtained solely from sources which a director is permitted to rely upon and will fall squarely within the confines of the business judgment rule.⁶⁶ The only hurdle left is to determine whether AID exercised the care that a reasonably prudent person would under the circumstances.

In the past, as demonstrated in *Van Gorkom*, whether reasonable care can be applied is at least in part a matter of how much time was spent deliberating over the information provided.⁶⁷ Given how AID works, courts may go one of two ways on this point. They could either recognize that

⁶⁶ See Michele Healy Ubelaker, *Director Liability under the Business Judgement Rule: Fact or Fiction*, 35 SMU L. REV. 775, 788 (1981) (explaining that the ABA's model business code provision contains "a right to rely on information, opinions, reports, and statements from board committees, officers and other corporate employees, attorneys, accountants, and other experts whom the director reasonably believes to be reliable.").

⁶⁷ See *Smith v. Van Gorkom*, 488 A.2d 858, 874 (Del. 1985) (finding the directors were grossly negligent in approving the 'sale' of the Company upon two hours' consideration, without prior notice, and without the exigency of a crisis or emergency.").

artificial intelligence devices require less time to analyze vast amounts of data compared to their human counterparts and find reasonable care is exercised, or they could find that the minimal time required for the devices to make the decisions does not evidence reasonable prudence. In either case, a court's finding will set precedent regarding the amount of time an artificial intelligence director must spend making a decision to be protected by the business judgment rule for every future application. Assuming reasonable prudence is found, the business judgment rule will apply and the court will dismiss the case on summary judgment.

Changing the facts slightly also leads to straightforward results. In another scenario, imagine that AID received the information about the potential merger from accountants, lawyers, investment bankers, and an unreliable reporter for a local tabloid who has access to AID and owns stock in ABC Inc. As long as AID is able to articulate where its information comes from and is unable to filter out the unscrupulous reporter's information, then courts will clearly apply the entire fairness review standard—just as courts would if a human made the same error of judgment.⁶⁸ It may in fact be easier to make this case against AID than a human director who can ostensibly lie about the information relied upon, which is interesting given that shareholders are less likely to trust artificial intelligence devices than their human counterparts.⁶⁹

Just like human directors, artificial intelligence devices utilizing emergent technology will make somewhat unpredictable decisions based on the information provided. As with any failed decision made by humans, judgments made by artificial intelligence devices that have negative consequences will likely come under the scrutiny of courts via suits from dissatisfied shareholders. Thus, courts will likely apply the business judgment rule in much the same way as they currently do. That is, as long as the information the artificially intelligent device relies upon comes from reliable sources, there is no finding of bad faith. Then, if the decision is made with the care of a reasonably prudent person under the circumstances, the business judgment rule will apply and shield the artificial intelligent device from liability for the resulting damages. If one of those elements is not met, then the court will apply an entire fairness standard of review.

B. Policy Considerations

The fact that the black letter law of the business judgment rule could apply to an artificial intelligence device is not the end of the story. Indeed, the

⁶⁸ See Sharfman, *supra* note 2, at 39–42.

⁶⁹ See *supra* Section III.C.

business judgment rule should not apply to artificial intelligence devices because it is not necessary to accomplish the policy objectives underlying the rule. The primary policy goals behind the business judgment rule revolve around inducing individuals to serve as directors and encouraging the necessary amount of risk-taking.⁷⁰

Artificially intelligent devices, however, do not need incentives to serve on boards. Not only are they unpaid for their work, the company will likely purchase or create the device directly for the purpose of serving on the board. The device does not have an active decision to make—it is either bought to serve on the board or it is not. Thus, an artificially intelligent device does not need an assurance that it will not be held liable for poor judgments to agree to serve on the board—it will serve on the board at the will of the purchaser. Additionally, the artificially intelligent device, having no assets of its own, will have no reason to fear liability. In the event that an artificially intelligent device makes a poor decision, the company can have someone who understands the algorithm (perhaps even its creator) try to determine why the machine made that decision and adjust the algorithm to prevent such decisions from being made in the future. The artificial intelligence device, incapable of having emotions, cannot fear the repercussions of poor judgments, thus the business judgment rule is unnecessary to encourage the appropriate amount of risk-taking necessary to maximize shareholder wealth.

There is, however, one policy reason that still applies to artificial intelligence devices under these circumstances—judicial efficiency. Regardless of who or what is serving as a member of the board, judges will still not want to be in the position to second-guess business decisions.⁷¹ Additionally, the number of suits brought against artificial intelligence devices could clog up court dockets just as easily as traditional suits against human directors. At the same time, if artificial intelligence directors work as they are meant to, the number of decisions that will need to be second-guessed should be far lower than human errors of judgment. Additionally, because the process of determining what information a machine relied upon is clearer-cut than it is with human directors, judges will have an easier time discerning what information is relevant. Ultimately, suits against artificial intelligence directors will likely be far less common than those against human directors for one simple reason: the artificial intelligence director is judgment-proof. If the director has no assets, then there is nothing to be gained from holding them personally liable for their poor judgments. Unless there is some sort of guarantor for the device, such as the developer, then cases will be seldom brought. Even in the event that its developer is liable, the correct approach

⁷⁰ See *supra* Section II.B.

⁷¹ See *id.*

may not be a business decision review at all but something more akin to product defect in which case the business judgment rule is irrelevant.⁷²

V. CONCLUSION

The business judgment rule plays a crucial role in the smooth operation of modern businesses. Under the rule, courts will not second guess decisions made by directors if they are made in good faith, with the honest belief that the decisions are for the good of the company and are made with the care of a reasonably prudent person under the circumstances. As technology advances and artificial intelligence devices with unpredictable emergent behavior begin to serve as directors, this rule will warrant a review of its own. While the application of the business judgment rule to artificial intelligence directors is straightforward, the policy underlying the need for the rule is obsolete with regards to artificial intelligence devices. Artificial intelligence devices do not suffer from the same fears as their human director counterparts. They have no assets nor the ability to turn down a position on the board and thus do not need the same assurance that they will not be held personally liable for the decisions that they make. While the social valance surrounding artificial intelligence is currently one of distrust, the fact that they are judgment-proof will still ensure that courts are not bogged down in suits against them. Just as human directors may become obsolete due to the rise of machines in the future, so too will the business judgment rule become obsolete for artificial intelligence devices serving on boards.

⁷² However, it is still unclear how product liability will function in regard to artificial intelligence devices. See Greg Swanson, *Non-Autonomous Artificial Intelligence Programs and Products Liability: How New AI Products Challenge Existing Liability Models and Pose New Financial Burdens*, 42 SEATTLE U. L. REV. 1201 (2019).

CFIUS AND A.I.: DEFENDING NATIONAL SECURITY WHILE ALLOWING FOREIGN INVESTMENT

Theodore Bruckbauer*

CITE AS: 4 GEO. L. TECH. REV. 279 (2019)

TABLE OF CONTENTS

I. INTRODUCTION	280
II. CFIUS BACKGROUND AND PROCESS	281
A. CFIUS Background.....	281
B. CFIUS Review Process.....	282
III. FIRMA AND THE IMPLICATION OF ARTIFICIAL INTELLIGENCE.....	285
A. Overview.....	285
B. Legislative History and Support	285
IV. MITIGATION MEASURES	287
A. Description & Purpose.....	287
B. Traditional Mitigations as They Pertain to Artificial Intelligence..	288
1. <i>Integrity Assurance</i>	288
2. <i>Exclusion of Sensitive Assets from the Transaction</i>	290
3. <i>Access Restrictions</i>	291
4. <i>Access Restrictions Compared to Other Mitigations</i>	292
C. Need for New Mitigation Options	293
V. CONCLUSION.....	295

* J.D., Northwestern Pritzker School of Law, 2019; M.B.A., Kellogg School of Management at Northwestern, 2019; BPhil, Northwestern School of Continuing Studies, 2013. I would like to thank my wife, Kirsten, for her wisdom and partnership throughout my academic pursuits, and our son, Evan, for always finding time in my schedule to play. I extend my gratitude to the Georgetown Law Technology Review and its editors for graciously facilitating the publication of this note.

I. INTRODUCTION

Artificial intelligence is one of the most important innovations to impact national security in recent years.¹ Among the national security concerns of the United States is that foreign countries seek to erode America's leadership in artificial intelligence development by buying or merging with U.S. companies.² In response to this perceived threat, Congress passed legislation in 2018 to expand the authority of the Committee on Foreign Investment in the United States (CFIUS).³ CFIUS conducts reviews for national security concerns on certain transactions where a foreign-controlled entity aims to acquire a stake in a U.S.-based company.⁴ If national security concerns are not resolved during its review, CFIUS may recommend the President block the transaction.⁵

For foreign companies that pose perceived threats to national security, an important part of the CFIUS review process is mitigations—binding agreements between CFIUS and the transacting companies that resolve security concerns and allow the transaction to receive approval. Foreign companies with proposed investments in U.S. artificial intelligence that could be perceived to threaten national security should expect CFIUS to approve transactions only after mitigations are in place. Those companies must therefore understand the unique national security concerns posed by artificial intelligence and which mitigations might resolve them. This paper seeks to provide that guidance.

Part II of this paper provides an overview of CFIUS and its review process. It explains the stages of CFIUS review and places mitigations inside this larger context. Part III introduces the 2018 legislation that targeted artificial intelligence transactions and reviews its legislative history. Part IV is devoted to mitigations. It first assesses the usefulness of traditional mitigations in artificial intelligence transactions and then recommends sources to aid in development of new categories of mitigations.

¹ See generally DANIEL S. HOADLEY & KELLEY M. SAYLER, CONG. RESEARCH SERV., R45178, ARTIFICIAL INTELLIGENCE AND NATIONAL SECURITY (2019) [hereinafter CRS AI & NATIONAL SECURITY REPORT].

² See Paul Mozur & John Markoff, *Is China Outsmarting America in A.I.?*, N.Y. TIMES (May 27, 2017), <https://www.nytimes.com/2017/05/27/technology/china-us-ai-artificial-intelligence.html> [<https://perma.cc/4MFJ-GHAA>].

³ JAMES K. JACKSON, CONG. RESEARCH SERV., IF10952, IN FOCUS: CFIUS REFORM: FOREIGN INVESTMENT NATIONAL SECURITY REVIEWS 1 (2018).

⁴ JAMES K. JACKSON, CONG. RESEARCH SERV., RL33388, THE COMMITTEE ON FOREIGN INVESTMENT IN THE UNITED STATES (CFIUS) 1 (2018) [hereinafter CRS CFIUS REPORT].

⁵ *Id.* at 13.

II. CFIUS BACKGROUND AND PROCESS

A brief history and description of CFIUS, its purpose, and the CFIUS review process provide context for subsequent discussions of recent changes in the review process.

A. CFIUS Background

The Committee on Foreign Investment in the United States (CFIUS) is an interagency committee that serves the President in overseeing the national security implications of foreign direct investment (FDI) in the U.S. economy.⁶ It consists of nine mostly Cabinet-level officials, including the Secretaries of State, the Treasury, Defense, Homeland Security, Commerce, and Energy; the Attorney General; the United States Trade Representative; and the Director of the Office of Science and Technology Policy.⁷ The Secretary of Labor and the Director of National Intelligence serve as *ex officio* members of the Committee.⁸ President Gerald Ford established the Committee in a 1975 Executive Order and granted to it “primary continuing responsibility within the Executive Branch for monitoring the impact of foreign investment in the United States.”⁹ Among other responsibilities, the Order instructed the Committee to review investments which “might have major implications for United States national interests.”¹⁰ President Ford created the Committee to encourage foreign investment and “dissuade Congress from enacting new restrictions.”¹¹

In 1988, the Exon–Florio Amendment to the Defense Production Act (“Exon–Florio”) codified the process CFIUS used to review foreign investment transactions.¹² Importantly, Exon–Florio also granted power to the President to block mergers, acquisitions, and takeovers that threaten to impair national security. The Amendment grants this authority whenever the President has “credible evidence” that the investment will impair national security and no other U.S. laws adequately protect U.S. security interests.¹³

A contemporaneous Executive Order delegated the President’s power to conduct reviews, undertake investigations, and make recommendations to

⁶ *Id.* at 1.

⁷ *Id.* at 14.

⁸ *Id.*

⁹ Exec. Order No. 11858, 3 C.F.R. § 990 (1971–1975).

¹⁰ *Id.*

¹¹ *The Operations of Federal Agencies in Monitoring, Reporting on, and Analyzing Foreign Investments in the United States: Hearings Before the Subcomm. on Commerce, Consumer, and Monetary Affairs*, 96th Cong. 334–335 (1979).

¹² CRS CFIUS REPORT, *supra* note 4, at 6.

¹³ *Id.* at 7.

CFIUS.¹⁴ Thus, CFIUS now performs investigations and makes recommendations to the President when it believes a transaction should be blocked. The 1992 Byrd Amendment to the Defense Production Act went on to mandate reviews whenever a foreign acquirer acts on behalf of a foreign government.¹⁵

2007 brought a major update to CFIUS through the Foreign Investment and National Security Act of 2007 (FINSA).¹⁶ FINSA codified CFIUS's position in the review and recommendation process and increased the number of factors the President could consider in making a decision.¹⁷ It also required high-level CFIUS members to certify to Congress that no unresolved national security issues exist in reviewed transactions.¹⁸ Certification must be made by a person ranking no lower than Assistant Secretary level for reviewed transactions and Secretary or Deputy Secretary level for investigated transactions.¹⁹

The next major update to CFIUS took place in 2018.²⁰ The Foreign Investment Risk Review Modernization Act of 2018 (FIRRMA) was passed as part of the John S. McCain National Defense Authorization Act for Fiscal Year 2019.²¹ FIRRMA made numerous changes that expanded the scope of transactions that fall under CFIUS's review by redefining "covered transactions" to include joint ventures and noncontrolling investments in critical—and emerging—technology companies.²² Additionally, FIRRMA mandates filing of certain transactions with CFIUS.²³ These changes were put in effect through a Treasury Department pilot program on November 10, 2018.²⁴

B. CFIUS Review Process

The CFIUS review process is comprised of one informal step and three formal steps.²⁵ The informal step is one of indefinite length during which

¹⁴ *Id.* at 6.

¹⁵ *Id.*

¹⁶ Foreign Investment and National Security Act of 2007, Pub. L. No. 110-49, 121 Stat. 246 (codified as amended in scattered sections of 50 U.S.C.).

¹⁷ CRS CFIUS REPORT, *supra* note 4, at 12, 14.

¹⁸ *Id.* at 14.

¹⁹ *Id.*

²⁰ JAMES K. JACKSON, CONG. RESEARCH SERV., IF10177, IN FOCUS: THE COMMITTEE ON FOREIGN INVESTMENT IN THE UNITED STATES 1 (2019) [hereinafter CRS IN FOCUS: CFIUS].

²¹ CRS CFIUS Report, *supra* note 4, at 11.

²² CRS IN FOCUS: CFIUS, *supra* note 20, at 2.

²³ *Id.*

²⁴ CRS CFIUS REPORT, *supra* note 4, at 18.

²⁵ *Id.* at 10.

transactions that will potentially require a formal investigation are screened.²⁶ During this phase, firms considering foreign direct investment work with CFIUS to identify potential issues prior to the beginning of an investigation.²⁷ This period allows firms to address issues that might be raised before risking any negative publicity that might accrue if the transaction is blocked on national security grounds.²⁸

A transaction enters the formal review process when the transacting companies notify CFIUS of their proposed transaction or when CFIUS initiates a review. In the past, many transacting companies entered the formal review process willingly by notifying CFIUS of their proposed investment, merger, or acquisition.²⁹ Firms subjected themselves to scrutiny because transactions completed without CFIUS review are subject to forced divestment by the President at any time in the future.³⁰ With the passage of FIRREA and corresponding regulations, the filing of transactions with CFIUS became mandatory for foreign investments in U.S. businesses that produce, design, test, manufacture, fabricate, or develop one or more critical technologies.³¹ Some lower risk transactions will, however, benefit from a new expedited review process that forgoes a formal review.³² In addition to the parties to an investment transaction, any member of CFIUS or the President may initiate a formal review of that transaction.³³

Once a transaction reaches the formal review process, the process can proceed potentially through three steps: (1) the National Security Review, (2) the National Security Investigation, and (3) Presidential Determination.³⁴ At each step, CFIUS members weigh factors and work with the transacting companies to identify security concerns and attempt to clear the transaction.³⁵ CFIUS can grant approval at either of the first two formal steps and the President can approve or block a transaction at step three.³⁶

The first formal step is the National Security Review. During this step, CFIUS is required to conduct a review if the investment threatens to impair national or homeland security, critical infrastructure, or critical technologies; and the transaction would result in foreign control of a U.S. entity.³⁷ The

²⁶ *Id.* at 12.

²⁷ *Id.* at 11.

²⁸ *Id.*

²⁹ CRS IN FOCUS: CFIUS, *supra* note 20, at 1.

³⁰ *Id.*

³¹ CRS CFIUS REPORT, *supra* note 4, at 17.

³² *Id.*

³³ CRS CFIUS REPORT, *supra* note 4, at 12.

³⁴ CRS IN FOCUS: CFIUS, *supra* note 20, at 1.

³⁵ *Id.*

³⁶ *Id.*

³⁷ *Id.*

Director of National Intelligence, an ex officio member of CFIUS, reviews the national security implications of the foreign investment and CFIUS assesses the impact of the investment using factors enumerated by Congress.³⁸ The Secretary of the Treasury may unilaterally exempt a transaction from the process at this stage if he or she determines the transaction will not impair security.³⁹

The second step is the National Security Investigation. CFIUS launches an Investigation if any of its members determine during the first step that the transaction threatens to impair national security.⁴⁰ The second step is a more thorough review of the national security implications of allowing the transaction to proceed. Critically, this phase allows CFIUS and the parties to the transaction to agree to mitigations.

CFIUS can negotiate, enter into, impose, and enforce “any agreement or condition with any party to the covered transaction in order to mitigate any risk to the national security of the United States that arises as a result of the covered transaction.”⁴¹ Details of agreed upon mitigations are generally not public. CFIUS, however, includes a high-level list of mitigations used during a calendar year in its annual report to Congress. The list is largely static from year-to-year which suggests that many mitigations are commonly used. Section IV below discusses mitigations in detail.

The final step is Presidential Determination. If the transacting parties and CFIUS cannot agree on mitigations, CFIUS can recommend that the President suspend or block the transaction.⁴² The Office of the President has used its authority to block only five transactions under this legal regime.⁴³ The rarity of this action suggests that most companies either agree to mitigations or withdraw once CFIUS intends to recommend Presidential intervention.

In summary, companies proposing transactions that implicate national security are funneled into a process where they must negotiate mitigations that relieve national security concerns or back out of the transaction (independently or by Presidential action). For these companies, finding effective mitigations during the National Security Investigation (or sooner) is key to successful CFIUS approval.

³⁸ *Id.*

³⁹ CRS IN FOCUS: CFIUS, *supra* note 20, at 1.

⁴⁰ *Id.*

⁴¹ 50 U.S.C. § 4565(l)(3)(A)(i) (2018).

⁴² CRS CFIUS REPORT, *supra* note 4, at 13.

⁴³ *Id.* at 7.

III. FIRRMA AND THE IMPLICATION OF ARTIFICIAL INTELLIGENCE

The expanded powers of CFIUS and legislative history of FIRRMA show the legislative concerns that led to CFIUS's expansion.

A. Overview

The Foreign Investment Risk Review Modernization Act of 2018 (FIRRMA) expanded the role of CFIUS.⁴⁴ FIRRMA redefined “covered transactions” to include joint ventures and noncontrolling investment in critical and emerging technology companies.⁴⁵ It also made filing with CFIUS mandatory for certain transactions, lengthened the review periods, and provided CFIUS with additional funding and staff.⁴⁶ The broad changes appear to have a narrow focus—capture strategic investments by foreign countries in emerging technology companies and give CFIUS the tools to review them thoroughly. Despite there being no mention of artificial intelligence within the text of FIRRMA, the analysis below demonstrates that foreign investment in artificial intelligence companies is Congress' primary concern.

B. Legislative History and Support

An emerging consensus in the government that artificial intelligence is a critical emerging technology and requires protection led to FIRRMA's implementation and usage. In a 2017 hearing before the Senate Committee on Banking, Housing, and Urban Affairs, the Senate Committee and witnesses discussed a report drafted by the Defense Innovation Unit Experimental (DIUx)⁴⁷ about China's technology acquisition strategy.⁴⁸ The report provided data and arguments which serve to support FIRRMA.⁴⁹ The main findings of

⁴⁴ CRS IN FOCUS: CFIUS, *supra* note 20, at 2.

⁴⁵ *Id.*

⁴⁶ *Id.*

⁴⁷ The Defense Innovation Unit Experimental (DIUx), now DIU, is an entity, founded by the Department of Defense, that invests in private sector companies to solve national security issues. See Billy Mitchell, ‘No longer an experiment’—DIUx Becomes DIU, *Permanent Pentagon Unit*, FEDSCOOP (Aug. 9, 2018), <https://www.fedscoop.com/diu-permanent-no-longer-an-experiment/> [<https://perma.cc/ETB7-534H>].

⁴⁸ *Examining The Role Of The Committee On Foreign Investment In The United State: Hearing Before the Subcomm. On Monetary Policy & Trade of the H. Comm. On Fin. Servs.*, 115th Cong. 28 (2017).

⁴⁹ See MICHAEL BROWN & PAVNEET SINGH, DEF. INNOVATION UNIT EXPERIMENTAL, CHINA'S TECHNOLOGY TRANSFER STRATEGY: HOW CHINESE INVESTMENTS IN EMERGING TECHNOLOGY ENABLE A STRATEGIC COMPETITOR TO ACCESS THE CROWN JEWELS OF U.S. INNOVATION (2018).

the report were that Chinese investments in venture-backed startups were at record levels and growing rapidly.⁵⁰ Additionally, the report found that the technologies targeted by these investments are the same ones where U.S. firms are investing and will be foundational to future innovations.⁵¹ The report recommends expanding the role of CFIUS to counter these trends.⁵² The technology discussed most prominently in the report and hearing is artificial intelligence.

Both houses of Congress introduced FIRRMA bills on November 8, 2017.⁵³ The House reintroduced a FIRRMA bill on May 16, 2018, where it passed with a vote of four hundred to two.⁵⁴ A final version became law as part of Title XVII of the John S. McCain National Defense Authorization Act for Fiscal Year 2019.⁵⁵

Another example of Congress' concern with foreign investment in U.S. artificial intelligence appeared as Title X of the same act. Title X created the National Security Commission on Artificial Intelligence to review, *inter alia*, “[d]evelopments and trends in international cooperation and competitiveness, including foreign investments in artificial intelligence, related machine learning, and computer science fields that are materially related to national security and defense.”⁵⁶

President Trump joined Congress in his support for using CFIUS to regulate artificial intelligence transactions. In 2017, he followed a recommendation from CFIUS and blocked the foreign acquisition of Lattice Semiconductor, a company that manufactures chips critical to artificial intelligence development.⁵⁷ In 2018, he voiced support for the idea of using CFIUS to protect technology developed by Silicon Valley.⁵⁸ Finally, in 2019, President Trump issued an Executive Order that encouraged concerted efforts to protect America's leadership in artificial intelligence.⁵⁹

⁵⁰ *Id.* at 5.

⁵¹ *Id.* at 2.

⁵² *Id.*

⁵³ Foreign Investment Review and Modernization Act of 2017, H.R. 4311, 115th Cong. (2017); Foreign Investment Risk Review Modernization Act of 2017, S. 2098, 115th Cong. (2017).

⁵⁴ 164 CONG. REC. H4137 (daily ed., May 16, 2018) (introduction of H.R. 5841 by Rep. Robert Pittenger).

⁵⁵ John S. McCain National Defense Authorization Act for Fiscal Year 2019, Pub. L. No. 115-232, 132 Stat. 1636 (2018).

⁵⁶ *Id.* § 1051(b)(2)(C).

⁵⁷ CRS AI & NATIONAL SECURITY REPORT, *supra* note 1, at 7.

⁵⁸ See Shawn Donnan, *Donald Trump Softens Tone on Chinese Investments*, FIN. TIMES (Jun. 26, 2018) <https://www.ft.com/content/3ce53380-798f-11e8-bc55-50daf11b720d> [<https://perma.cc/Y7CT-D7T9>].

⁵⁹ Exec. Order No. 13859, 84 Fed. Reg. 3967 (Feb. 11, 2019).

Together, these actions point to a broad consensus in the government that artificial intelligence is a critical emerging technology that requires protection. Congress intends CFIUS to provide that protection and passed FIRRMA to ensure the Committee has the authority and resources needed to do so. With review of artificial intelligence transactions certain to become the norm, CFIUS and transacting companies must now learn to mitigate the security risks that arise from such transactions.

IV. MITIGATION MEASURES

This section describes mitigations and their place in the CFIUS approval process, discusses how traditional mitigations may be applied to artificial intelligence transactions, and suggests sources for new mitigations that address the unique traits of artificial intelligence.

A. Description & Purpose

Mitigations are binding agreements between CFIUS members and the transacting companies that resolve security concerns to allow transaction approval. CFIUS's declassified yearly report includes a bullet-point list of mitigation measures "negotiated and adopted" that "required the businesses involved to take specific and verifiable actions."⁶⁰ For foreign companies that pose perceived threats to national security, negotiated mitigations are perhaps the most important step in the CFIUS process. When the Committee determines national security concerns exist, mitigations are required before transaction approval. Without mitigations, the only path to approval in these cases is by presidential ruling against the advice of the Committee.

As already noted, the mitigations are negotiated between the transacting companies and CFIUS. Neither side has dictatorial control of the terms. Instead, CFIUS looks to reach an agreement that meets an acceptable security standard and the companies try to accommodate that standard at minimal cost to themselves. From 2013 to 2015, forty cases resulted in the use of legally binding mitigation measures.⁶¹ But, finding workable mitigations is by no means guaranteed. In 2015, at least three transactions were abandoned after CFIUS and the parties could not identify acceptable mitigations.⁶²

CFIUS does not publicly divulge the national security concerns that lead to the introduction of mitigations into the review of a transaction. Thus, an understanding of CFIUS's intent must be extracted from the mitigations themselves. A pragmatic interpretation is useful here and will reveal three

⁶⁰ COMM. FOREIGN INVEST. U.S., ANNUAL REPORT TO CONGRESS (CY 2015) 21 (2017).

⁶¹ *Id.*

⁶² *Id.* at 20.

traits about mitigations. First, each mitigation alludes to one or more concern CFIUS has with a transaction. If there were no corresponding concern, there would be no cause for the mitigation to exist. Second, CFIUS views the mitigation as acceptably resolving the concern or it would not have approved it. Finally, the imposed control is not so onerous on the companies or so material to the transaction that the parties prefer to abandon the transaction instead of agreeing. In other words, the inclusion in the annual report of a mitigation shows that at least some companies accepted the mitigation and its associated costs.

These three traits are demonstrated in an example of a mitigation from a recent CFIUS report: that the business must “[n]otify security officers or relevant [U.S. Government] parties in advance of foreign national visits to the U.S. business for approval.”⁶³ First, this mitigation reveals CFIUS’s concern about access to domestically located secrets of the U.S. company by foreign nationals. Second, it reveals that vetting foreign visitors resolves that concern. Lastly, agreement to this mitigation by the transacting companies shows that getting pre-approval for foreign visitors is not an overly large burden, perhaps because it is an infrequent occurrence within these companies.

This understanding of the purpose and interpretation of mitigations makes it possible to evaluate how traditionally used mitigations might be used in future artificial intelligence transactions.

B. Traditional Mitigations as They Pertain to Artificial Intelligence

This section will address how traditional mitigations commonly found in CFIUS reports—integrity assurance, exclusion of sensitive assets from transactions, and access restrictions—can be applied to artificial intelligence. While these mitigations could be used in transactions dealing with artificial intelligence and national security, issues unique to artificial intelligence pose problems in applying these traditional mitigations.

1. *Integrity Assurance*

Of the previously used mitigations, the only to mention software is integrity assurance: “[s]ecurity protocols to ensure the integrity of goods or software sold to the [U.S. Government].”⁶⁴ Though CFIUS does not list the specific concerns that spawned a mitigation, this mitigation is likely designed to address a concern that a foreign controlled company that provides software may make dangerous or undesirable changes without the U.S. government’s detection.

⁶³ *Id.* at 21.

⁶⁴ *Id.*

A scenario leading to this mitigation is easy to imagine. For example, a national security concern would arise if an untrusted foreign company acquired a U.S. software company that develops software for the U.S. government. The software might be custom and application specific—such as code for weapons systems—but could also be mundane business software, like word processors or spreadsheet software. In either case, there is a risk that foreign actors in control of strategic software development could make changes, malicious or otherwise, that have adverse effects. Both intentionally and inadvertently introduced vulnerabilities can compromise the integrity of mission-critical software.

Principles used to protect the integrity of traditional software are transferable to artificial intelligence software, with one notable exception. Where traditional software is expected to be deterministic (i.e. it returns the same correct output for each possible input), artificial intelligence is useful because of its efficacy at tasks that cannot be described by deterministic rules.⁶⁵ Furthermore, correctly functioning artificial intelligence systems always display some rate of error.⁶⁶ Because of these traits, a method of integrity assurance other than testing inputs and outputs is required.

One possible solution to this problem would be limiting changes to the software's development. If the U.S. government trusts software prior to a foreign acquisition, it must believe that the development process will not create unwanted changes. For that trust to continue after an acquisition, that process must be protected. This type of integrity assurance might be possible, but it has three limitations.

First, even the strictest agreement to protect the development process may not satisfy CFIUS. Bias introduced intentionally or unintentionally might result from new or different programmers, training data, preferences for different statistical models, or tuning decisions. Bias in artificial intelligence can be extremely hard to detect.⁶⁷ Without extreme controls, CFIUS may not view development protections as adequate.

Second, introducing controls that meet CFIUS's standards may be too costly for the parties to the transaction. Monitoring becomes more costly as the number of decisions under scrutiny increases. At some level of granularity, the cost of monitoring will outweigh the value of the project. Said another

⁶⁵ HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, EUR. COMM., A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES 3–5 (2019).

⁶⁶ *Id.* at 4.

⁶⁷ See Jeremy Kun, *Big Data Algorithms Can Discriminate, And It's Not Clear What To Do About It*, THE CONVERSATION (Aug. 13, 2015) <http://theconversation.com/big-data-algorithms-can-discriminate-and-its-not-clear-what-to-do-about-it-45849> [<https://perma.cc/5HSP-P5FL>].

way, it might be so expensive to appease CFIUS that companies may choose to withdraw from the transaction rather than agree to the controls.

Finally, any controls might materially reduce the value of the acquired artificial intelligence if they prevent future development or new uses. Unlike the above point, the controls do not need to be extreme to have an effect. Even inexpensive controls, such as limiting training data to a trusted source, might create a competitive disadvantage against others operating without similar limitations. Limitations that greatly reduce the value of a targeted company's technology will cause acquirers to abandon transactions.

In conclusion, non-deterministic outputs, the risk of undetectable bias, and costs arising from development controls make it unlikely that CFIUS and transacting companies will find workable mitigations based on integrity assurance.

2. *Exclusion of Sensitive Assets from the Transaction*

Another mitigation that appears in the annual reports is the "exclusion of sensitive assets from the transaction."⁶⁸ Interestingly, the use of this mitigation shows that companies are willing to complete at least some transactions even when the most sensitive technologies of the acquired company are withheld and sold to third parties.

Machine learning, a form of artificial intelligence, is a combination of statistical models and processes encoded in software, trained on data, often with manual refinements and tuning made after deployment.⁶⁹ Any one of these components might on its own be considered too sensitive to sell to a foreign power. For example, CFIUS might decide that the data underlying the models is too sensitive but the models themselves are uncontroversial, or precisely the opposite. It is also possible that isolating a single component of concern is not possible: some artificial intelligence may be more than the "sum of its parts."

This difficulty can be illustrated with two examples. First, consider a hypothetical case where an artificial intelligence is part of a larger system: a company that gives conventional missiles devastatingly effective results through the inclusion of custom artificial intelligence targeting software. If a foreign entity attempted to buy this company, CFIUS might insist on excluding the artificial intelligence software over concerns it be used against the United States. Assuming the buyer only values the capabilities of a complete system, they are unlikely to agree to this limitation. Where artificial

⁶⁸ COMM. FOREIGN INVEST. U.S., *supra* note 60, at 22.

⁶⁹ *Data Science and Machine Learning*, IBM, <https://www.ibm.com/analytics/machine-learning> (accessed Nov. 1, 2019) [<https://perma.cc/6R34-9ZMG>].

intelligence is a critical part of a system, attempts to exclude it from the transaction will lead companies to abandon otherwise viable transactions.

Second, consider the case where a U.S. company makes only general-purpose artificial intelligence and does not produce a product with identifiable sensitive components or with readily apparent uses. How might CFIUS attempt to judge the threat posed by a foreign acquisition in this case? One option is to attempt to predict potential uses of the acquired technology to block transactions that pose a future threat. It seems unlikely, however, that these predictions will form a reliable basis for CFIUS intervention. Alternatively, CFIUS might attempt to block transactions involving technology so advanced that it is *de facto* sensitive regardless of its potential uses. Aside from the difficulty in making such a determination, this method potentially projects that the United States only allows foreign investment in sub-par technologies. It is unclear how this might be resolved.

These two cases show the challenges with trying to mitigate transactions by excluding all or part of a company's artificial intelligence assets. The barriers present in the sale of stand-alone artificial intelligence technology and those present in the sale of integrated artificial intelligence systems will often occur together, further complicating the conversation. Where artificial intelligence is material to the transaction, any attempt to exclude it either wholly or in-part is likely to adversely affect transactions.

3. *Access Restrictions*

Several of the mitigations listed in CFIUS's annual reports revolve around the concept of limiting untrusted parties' access to sensitive technology, products, and services.⁷⁰ Specifically, these mitigations require "[e]nsuring that only authorized persons have access to certain technology;" "[e]nsuring that only U.S. citizens handle certain products and services, and ensuring that certain activities and products are located only in the [U.S.];" and "[n]otifying security officers or relevant [U.S. Government] parties in advance of foreign national visits to the U.S. business for approval."⁷¹ Collectively, these mitigations show that some negotiated approvals resulted in companies erecting screens between sensitive information and non-U.S. personnel.

Limiting untrusted parties' access to sensitive technology limits their ability to tamper with the technology and their ability to capture and transfer

⁷⁰ See, e.g., COMM. FOREIGN INVEST. U.S., ANNUAL REPORT TO CONGRESS (CY 2012) 18 (2013); COMM. FOREIGN INVEST. U.S., ANNUAL REPORT TO CONGRESS (CY 2013) 20 (2015); COMM. FOREIGN INVEST. U.S., ANNUAL REPORT TO CONGRESS (CY 2014) 21 (2016); COMM. FOREIGN INVEST. U.S., *supra* note 60.

⁷¹ COMM. FOREIGN INVEST. U.S., *supra* note 60.

the technology. In other words, access restrictions are alternative means to achieve some of the objectives outlined in Sections IV.B.1 and IV.B.2 above. However, unlike the previously mentioned integrity assurance or exclusion mitigations, access restrictions achieve different outcomes by targeting people rather than processes or assets.

Applying access restrictions to protect integrity is different than assuring integrity. When only trusted individuals influence the creation process, the end-product should be trustworthy, even if not verifiably so. That logic does not foreclose the usefulness of integrity assurance, but properly implemented restrictions may obviate it to some degree.

Conversely, access restrictions to prevent loss of sensitive technology to foreign adversaries is a mirror image of asset exclusion. Instead of excluding the assets from the transaction, the people are excluded. It would not make sense to have both regarding the same piece of technology.

4. *Access Restrictions Compared to Other Mitigations*

Comparing the effectiveness of the above mitigations as applied to artificial intelligence, access restriction mitigations are more appropriate and effective than the assurance and exclusion mitigations, as four unsolved issues that arise from integrity assurance and asset exclusion can be partially resolved through access restrictions.

The first issue with integrity assurance raised in Section IV.B.1 is that it may be impossible to reduce unwanted foreign influence on development enough to satisfy the integrity demands of CFIUS. This is because even small interactions might introduce an undetectable bias. The desired reduction in foreign influence may be accomplished by altogether removing access. Access restrictions have a distinct advantage over assurance because it preempts any opportunity for bias or tampering to seep in.

The second issue raised with integrity assurance, that the cost of monitoring the development of artificial intelligence systems might be too high, is another situation in which access restrictions have a distinct advantage. A system that approves clearance for known personnel is more efficient than oversight of individual design choices; it is also easier to enforce and audit.

The third issue with integrity assurance is that even minimal controls might materially affect the value of the acquired technology. Section 1 noted that artificial intelligence might develop less competitively under method or data limitations. Likewise, artificial intelligence will develop less efficiently if access restrictions prevent access by highly skilled but foreign data scientists. Access restrictions, however, may target persons other than the scientists and developers. It is therefore possible that some restrictions would

satisfy CFIUS's national security concerns without negatively impacting development.

Lastly, access restrictions may be more effective than exclusion mitigations. Section IV.B.2 discusses attempts to protect sensitive artificial intelligence from foreign ownership by excluding it in whole or in-part from transactions and concludes that there may be insurmountable problems with that approach. As in Section 2, it is necessary in restricting access to evaluate artificial intelligence by its individual subcomponents, and thus the difficulty of isolating components of concern remains. However, if CFIUS successfully isolates the sensitive components, access restrictions work better than asset exclusion.

Section 2 argued that a buyer who values the capabilities of a complete system where artificial intelligence is a critical piece would object to a mitigation that removes the artificial intelligence from the transaction. Access restrictions do not raise the same concern because they allow the complete system to remain intact. They therefore create a possible avenue to acceptable mitigations in scenarios where CFIUS's concerns can be addressed by, for example, limiting access to U.S. citizens.

Overall, access restrictions might be more appropriate for artificial intelligence transactions than either integrity assurance or asset exclusion. Though they do not resolve all issues, access restrictions are more appropriate when a system is holistic and hard to segment into discrete components.

C. Need for New Mitigation Options

While the mitigations revealed in the CFIUS annual reports largely remain static from year to year, new mitigations could—and should—be used if the circumstance requires. The above discussion illuminates both the failure of traditional mitigations and some of the unique traits of artificial intelligence that make it difficult to police either during development or after deployment. The non-deterministic outputs and injection of bias described in Section IV.B.1 and the difficulty of locating the specific components that drive outcomes described in Section IV.B.2 create issues for traditional mitigations.

These difficult-to-manage traits exist in all artificial intelligence systems and are not limited to those under CFIUS's scrutiny. Commenters have raised concerns around discrimination from bias, protection of Fourth Amendment rights in the face of unexplainable algorithms, and the impact on Due Process of inaccurate artificial intelligence.⁷² The solutions presented to

⁷² See, e.g., David Lehr and Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn about Machine Learning*, 51 U.C.D. L. REV. 653, 658, 662, 664, 703-705 (2017) (discussing current legal scholarship regarding automated suspicion algorithms and the Fourth Amendment, due process for automated predictions, and the disparate impact of big data).

these problems, though discussed as accountable or ethical principles⁷³ or guidelines,⁷⁴ may represent the foundation of a new class of mitigations for transactions.

In forming new principled mitigations, CFIUS and transacting companies should consider two objectives from an ethical framework to guide new families of mitigations that address artificial intelligence and national security. The first is artificial intelligence robustness assurances.

A robust artificial intelligence makes decisions that are both accurate and reproducible.⁷⁵ Accuracy pertains both to a system's ability to make correct judgments and to its ability to display an error rate lower than a predetermined acceptable level.⁷⁶ Error rates are useful because, unlike in traditional software, it is not always possible to explain why an artificial intelligence system generated a particular output.⁷⁷ This means that testing an artificial intelligence system can reveal the frequency of errors but not always the cause. Closely related to accuracy is reproducibility. An artificial intelligence system that displays perfect reproducibility generates matching outputs in initial and subsequent runs while inputs remain the same.⁷⁸ An input pattern that always causes an undesirable or incorrect output, for example, can be documented even though it remains unclear why the system acts undesirably in the particular circumstance. Together, accuracy and reproducibility describe a system's ability to meet objectives, the rate at which it does so, and the conditions under which it succeeds and fails. Robustness assurance mitigations that impose high accuracy targets, low error rate limits, and require reproducible results can be enforced and will therefore help ensure developers proceed conservatively and test to stay within compliance boundaries.

A second class of objectives to consider are artificial intelligence transparency assurances. To assure transparency, companies should adopt (or be required to adopt by CFIUS) audit-friendly practices, including explainability and traceability. Explainability is the ability for the developer or deployer to explain decisions made by the artificial intelligence to an outside auditor.⁷⁹ Traceability is the ability to follow a trail backwards to the

⁷³ See *Principles for Accountable Algorithms and a Social Impact Statement for Algorithms*, FAT/ML, <http://www.fatml.org/resources/principles-for-accountable-algorithms> (accessed Nov. 1, 2019) [<https://perma.cc/VB98-P6Y9>].

⁷⁴ See HIGH-LEVEL EXPERT GROUP AI, EUR. COMM., *ETHICS GUIDELINES FOR TRUSTWORTHY AI* (2019).

⁷⁵ *Id.* at 17.

⁷⁶ *Id.*

⁷⁷ *Id.* at 13.

⁷⁸ *Id.* at 17.

⁷⁹ *Id.* at 18.

root of the decision.⁸⁰ An artificial intelligence system developed in accordance with these objectives would be easier to investigate if its integrity was questioned. The traits of explainability and traceability also aid the use of access restrictions by helping to isolate sensitive components.

These two families of objectives are merely examples of the kinds of mitigations that CFIUS might derive from existing frameworks for ethical artificial intelligence. Ethical frameworks address problems caused by artificial intelligence's unique characteristics and can aid CFIUS and artificial intelligence companies in doing the same.

V. CONCLUSION

As foreign investment into U.S. artificial intelligence companies continues to rise, so will the national security concerns it brings. Congress granted CFIUS new and expansive powers to push back against investments that have the potential to undermine national security. For artificial intelligence companies courting foreign investment and for foreign investors looking to acquire U.S. know-how, the stakes have been raised. Furthermore, when artificial intelligence transactions raise national security concerns, CFIUS must identify new mitigations to cope with the unique difficulties posed by the new technology. CFIUS and transacting companies should look to existing ethical, accountable, and accuracy frameworks to create a new family of mitigations that addresses the unique characteristics of artificial intelligence.

⁸⁰ *Id.*

TECHNOLOGY EXPLAINERS

SEARCH ENGINE OPTIMIZATION: WHAT WE SEE AND WHY WE SEE IT

Joseph Baillargeon *

CITE AS: 4 GEO. L. TECH. REV. 299 (2019)

TABLE OF CONTENTS

I. INTRODUCTION: WHAT IS SEARCH ENGINE OPTIMIZATION AND WHY IS IT IMPORTANT.....	299
II. CRAWLER SEARCH ENGINES: THE MECHANICS OF “GOOGLING”	301
III. GETTING RESULTS: ON AND OFF-PAGE SEO.....	303
IV. CONCLUSION	305

I. INTRODUCTION: WHAT IS SEARCH ENGINE OPTIMIZATION AND WHY IS IT IMPORTANT

Search Engine Optimization (SEO) is a technique to improve the ranking of a website on the results page of a search engine for the purpose of increasing the traffic to that site.¹ This technique has become crucial in the modern Internet era, where standing out amongst the over 1.7 billion websites can be a challenge.² This challenge derives from the fact that many of these websites are competing for the top spots of the same search queries. While such a competition may not seem that important, its outcome can actually be pivotal to the success of an online business. The unfortunate reality is that only about thirty-seven percent of online shoppers look past the top three results.³ These top three search positions have come to be known as “the golden triangle” and are highly sought after.⁴

* Georgetown University Law Center, J.D. Candidate 2022; Virginia Polytechnic Institute, B.S. Engineering Science and Mechanics, B.A. Physics. Thank you to the wonderful editors of GLTR, without whom this piece would not have been possible.

¹ Dushyant Sharma et al., *A Brief Review on Search Engine Optimization, in CONFLUENCE 2019: 9TH INT’L CONF. ON CLOUD COMPUT., DATA SCI. & ENG’G* 687, 687 (2019).

² INTERNET LIVE STATS, <https://www.internetlivestats.com/> (website live tracks total number of websites and constantly updates total) (accessed Nov. 22, 2019) [<https://perma.cc/ZR7A-BA8L>].

³ See Sharma et al., *supra* note 1, at 689.

⁴ *Id.*

The importance of this behavioral phenomenon is only increasing with time as e-commerce continues to play a larger and larger role in the economy. Studies have shown that in today's marketplace sixty-one percent of Internet users research products online and forty-four percent of online shoppers use a search engine to begin the search for a seller.⁵ Furthermore, of that forty-four percent of online shoppers, seventy-five percent never even click past the first Search Engine Results Page (SERP).⁶ Companies like Google do provide services, such as AdWords, which allow a website to be placed at the top of certain relevant searches without having to worry about SEO.⁷ However, studies have shown that seventy percent of the links that online shoppers click on are organic—meaning the links were recommended by Google's algorithm.⁸

The behavioral impact of search engine rankings is significant and certainly not limited to economics thereto. A study in India used a mock search engine to intentionally return biased results to its participants, such that the first SERP only exposed them to the positive news articles of a target candidate when that candidate was searched for.⁹ The study concluded it would be relatively easy to persuade around twenty percent of undecided voters to a target candidate, merely by changing what the voter saw in his or her search results.¹⁰

What an individual sees on the Internet is significantly determinative of what that individual will buy and how the individual will think. It has therefore become imperative to understand the algorithms that dictate what we see every time we “Google” something so we can appreciate their fallibility. Part II of this paper will help accomplish this by outlining the basics of how a search engine works. Subsequently, Part III will focus on some of the ways webmasters use search engines to gain a competitive edge through SEO. Finally, Part IV will conclude this paper by reiterating the importance of understanding these topics in a modern context.

⁵ Venkat N. Gudivada et al., *Understanding Search Engine Optimization*, COMPUTER, Oct. 2015, at 43, 43 (2015).

⁶ *Id.*

⁷ Aranyak Mehta et al., *AdWords and Generalized On-Line Matching*, in 46TH ANNUAL IEEE SYMPOSIUM ON FOUNDATION OF COMPUTER SCIENCE 1, 1 (2005).

⁸ Gudivada et al., *supra* note 5, at 43.

⁹ Robert Epstein & Ronald E. Robertson, *The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections*, 112 PNAS 1, 6 (2015).

¹⁰ *Id.* at 9.

II. CRAWLER SEARCH ENGINES: THE MECHANICS OF “GOOGLING”

Most people are accustomed to using crawler-based search engines, which include search engines like Bing, Yahoo!, and Google.¹¹ This paper will focus on Google, as Google handles around seventy-five percent of Internet searches.¹² In general, the crawler-based search engine takes four major steps to try to find and display what a user is looking for: crawling, indexing, calculating relevancy, and retrieving the results.¹³ First, the search engine must run a large number of computer programs, known as bots, that scour the Internet for as many webpages as they can find.¹⁴ Next the search engine stores a copy of the webpages into a giant document database, which Google calls Caffeine.¹⁵ Once the database of webpages has been created or updated, the contents of those websites are formatted to remove unnecessary information and keywords and phrases are extracted.¹⁶ The document database then uses these keywords and phrases to create indexes to be later queried by the user.¹⁷ Finally, when the user enters a keyword or phrase into the browser, the text string is transformed into its canonical form—a form that allows the search engine to take into account misspellings, capitalization, related words, noise words, etc.—and the Google search engine searches the indexed document database, Caffeine, for this canonical form, returning the most relevant URLs using Google’s ranking algorithms.¹⁸

The ranking algorithm is one of the most important aspects of the search engine. Since its inception, Google has been working diligently to perfect this algorithm by improving its accuracy. The original Google ranking algorithm, called PageRank, ranks webpages based on: (1) the number of outgoing links on a page, the number of backlinks, or pages that link into that

¹¹ Sharma et al., *supra* note 1, at 687.

¹² Dave Davies, *Meet the 7 Most Popular Search Engines in the World*, SEARCH ENGINE J. (Jan. 7, 2018), <https://www.searchenginejournal.com/seo-101/meet-search-engines/#close> [<https://perma.cc/8CB3-AXJ6>].

¹³ Sharma et al., *supra* note 1, 687–88.

¹⁴ Gudivada et al., *supra* note 5, at 44.

¹⁵ Sharma et al., *supra* note 1, at 687; Dave Davies, *Google’s Caffeine Update: Better Indexing & Fresher Search Results*, SEARCH ENGINE J. (Nov. 20, 2017), <https://www.searchenginejournal.com/google-algorithm-history/caffeine-update/#close> [<https://perma.cc/8CB3-AXJ6>].

¹⁶ Gudivada et al., *supra* note 5, at 44.

¹⁷ *Id.*

¹⁸ Crosby Grant, *Canonical Form: The Hidden Keywords in Paid Search*, SEARCH ENGINE LAND (Dec. 26, 2011), <https://searchengineland.com/canonical-form-the-hidden-keywords-in-paid-search-100603> [<https://perma.cc/H7JR-DUUE>]; Gudivada et al., *supra* note 65, at 44.

page; and (2) the quality of those links.¹⁹ This method allows the internet to “vote” on which webpage it deems the most relevant.²⁰ Once people realized this, however, artificial rank inflation coincided, as people exploited the algorithm by doing things like loading their websites with hidden links in text colors that matched the backgrounds.²¹ These tactics undermine the quality of the results returned by the search engine and are known as “black hat” SEO practices because they are contrary to the Google guidelines for creating websites.²² The competing interests of search engines seeking to produce high quality results and companies seeking placement at the top of the SERP have created an arms race that has proliferated the number and elevated the quality of Google’s various ranking algorithms.

Some such ranking algorithms include: Panda, Penguin, the Pirate, Payday Loan, Hummingbird, Pigeon, Possum, and Fred. Panda is designed to evaluate the quality of information on a website—such as having too little or too much information, poor format, grammar error, spelling mistakes, unreliable information, or low-quality content.²³ Penguin checks for contextual page linking (links that are surrounded by text), and determines if the links are from trustworthy sources.²⁴ The Pirate algorithm blocks or de-ranks sites that have received several reports for copyright infringement.²⁵ Payday Loan filters out pornography, casino, and high interest loan sites.²⁶ Hummingbird attempts to interpret the user’s intent when they input a keyword or phrase into the browser.²⁷ Pigeon finds the user’s location and alters ranking based on the location of companies in the user’s area.²⁸ Possum improves the ranking of websites or businesses in the top position of the search result to coincide with those nearest to the location of the searcher.²⁹ Finally,

¹⁹ Ian Rogers, *The Google Pagerank Algorithm and How it Works*, PRINCETON UNIV., <https://www.cs.princeton.edu/~chazelle/courses/BIB/pagerank.htm> [<https://perma.cc/7WAK-7WQJ>].

²⁰ *Id.*

²¹ Sharma et al., *supra* note 1, at 688.

²² *Id.*

²³ Abuzar Khan, *Top 12 Google Algorithm Updates You Need to Know in 2019*, SEO BASICS (May 29, 2019), <https://www.seobasics.net/google-algorithm-updates> [<https://perma.cc/5RP3-9XRE>].

²⁴ *Id.* at §§ 6–7.

²⁵ *Id.* at § 8.

²⁶ *Id.* at §§ 10–11.

²⁷ *Id.* at §§ 11–12.

²⁸ *Id.* at §§ 13–15.

²⁹ *Id.*

Fred looks for excessive ads, low-value content, and websites that offer very little user benefit.³⁰

The aggregate effect of these algorithms makes sure that users get the highest quality information relevant to what they are looking for. These algorithms take into account over 200 factors³¹ to accomplish their respective functions, which is where legitimate, “white hat” SEO practices come into play. SEO is not about gaming the system, but rather, optimizing a website to help search engines return the most pertinent results to the user.³²

III. GETTING RESULTS: ON AND OFF-PAGE SEO

SEO is not an exact science. It requires quite a bit of guess work because Google’s ranking algorithms are constantly being updated and tested, and their exact algorithms are guarded as a trade secret.³³ Google commonly sequesters a small subsection of the Internet to test new algorithms, so it is possible for two people using Google at the same time to be using different algorithms.³⁴ The SEO community is left to rely on a mishmash of vague guidance and second hand sources to determine best practices. One of the best sources comes from Google itself, which publishes its own guidelines on how to design a website with SEO in mind.³⁵ Additionally, because Google patents some portions of its algorithms, those portions are publicly available and provide a broad overview of what is going on behind the scenes.³⁶ But outside of what Google chooses to disclose, the only other reliable SEO strategies come from trial and error and approximations (using mathematical and machine learning techniques).³⁷ The information from all these different

³⁰ Aleh Barysevich, *A Cheat Sheet to Google Algorithm Updates from 2011 to 2018*, SEARCH ENGINE WATCH (Oct. 10, 2018), <https://www.searchenginewatch.com/2018/10/10/a-cheat-sheet-to-google-algorithm-updates-from-2011-to-2018/> [<https://perma.cc/74WJ-UY6J>].

³¹ Gudivada et al., *supra* note 5, at 43.

³² *Id.* at 46.

³³ Kristine Forderer, *Trade Secrets: A Valuable Tool in Your IP Protection Strategy*, COOLEY GO (Feb. 4, 2019), <https://www.cooleygo.com/trade-secrets-a-valuable-tool-in-your-ip-protection-strategy> [<https://perma.cc/HU33-PSNP>]; Steven Levy, *Exclusive: How Google’s Algorithm Rules the Web*, WIRED (Feb. 22, 2010), https://www.wired.com/2010/02/ff_google_algorithm/ [<https://perma.cc/KU2L-XD7L>].

³⁴ Levy, *supra* note 33, at 687.

³⁵ GOOGLE, SEARCH ENGINE OPTIMIZATION STARTER GUIDE (2010), <http://static.googleusercontent.com/media/www.google.com/en/us/webmasters/docs/search-engine-optimization-starter-guide.pdf> [<https://perma.cc/RDK3-NRP4>].

³⁶ Lawrence Page, US Patent 6,285,999 B1, <https://patentimages.storage.googleapis.com/37/a9/18/d7c46ea42c4b05/US6285999.pdf> [<https://perma.cc/3M2B-NA26>].

³⁷ Hengameh Banaei & Ali Reza Honarvar, *Web Page Rank Estimation in Search Engine Based on SEO Parameters Using Machine Learning Techniques*, 17 INT’L J. COMP. SCI. &

sources has culminated into a plethora of SEO strategies that are often divided up into two distinct groups: (1) on-page, which refers to the actual content and code the makes up a person's website; and (2) off-page, which refers to all of the information about your website that is not contained on your website.³⁸ The strategies in each of these two groups can be further categorized as either a white hat or black hat strategy.³⁹

A crucial first step in implementing white hat, on-page SEO is to conduct "keyword research."⁴⁰ Keyword research is the process of finding commonly searched phrases and topics whose Google search results would be best to appear in.⁴¹ For instance, if someone had a Mexican restaurant in Washington DC, they might want the restaurant's webpage to be the top result in a search for, "best Mexican restaurants in DC." Google provides tools to determine the competitiveness and frequency of use for these searched phrases,⁴² to help decide which searches would be best to get highly ranked for. Once a phrase has been decided on, adjustments can be made to the webpage, including: the title tag, header tags, the URL, and the content of the webpage, such that they all gravitate around the target phrase.⁴³ It is ideal for the keyword to occur around every 5 to 7 words per 100 words.⁴⁴ Other things the Google algorithms are known to look at are: organic looking anchor text (text that is also a URL), a custom 404 page (what happens when you link to a page that does not exist), and a custom privacy policy.⁴⁵

Examples of off-page SEO techniques include things like placing backlinks to your website on other high-quality websites via blog comments⁴⁶ and social media sites like LinkedIn and Twitter.⁴⁷ This strategy is referred to as Link Building.⁴⁸ Another crucial off-page SEO technique is ensuring that Google has an accurate sitemap (information about and structure of every

NETWORK SEC. 95, 96 (2017) (discussing the extensive research that he been done since 2005 to computationally approximate web ranking),

http://paper.ijcsns.org/07_book/201705/20170513.pdf [<https://perma.cc/LC7J-NL6G>].

³⁸ Sharma et al., *supra* note 1, at 688–89.

³⁹ Gudivada et al., *supra* note 5, at 44.

⁴⁰ Sharma et al., *supra* note 1, at 688.

⁴¹ *Id.*

⁴² Google, *How Keyword Planner Works*, GOOGLE ADS, https://ads.google.com/intl/en_en/home/tools/keyword-planner/ [<https://perma.cc/9G7Y-DS39>].

⁴³ Sharma et al., *supra* note 1, at 688.

⁴⁴ Muhammad Naeem Ahmed Khan & Azhar Mahmood, *A Distinctive Approach to Obtain Higher Page Rank Through Search Engine Optimization*, SĀDHANĀ 1, 3 (Mar. 2018).

⁴⁵ Gudivada et al., *supra* note 5, at 48–49.

⁴⁶ Sharma et al., *supra* note 1, at 689.

⁴⁷ Khan & Mahmood, *supra* note 44, at 8.

⁴⁸ *Id.*

webpage on a website) for the website.⁴⁹ When Google sends its crawler bot to a website to be copied, it creates a sitemap as it traverses all of the links on the page.⁵⁰ By submitting its own sitemap directly to Google, a website can ensure that all of its webpages have been properly indexed and can be found through the search engine.⁵¹

Unfortunately, as with other automated systems, once the underlying algorithms are dissected, the system can be exploited. Exploiting these ranking algorithms to raise a website's rank is known as black hat SEO. These practices include things like keyword stuffing, where the website is filled with many different, often unrelated, keywords to increase the number of search engine queries for which a website is returned.⁵² Cloaking is a method where hackers cause the crawler bots and the users to see different content, usually to hide malicious content.⁵³ Hackers also often write computer code that causes automatic webpage generation, copying content word for word from well-known sites in attempts to rank for certain keywords.⁵⁴ These underhanded methods end up costing the businesses they displace an estimated \$130 billion annually.⁵⁵

IV. CONCLUSION

SEO is something that cannot be ignored by those who have or want to have an online presence. Until human behavior changes and people begin consistently looking past the first few search results, one of the best ways to increase traffic to a website is making sure the search engine algorithms are working in the site's favor. Search engines, like Google, have been increasing their accuracy by leaps and bounds over the past decade. However, at the end of the day these algorithms are still, and probably will always be, imperfect systems that are based off of imperfect metrics. Unless a website invests, at least minimally, in SEO, it is vulnerable to outranking by savvy webmasters with lower quality content who aspire to make a quick buck. The implications of this go far beyond just e-commerce and have profound effects on how the public is informed. Misinformation and disinformation can easily be ranked higher than accurate information.⁵⁶ Therefore, it is—now more than ever—

⁴⁹ Gudivada et al., *supra* note 5, at 50.

⁵⁰ *Id.*

⁵¹ *Id.*

⁵² Sharma et al., *supra* note 1, at 688.

⁵³ Gudivada et al., *supra* note 5, at 45.

⁵⁴ Gudivada et al., *supra* note 5, at 45.

⁵⁵ *Id.* at 44.

⁵⁶ STACIE HOFFMAN ET AL., OXFORD TECH. & ELECTIONS COMM'N, THE MARKET OF DISINFORMATION 11 (2019).

crucial to inoculate the populous to these manipulations by making sure they understand the fallibility and limitations that underlie their search results.

SOCIAL SPAMBOTS

Richard Bernache*

CITE AS: 4 GEO. L. TECH. REV. 307 (2019)

TABLE OF CONTENTS

I. INTRODUCTION: WHAT IS A SPAMBOT?	307
II. TYPES OF SPAMBOTS	308
A. Email Spambots	308
B. Social Media Spambots.....	308
C. Cross Platform Spam Attacks	310
III. EVOLUTION OF SPAMBOTS	311
IV. SPAMBOT DETECTION AND PREVENTION—CAPTCHA TESTS	312
V. CONCLUSION.....	314

I. INTRODUCTION: WHAT IS A SPAMBOT?

Spam, unsolicited commercial content spread at a mass scale, is everywhere on the Internet, in our email and on the sites we visit.¹ At a minimum, spam is a nuisance, just another email or post online to delete or ignore. At its worst, falling victim to a spam attack risks infecting a system with malware or compromising an individual's identity. This proliferation is due in large part to spambots, Internet robots designed to spread spam autonomously. Spambots can be programmed in any number of modern coding languages and are created for a variety of purposes including the mass spreading of advertising content, proliferation of credible or false information,

* Georgetown University Law Center, J.D. Candidate 2021. Saint Michael's College, B.A. in Theatre 2016. First and foremost, I would like to thank the staff and editors of the Georgetown Law Technology Review for their work and patience during the writing process. Thanks also to my friends and family for their unyielding support while I undertake this new part of my professional life..

¹ Pedram Hayati et al., *Characterisation of Web Spambots Using Self Organising Maps*, 2 INT'L J. COMPUTER SCI. & ENGINEERING 87, 87–88 (2011). [<https://perma.cc/83HH-7267>].

or the spread of malware and other harmful software.² This explainer focuses on two types of spambots that ordinary human users are likely to encounter: email spambots, which scrape the web looking for email addresses and then auto-send mail with malware, and social media spambots, which pose as human users and autogenerate content containing spam. Web developers have spent significant resources into developing spam prevention mechanisms. However, the most relied upon spambot prevention tool, the CAPTCHA, has not been able to keep up with the evolution of current spambots, leaving our email and social media accounts vulnerable.

II. TYPES OF SPAMBOTS

A. Email Spambots

Email was among the earliest targets of online spam campaigns. The dispersion of spam through email quickly proliferated, and now the majority of all email sent contains spam.³ The most basic types of email spam are not context specific, meaning they are not connected to the human target's personal history. These emails seek to either convince recipients to follow a link containing malware or to steal a human user's identity.⁴ In order to maximize their success, many email spammers target massive numbers of human users in the hope that even a relatively low number of recipients open and follow the malicious link.⁵ Deploying spambots to scrape websites for email addresses and generating and auto-sending spam emails is inexpensive.⁶ This allows spammers to target more users.

B. Social Media Spambots

While email spambots primarily spread malware, social media bots perform a number of different tasks. For example, Twitter has become a focus of the public conversation around spambots, in large part due to efforts by Russian intelligence agencies to spread misinformation during the 2016

² *Id.* at 88; Rob Dubbin, *The Rise of Twitter Bots*, NEW YORKER (Nov. 14, 2013), <https://www.newyorker.com/tech/annals-of-technology/the-rise-of-twitter-bots> [<https://perma.cc/X5KF-FQ92>]; Chencheng Shao et al., *The Spread of Low-credibility Content by Social Bots*, 9 NATURE COMMS. 1, 5 (2018), <https://www.nature.com/articles/s41467-018-06930-7.pdf> [<https://perma.cc/VW5Z-CWK9>].

³ Cristian Lumezanu & Nick Feamster, *Observing Common Spam in Tweets and Emails*, INTERNET MEASUREMENT CONF. PROC. (2012), <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.297.7011&rep=rep1&type=pdf> [<https://perma.cc/9RK7-BX58>].

⁴ *Id.*; Dubbin, *supra* note 2.

⁵ Lumezanu & Feamster, *supra* note 3.

⁶ Hayati et al., *supra* note 1, at 88.

United States presidential election.⁷ In total, Twitter identified 50,258 bot accounts linked to the Internet Research Agency (IRA), a Russian propaganda organization.⁸ These bots were used to spread false or misleading information regarding the U.S. election and ultimately approximately 1.4 million human users interacted with an IRA bot by liking, retweeting, replying to, or following the bot.⁹ These adversarial bot accounts were able to flourish on Twitter due to the website's open application program interface (API).¹⁰ Twitter allows companies and individuals to apply for access to Twitter's various APIs.¹¹ APIs connect a developer's computer program with Twitter's "endpoints" which are the various types of information (users, tweets, direct messages, and ads) that a developer is able to access via use of Twitter APIs.¹² Accessing Twitter's tweet and reply APIs allows a developer to collect tweets of a certain topic and program bots to post tweets via the API connection.¹³ Exploitation of Twitter's API by spammers caused Twitter to add additional steps to the API application process and the site has begun to more aggressively police violations of its spam policy.¹⁴ However, the effectiveness of these new measures has yet to be analyzed.

Additionally, due to advances in publicly available source codes for social media bots and advanced artificial intelligences (AIs) capable of conversational text, creating social media bots now requires less individual programming knowledge.¹⁵ These source codes are readily available on various tech blogs, allowing an unsophisticated human spammer to use spambots to spread malicious content.¹⁶ Simple spambots may post the same pre-scripted content written by the human spammer according to a programmed schedule.¹⁷ More sophisticated social media bots may integrate

⁷ *Update on Twitter's Review of the 2016 US Election*, TWITTER BLOG, (last updated Jan. 31, 2018), https://blog.twitter.com/official/en_us/topics/company/2018/2016-election-update.html [<https://perma.cc/XST3-WMRU>].

⁸ *Id.*

⁹ *Id.*

¹⁰ Dubbin, *supra* note 2.

¹¹ *About Twitter's API*, TWITTER, <https://help.twitter.com/en/rules-and-policies/twitter-api> (accessed Oct. 31, 2019) [<https://perma.cc/74WV-XP65>].

¹² *Id.*

¹³ *Id.*

¹⁴ Yoel Roth & Rob Johnson, *New Developer Requirements to Protect Our Platform*, TWITTER BLOG (July 24, 2018), https://blog.twitter.com/developer/en_us/topics/tools/2018/new-developer-requirements-to-protect-our-platform.html [<https://perma.cc/XUP6-QSL7>].

¹⁵ See generally Emilio Ferrara, Univ. S. Cal. Info. Sci. Inst., *Bots, Elections, and Social Media: A Brief Overview*, (arXiv:1910.01720v1, Oct. 3, 2019), <https://arxiv.org/pdf/1910.01720.pdf> [<https://perma.cc/QRX5-JC8V>].

¹⁶ *Id.*

¹⁷ Emilio Ferrara, *The History of Digital Spam*, 62 COMMS. ACM 82, 88 (Aug. 14, 2019), <https://arxiv.org/pdf/1908.06173.pdf> [<https://perma.cc/YX6X-AWJF>].

text generating AI such as those provided by companies like ChatBots.io in order to auto-generate conversational text and engage with and respond to real human users.¹⁸ These AIs allow social media bots to better mimic human user behavior, increasing the effectiveness of spam campaigns.¹⁹

While many Twitter bots are used for nefarious purposes, many early Twitter bots used Twitter's API solely for innocuous, often times comedic, purposes. One comedic bot, Pentametron, was programmed to roam Twitter searching for rhyming couplets, and retweets the rhyme to the bot's twenty-six thousand followers.²⁰ Other bots seek to spread malware or harvest and steal user identities just as email spambots do.²¹ Others can be purchased in order to boost a user's follower count to project a greater following than the individual actually possess. In the process, these bots follow large numbers of other human users in order to mask the fact that they are not actually human and to protect the identity of whoever purchased the follower boost.²²

C. Cross Platform Spam Attacks

The effectiveness of spambots increases when the spam is spread across multiple platforms.²³ Such cross-platform spam attacks can be characterized as "context aware spam."²⁴ By using bots to web-scrape social media and other online profiles of human users, spambots are able to generate content which is harder to identify as spam. For example, a spambot may scrape sufficient information from a user's Facebook profile to determine the user's friends, date of birth, and email address.²⁵ The bot is then able to use this information to create a spam email with a malware link that is context specific to the individual target. This could be synthesized into an email from a user's Facebook friend with a link to a birthday e-card.²⁶ By capitalizing on the user's personal data, the spammer is able to increase the number of users who follow the malicious link.²⁷

¹⁸ Ferrara, *supra* note 15, at 3.

¹⁹ Ferrara, *supra* note 17, at 84.

²⁰ Dubbin, *supra* note 2.

²¹ Alexis Madrigal, *Here's Why 9,000 Porny Spambots Descended on a High Schooler's Twitter Account*, ATLANTIC (Nov. 25, 2013), <https://www.theatlantic.com/technology/archive/2013/11/why-did-9-000-porny-spambots-descend-on-this-san-diego-high-schooler/281773> [<https://perma.cc/D8XQ-E4YY>].

²² *Id.*

²³ Lumezanu & Feamster, *supra* note 3.

²⁴ Garrett Brown et al., *Social Networks and Context-Aware Spam*, in *COMPUTER SUPPORTED COOPERATIVE WORK CONF.* 403 (2008), <http://www-personal.umich.edu/~kborders/p403-brown.pdf> [<https://perma.cc/3JL9-TD6U>].

²⁵ *Id.* at 407.

²⁶ *Id.*

²⁷ *Id.*

III. EVOLUTION OF SPAMBOTS

Early spambots were typically easily identified by human users because bots shared many common behavioral patterns which did not mirror human to human online communication. For example, early email spambots frequently misspelled or inserted random extra characters to commonly filtered words in order to avoid detection by spam filters.²⁸ Early social media spambots often posted the same advertisement or link incessantly and featured nonsensical usernames.²⁹

However, spambots have rapidly improved their behavior such that it has become significantly more difficult for human users to detect bot accounts on their own. Modern social media spambots have seen dramatic improvements in their ability to mimic human behavior. Spambots attempt to hide their presence by following or friending large numbers of real human users.³⁰ This tactic is made more effective when these real human users follow the spambot back, deceiving detection mechanisms which rely on targeting bots' tendency to have limited followers while generating a large amount of content.³¹ This technique has been improved by reducing the amount of content a bot produces each day in order to more closely approximate the typical human user's Twitter usage.³² Bots also focus on trending topics, such as popular songs and YouTube videos and generate tweets of their own which appear to follow online trends.³³

Identifying these more advanced bots typically requires the ability to observe bots as a group, which is often beyond the capabilities of a normal human user. For example, an examination of bots used in the 2014 mayoral election in Rome revealed that these bots were unlikely to be detected without the ability to track the common behaviors of a large group of bots.³⁴ These bots followed a large number of ordinary users, tweeted at a relatively infrequent pace to mimic common human user patterns, and normally tweeted

²⁸ Li Zhuang et al., *Characterizing Botnets from Email Spam Records*, in FIRST USENIX WORKSHOP ON LARGE-SCALE EXPLOITS & EMERGENT THREATS 3 (2008), http://static.usenix.org/events/leet08/tech/full_papers/zhuang/zhuang.pdf [<https://perma.cc/ZU6N-7P28>].

²⁹ Dubbin, *supra* note 2.

³⁰ Stefano Cresci et al., *On the Capability of Evolved Spambots to Evade Detection via Genetic Engineering*, 9 ONLINE SOC. NETWORKS & MEDIA 1, 4 (2019), <https://www.sciencedirect.com/science/article/pii/S246869641830065X> [<https://perma.cc/8QYK-RLQY>].

³¹ Xia Hu et al., *Online Social Spammer Detection*, in 28TH AAAI CONF. ARTIFICIAL INTELLIGENCE 59, 59 (2014), <https://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/viewFile/8467/8399> [<https://perma.cc/F673-9A47>].

³² Cresci et al., *supra* note 30 at 4.

³³ *Id.*

³⁴ *Id.*

about popular songs and media. However, when the candidate these bots were designed to promote tweeted, all bots retweeted the candidate's content within minutes of each other.³⁵ The ordinary user would be unable to determine this synchronized behavior unless they were able to observe the tweets of several bots. In fact, some human users actually replied to bot generated content during Rome's mayoral election.³⁶ In order to protect users from spam attacks, sites have long since used spam detection and prevention mechanisms, such as CAPTCHA, to stop the proliferation of spam to the site's users.³⁷ As spambots become increasingly more adept at deceiving the ordinary human users, these prevention mechanisms are even more crucial as tools to stop bots from reaching human users. However, even the most frequently used spam prevention tool, the CAPTCHA test, has become increasingly ineffective at preventing spambots from gaining access to sites and human users.

IV. SPAMBOT DETECTION AND PREVENTION—CAPTCHA TESTS

CAPTCHA, or Completely Automated Public Turing test to tell Computers and Humans Apart, is the most popular and well recognized anti-spambot tool.³⁸ CAPTCHAs are tests that a user must solve before accessing a webpage. CAPTCHAs come in different forms but all must be easily solvable by humans and easily generated, but must not be easily solved by bots.³⁹ Early CAPTCHAs were primarily text based and typically featured an image of a word or words, which were distorted through some combination of text warping or color overlay.⁴⁰ These images are decipherable by human users, who are able to read the text through the distortions, but unreadable by bots that rely optical character recognition (OCR) software to convert images to text.⁴¹ The OCR software would be unable to accurately decipher the text obscured by the CAPTCHA distortions. In 2014, Google announced its own

³⁵ *Id.*

³⁶ *Id.*

³⁷ Suphanee Sivakorn, *I'm Not a Human: Breaking the Google reCAPTCHA*, BLACK HAT (2016), <https://www.blackhat.com/docs/asia-16/materials/asia-16-Sivakorn-Im-Not-a-Human-Breaking-the-Google-reCAPTCHA-wp.pdf> [<https://perma.cc/L85W-E55P>].

³⁸ Hayati, *supra* note 1, at 89.

³⁹ Marti Motoyama et al., *Re: CAPTCHAs—Understanding CAPTCHA-Solving Services in an Economic Context*, in PROC. 19TH USENIX SEC. SYMPOSIUM 2 (2010), <http://cseweb.ucsd.edu/~savage/papers/UsenixSec10.pdf> [<https://perma.cc/Z7E5-PTSL>].

⁴⁰ Jeff Yan and Ahmad Salah El Ahmad, *Usability of CAPTCHAs or Usability Issues in CAPTCHA Designs*, in PROC. 4TH SYMPOSIUM ON USABLE PRIVACY & SEC. (2008), <https://prof-jeffyan.github.io/soups08.pdf> [<https://perma.cc/9CUJ-R9LM>].

⁴¹ Haley Tsukayama, *The Surprising Tool Bots Use to Get Around Those Pesky CAPTCHAs*, WASH. POST (June 9, 2016), <https://www.washingtonpost.com/news/the-switch/wp/2016/06/09/the-surprising-tool-bots-use-to-get-around-those-pesky-captchas/> [<https://perma.cc/LJ7M-CPHT>].

spambot prevention tool, “No CAPTCHA reCAPTCHA” (reCAPTCHA), in part because of the advances made by AI algorithms which were solving text-based CAPTCHAs with 99.8% accuracy.⁴²

Google’s reCAPTCHA has become the most popular CAPTCHA service.⁴³ reCAPTCHA functions by presenting a widget on a webpage that asks the user to certify “I’m not a robot” by clicking a box.⁴⁴ When the box is clicked, the user’s past usage behavior is automatically analyzed by a risk analysis algorithm to determine level of confidence that the user is human.⁴⁵ If this level of confidence is high, meaning the user is very likely to be human, the user is given permission to access the webpage without completing a CAPTCHA test.⁴⁶ While much of how this confidence level is estimated is proprietary information, research has shown that browsing history as determined through Google’s tracking cookie plays a critical role in determining the confidence level.⁴⁷

If the algorithm cannot determine with high confidence that the user is human, then the user is given either an image CAPTCHA or a text CAPTCHA. Image CAPTCHAs ask users to match images to a prompt.⁴⁸ For example, a user may be asked to “select all images which include wine.”⁴⁹ While human users are able to select the images that match the prompt, spambots are unable to accurately differentiate between the images generated by the program. Image CAPTCHAs are reCAPTCHA’s preferred form of test, and text CAPTCHAs have been gradually phased out.⁵⁰

However, Google’s reCAPTCHA has not been entirely effective at preventing bots from accessing sites. reCAPTCHA initially reused images frequently, allowing human users to collect and create databases of “tags” which can be used to program spambots to correctly select the prompted images of a reCAPTCHA.⁵¹ Image annotation services, such as Google Reverse Image Search, can generate lists of tags which describe the image, making the process of generating tags more efficient.⁵² Afterwards, the bots can be programmed to match the prompt of the reCAPTCHA (for example,

⁴² *Are You a Robot? Introducing “No CAPTCHA reCAPTCHA,”* GOOGLE SECURITY BLOG (Dec. 3, 2014), <https://security.googleblog.com/2014/12/are-you-robot-introducing-no-captcha.html> [<https://perma.cc/EA5S-3VH7>].

⁴³ Sivakorn, *supra* note 37.

⁴⁴ *Are You a Robot? Introducing “No CAPTCHA reCAPTCHA,”* *supra* note 41.

⁴⁵ *Id.*

⁴⁶ Sivakorn, *supra* note 37.

⁴⁷ *Id.*

⁴⁸ *Are You a Robot? Introducing “No CAPTCHA reCAPTCHA,”* *supra* note 41.

⁴⁹ Sivakorn, *supra* note 37.

⁵⁰ *Id.*

⁵¹ *Id.*

⁵² *Id.*

select all images which include wine) tags listed for each reused image and correctly select the matching images.⁵³ One such experiment, conducted in 2016, was able to use bots to solve image CAPTCHAs with seventy percent accuracy.⁵⁴

Even when spammers are unable to generate large image and tag databases on their own, they have been able to navigate around CAPTCHAs by employing human workers to solve text and image CAPTCHAs. These bad actors hire CAPTCHA solving services that employ workers to solve CAPTCHAs for the entirety of their workday.⁵⁵ These human workers and spambots are connected through a plug-in distributed by the CAPTCHA solving service.⁵⁶ When the bot encounters a CAPTCHA, the test is sent to a human user who solves it. The solution is then communicated back to the bot, which enters the solution and gains access to the site.⁵⁷ Integrating human solvers into spambots in this way thwarts the fundamental purpose of CAPTCHAs, which were designed to allow humans, but not the spam spreading bots, to access the sites.⁵⁸

V. CONCLUSION

Spambots are rapidly evolving and remain ever present on online platforms. As these bots evolve to outpace prevention mechanisms, the sheer amount of spam content online is likely to increase. Additionally, as these spam techniques become more effective in deceiving ordinary human users, more and more users may fall victim to malware attacks, identify theft, or disinformation campaigns. In order to keep pace with developments in spambot technology, prevention mechanisms must continue to evolve.

⁵³ *Id.*

⁵⁴ *Id.*

⁵⁵ Christopher Mims, *How Spammers Use Low-cost Labor to Solve CAPTCHAs*, MIT TECH. REV. (Aug. 11, 2010), <https://www.technologyreview.com/s/420200/how-spammers-use-low-cost-labor-to-solve-captchas/> [<https://perma.cc/9CWS-XQ5P>].

⁵⁶ Motoyama et al., *supra* note 39, at 6.

⁵⁷ *Id.*

⁵⁸ *Id.* at 5.

DATA BREACHES

Drew Diedrich *

CITE AS: 4 GEO. L. TECH. REV. 315 (2019)

TABLE OF CONTENTS

I. INTRODUCTION: DATA BREACHES IN MODERN SOCIETY	315
II. WHAT IS A DATA BREACH, AND HOW DOES IT OCCUR?	316
A. Data Breaches from Human Error	317
B. Data Breaches from System Glitches	318
C. Malicious or Criminal Data Breaches.....	319
1. <i>Malware Infections</i>	319
2. <i>Hacking</i>	321
3. <i>Social Engineering</i>	322
III. WHY MALICIOUS ACTORS CONDUCT DATA BREACHES	322
A. Financial Gain.....	323
B. Espionage.....	323
IV. CONCLUSION.....	324

I. INTRODUCTION: DATA BREACHES IN MODERN SOCIETY

To say that data breaches are ubiquitous within the modern digital society may be an understatement.¹ Over the past fifteen years, more than ten billion records have been breached from over 9,000 data breaches in the United States, impacting a majority of Americans.² In the first three months of 2019 alone, there were twenty data breaches that each exposed over ten million

* Georgetown University Law Center, J.D. Candidate 2020; Cornell University, B.A. Government, 2017. Thank you to all the editors on the Georgetown Law Technology Review for their help and advice.

¹ See *Roughly Half of Americans Do Not Trust the Federal Government or Social Media Sites to Protect Their Data*, PEW RES. CENTER (Jan. 23, 2017), https://www.pewinternet.org/2017/01/26/americans-and-cybersecurity/pi_01-26-cyber-00-02 [<https://perma.cc/7EUT-HAKW>].

² Aaron Smith, *Americans and Cybersecurity*, PEW RES. CENTER (Jan. 26, 2017), <https://www.pewinternet.org/2017/01/26/americans-and-cybersecurity> [<https://perma.cc/243G-VCQW>]; *Data Breaches*, PRIVACY RIGHTS CLEARINGHOUSE, <https://privacyrights.org/data-breaches> (accessed Oct. 20, 2019) [<https://perma.cc/AE3L-G3U4>].

records.³ Data breaches have hit all sectors and industries, from businesses and medical organizations to governments and educational entities.⁴ Within the past year, Dunkin' Donuts, Facebook, the Atlanta Hawks, Uniqlo, and the Federal Emergency Management Agency (FEMA) were some of the many victims.⁵ Even law firms have been targeted.⁶ In response to the prevalence of breaches, legal clients increasingly demand work in the realm of data breaches.⁷ A working knowledge of how they occur may prove useful to many attorneys.

II. WHAT IS A DATA BREACH, AND HOW DOES IT OCCUR?

Data breaches are not consistently defined across jurisdictions. Federal legislation for specific sectors, regulations from agencies, state statutes, and even international bodies establish rules governing data breaches, resulting in a variety of data breach definitions.⁸ Further, the private data security sector

³ RISK BASED SEC., INC., DATA BREACH QUICKVIEW REPORT-FIRST QUARTER 2019 DATA BREACH TRENDS 3 (Apr. 30, 2019), <https://pages.riskbasedsecurity.com/hubfs/Reports/2019/2019%20Q1%20Data%20Breach%20QuickView%20Report.pdf> [<https://perma.cc/JUV4-G8M4>].

⁴ *Id.* at 1, 10.

⁵ Steve Turner, *2019 Data Breaches—The Worst so Far*, IDENTITY FORCE (Jan. 3 2019), <https://www.identityforce.com/blog/2019-data-breaches> [<https://perma.cc/9WKK-2WUV>].

⁶ Debra C. Weiss, *More Than 100 Law Firms Have Reported Data Breaches; 2 BigLaw Firms Affected*, ABA J. (Oct. 18, 2019), <http://www.abajournal.com/news/article/more-than-100-law-firms-have-reported-data-breaches-2-biglaw-firms-affected> [<https://perma.cc/53VA-L7CX>].

⁷ *Five Legal Trends to Watch in 2019*, LEXISNEXIS (Feb. 28, 2019), <https://www.lexisnexis.com/community/lexis-legal-advantage/b/trends/posts/five-legal-trends-to-watch-in-2019> [<https://perma.cc/H9CX-HCLY>]; see Monique C.M. Leahy, *Litigation of Data Breach*, 14 AM. JUR. TRIALS 327 § I(1), para. 7 (2015); Chris Cwalina et al., *Nine States Pass New and Expanded Data Breach Notification Laws*, NORTON ROSE FULBRIGHT (Jun. 27, 2019), <https://www.dataprotectionreport.com/2019/06/nine-states-pass-new-and-expanded-data-breach-notification-laws/> [<https://perma.cc/T8L7-485G>].

⁸ Regulation 2016/679, art. 4, 2016 O.J. (L 119) 33 (EU), <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2016:119:FULL> [<https://perma.cc/8DD4-WUYK>]; FEDERAL DEPOSIT INSURANCE CORP., FIL-27-2005, FINAL GUIDANCE ON RESPONSE PROGRAMS FOR UNAUTHORIZED ACCESS TO CUSTOMER INFORMATION AND CUSTOMER NOTICE (2005), <https://www.fdic.gov/news/news/financial/2005/fil2705.pdf> [<https://perma.cc/A3FL-ZX88>]; *State Data Breach Notification Laws*, FOLEY & LARDNER LLP, (Nov. 11, 2019), <https://www.foley.com/-/media/files/insights/publications/2019/11/19mc23532-data-breach-chart-update-101419.pdf> [<https://perma.cc/RN6E-4QAN>]; Dave Maxfield & Bill Latham, *Data Breaches: Perspectives from Both Sides of the Wall*, 25 S.C. LAW. 28, 30 (2014),

creates its own definitions while combatting data breaches.⁹ Generally, these sources define a data breach as an unintentional release of secure or personally identifiable information to an unsecure environment.¹⁰

This personally identifiable information is data that is attributable and identifiable to one person.¹¹ Email addresses and passwords are most frequently stolen, but other frequent targets include credit card numbers and financial information.¹² Separate from personally identifiable information, data breaches may result in the exposure of sensitive corporate information, such as business-sensitive trade secrets.¹³ The annual “Cost of A Data Breach” study conducted by the Ponemon Institute estimated that roughly one quarter of data breaches in 2019 resulted from human error (“breaches caused by neglect or error by a person”), one quarter followed system glitches (“breaches caused by technology failures”), and the remaining half were caused by malicious or criminal attacks.¹⁴

A. Data Breaches from Human Error

Human error, or “an error caused by neglect or error by a person,” that results in a data breach may be frustrating as it appears to be the most

⁹ *Data Breach*, TREND MICRO (2019), <https://www.trendmicro.com/vinfo/us/security/definition/data-breach>, (accessed Oct. 19, 2019) [<https://perma.cc/J45F-J48P>]; Steve Symanovich, *What is a Data Breach?*, NORTON LIFELOCK, <https://www.lifelock.com/learn-data-breaches-data-breaches-need-to-know.html> (accessed Nov. 20, 2019) [<https://perma.cc/C46V-M3RV>]; *What is a Data Breach*, KASPERSKY (2019), <https://usa.kaspersky.com/resource-center/definitions/data-breach> [<https://perma.cc/J55Y-H785>].

¹⁰ *State Data Breach Notification Laws*, *supra* note 8; Maxfield & Latham *supra* note 8, at 30.

¹¹ *What is Personally Identifiable Information (PII)?*, NORTON LIFELOCK, <https://www.lifelock.com/learn-identity-theft-resources-what-is-personally-identifiable-information.html> [<https://perma.cc/2QGG-YPQ6>]; *see State Data Breach Notification Laws*, *supra* note 8.

¹² RISK BASED SEC., INC., *supra* note 3, at 9; *What Is a Data Breach*, KASPERSKY (2019), <https://usa.kaspersky.com/resource-center/definitions/data-breach> [<https://perma.cc/J55Y-H785>].

¹³ RISK BASED SEC., INC., *supra* note 3, at 9.

¹⁴ IBM SECURITY, COST OF A DATA BREACH REPORT 2019 at 30 https://www.ibm.com/downloads/cas/ZBZLY7KL?_ga=2.249958785.687410232.1571437539-250560343.1569949773&_gac=1.16575106.1571437539.EA1aIQobChMI_rqhG02m5QIVQpyzCh2yBw32EAAYASAAEgLZ9fD_BwE (accessed Oct. 19, 2019) [<https://perma.cc/R2N6-SUBA>]; Larry Ponemon, *What's New in the 2019 Cost of a Data Breach Report*, SECURITYINTELLIGENCE (Jul. 23, 2019), <https://securityintelligence.com/posts/whats-new-in-the-2019-cost-of-a-data-breach-report> [<https://perma.cc/GV7T-2MJA>].

preventable.¹⁵ Illustrative examples include failing to fix known vulnerabilities, mistakenly emailing the wrong person sensitive information, and unintentionally storing information on a public database.¹⁶ Often, data breaches like these result from something as simple as a school administrator accidentally publishing student medical records on the school's intranet.¹⁷

B. Data Breaches from System Glitches

System glitches resulting in data breaches, or “breaches caused by technology failure,” are the result of purely system issues and are not directly connected to the actions of individuals.¹⁸ Examples of system glitches resulting in data breaches include “application failures, inadvertent data dumps, [and] logic errors in data transfer.”¹⁹ The data breach of First American Financial Corporation's insurance records, one of the largest so far in 2019, is a recent example of a system glitch that caused a data breach.²⁰ Early this year, a real estate developer was given links to documents that he had legitimate access to on the company website, but by switching one digit in the link, he was suddenly able to access millions of sensitive private records containing social security numbers and bank account information.²¹ Once notified of the incident, First American concluded that a “technological defect” had caused the data breach.²² It is estimated this data breach exposed 885 million records.²³

¹⁵ Ponemon, *supra* note 14; *Top 3 Causes of Data Breach Are Expensive*, CALYPTIX SEC. (Jun. 29, 2017), <https://www.calyptix.com/top-threats/top-3-causes-data-breach-expensive> [<https://perma.cc/WF4U-XXH8>].

¹⁶ *Top 3 Causes of Data Breach Are Expensive*, *supra* note 15.

¹⁷ Australian Associated Press, *Melbourne Student Health Records Posted Online in “Appalling” Privacy Breach*, GUARDIAN (Aug. 21, 2018, 8:47 PM), <https://www.theguardian.com/australia-news/2018/aug/22/melbourne-student-health-records-posted-online-in-appalling-privacy-breach> [<https://perma.cc/WQ5U-57VM>].

¹⁸ Ponemon, *supra* note 15.

¹⁹ Thor Olavsrud, *Most Data Breaches Caused by Human Error, System Glitches*, CSO (Jun. 17, 2013, 7:00 AM), <https://www.csoonline.com/article/2133631/most-data-breaches-caused-by-human-error--system-glitches.html> [<https://perma.cc/7GBA-DVW9>].

²⁰ *Frequently Asked Questions*, FIRST AM. FIN. CORP. (May 31, 2019), <https://www.firstam.com/incidentupdate/update20190531.html> [<https://perma.cc/D8NZ-RNLD>]; Brian Krebs, *First American Financial Corp. Leaked Hundreds of Millions of Title Insurance Records*, KREBS ON SECURITY (May 24, 2019), <https://krebsonsecurity.com/2019/05/first-american-financial-corp-leaked-hundreds-of-millions-of-title-insurance-records/> [<https://perma.cc/8T4N-GDAR>].

²¹ Krebs, *supra* note 20.

²² *Frequently Asked Questions*, *supra* note 20.

²³ Krebs, *supra* note 20.

C. Malicious or Criminal Data Breaches

Malicious and criminal attacks were the most common form of data breaches in fiscal year 2019 and are on the rise.²⁴ While cybercriminals are constantly trying to develop new methods to stay ahead of data security, it may be useful to understand some of the most prevalent techniques currently in use.²⁵ Some strategies used in malicious and criminal attacks are malware infections, hacking, and social engineering.²⁶

1. *Malware Infections*

Malware is a compound word referring to “malicious software” and is an umbrella term for any type of software designed to infiltrate a computer without the owner’s permission.²⁷ Twenty-eight percent of data breaches in 2018 involved malware installation.²⁸ Some of the most used techniques to install malware are email attachments, direct installations (where further malware is installed after the device is already compromised), and web drive-bys (where the user visits a compromised website that downloads malicious files onto the user’s computer).²⁹ Different types of malware are often used simultaneously and in conjunction with one another.³⁰

²⁴ IBM SECURITY, *supra* note 14, at 6, 21.

²⁵ IDENTITY THEFT RESOURCE CTR., 2018 END-OF-YEAR DATA BREACH REPORT 5 (2019), https://www.idtheftcenter.org/wp-content/uploads/2019/02/ITRC_2018-End-of-Year-Aftermath_FINAL_V2_combinedWEB.pdf [<https://perma.cc/9GFV-ALNH>]; EXPERIAN, DATA BREACH INDUSTRY FORECAST 2 (2019) <https://www.experian.com/assets/data-breach/white-papers/2019-experian-data-breach-industry-forecast.pdf> [<https://perma.cc/SHZ7-2WGJ>].

²⁶ VERIZON, 2018 DATA BREACH INVESTIGATIONS REPORT 7 (2018), https://enterprise.verizon.com/resources/reports/DBIR_2018_Report.pdf [<https://perma.cc/D2PT-7KPD>].

²⁷ *The Playpen Cases: Frequently Asked Questions*, ELECTRONIC FRONTIER FOUND., <https://www.eff.org/pages/playpen-cases-frequently-asked-questions#whatismalware> (accessed Oct. 20, 2019) [<https://perma.cc/YGX7-HHS3>]; *Threat Actions*, VERIS, <http://veriscommunity.net/actions.html> (accessed Oct. 20, 2019) [<https://perma.cc/9C2E-5X93>].

²⁸ VERIZON, 2019 DATA BREACH INVESTIGATIONS REPORT 5 (2019) <https://enterprise.verizon.com/resources/reports/2019-data-breach-investigations-report.pdf> [<https://perma.cc/R5F9-NFBD>].

²⁹ *Id.* at 12; *Drive-by Download*, TREND MICRO (2019), <https://www.trendmicro.com/vinfo/us/security/definition/drive-by-download> [<https://perma.cc/9EP8-GCZT>]; *What Is a Drive-By Download?*, KASPERSKY, <https://www.kaspersky.com/resource-center/definitions/drive-by-download> [<https://perma.cc/3FUB-PW3S>].

³⁰ *Threat Actions*, *supra* note 27.

Some of the most prevalent malware activities used in recent data breaches are the use of backdoors, spyware, and RAM scrapers.³¹ A computer “backdoor” is designed to create an alternative access point, or “backdoor,” around computer security measures, allowing actors unauthorized access to data and can be exploited via either malware or hacking.³² These backdoors are designed to circumvent security in a number of ways and allow actors to access and move files undetected and acquire user passwords.³³ A RAM scraper is malware that is designed to steal credit or debit card data while it is stored in the random access memory (RAM) of a point-of-sale terminal.³⁴ Once installed, the RAM scraper malware collects the card numbers in a readable text file which the malicious actor has access to.³⁵ Spyware is software that is installed in the computer and monitors user activity, sending the information back to the malicious actor.³⁶ The software records information such as the user’s surfing behavior and goes unnoticed by the computer’s owner.³⁷ Using information found by monitoring the user, the malicious actor is able to gain access to data on the computer.³⁸

³¹ VERIZON, *supra* note 28, at 12.

³² VERIZON, *supra* note 28, at 67; *Backdoor*, MALWAREBYTES, <https://www.malwarebytes.com/backdoor/> (accessed Oct. 20, 2019) [<https://perma.cc/UKB9-YR3P>]; Devin Coldewey, *WTF is a Backdoor?*, TECHCRUNCH (Jan. 29, 2017), <https://techcrunch.com/2017/01/29/wtf-is-a-backdoor/> [<https://perma.cc/MQ5B-EGH6>].

³³ DOVE CHIU, SHIH-HAO WENG, & JOSEPH CHIU, TREND MICRO, *BACKDOOR USE IN TARGETED ATTACKS* (2014), <https://www.trendmicro.de/cloud-content/us/pdfs/security-intelligence/white-papers/wp-backdoor-use-in-targeted-attacks.pdf> [<https://perma.cc/YU9A-TTWZ>]; U.S. GOV’T ACCOUNTABILITY OFFICE, *GAO-12-361, IT SUPPLY CHAIN: NATIONAL SECURITY-RELATED AGENCIES NEED TO BETTER ADDRESS RISKS* (2012), <http://www.gao.gov/assets/590/589568.pdf> [<https://perma.cc/AL24-945Z>]; *Backdoor*, *supra* note 32.

³⁴ David W. Opderbeck, *Cybersecurity, Data Breaches, and the Economic Loss Doctrine in the Payment Card Industry*, 75 MD. L. REV. 935, 959 (2016); Brian Riley, *Ram Scraper Malware: Why PCI DSS Can’t Fix Retail*, DARK READING (Jul. 23, 2014), <https://www.darkreading.com/attacks-breaches/ram-scraper-malware-why-pci-dss-cant-fix-retail/a/d-id/1297501> [<https://perma.cc/67D3-LJKS>].

³⁵ Numaan Huq, *A Look at Point of Sale RAM Scraper Malware and How it Works*, SOPHOS (July 16, 2013), <https://nakedsecurity.sophos.com/2013/07/16/a-look-at-point-of-sale-ram-scraper-malware-and-how-it-works/> [<https://perma.cc/ED7M-76FY>].

³⁶ *What is Spyware?—Definition*, KASPERSKY, <https://usa.kaspersky.com/resource-center/threats/spyware> (accessed Oct. 19, 2019) [<https://perma.cc/VV7Q-8JV6>].

³⁷ Stephen Y. Chow, *Conceptions of Privacy and Security in a Digital World*, in *DATA SECURITY AND PRIVACY IN MASSACHUSETTS* § 1.2.2(a) (2018).

³⁸ *What is Spyware*, *supra* note 36.

2. Hacking

Hacking refers to deliberate actions taken to access information by avoiding logical security mechanisms (including passwords and two-factor authentication).³⁹ This is distinct from malware, where the attack focuses on the downloading or use of malicious software.⁴⁰ Hacking techniques were used in fifty-two percent of data breaches in 2018.⁴¹ There are a plethora of activities that fall under the hacking umbrella but three of the more often used strategies are brute force attacks, RFI hacks, and SQL injections.⁴² A brute force attack is conceptually straight-forward. A malicious actor wants to gain access into a database or an encrypted file that is protected behind a username and password.⁴³ The actor utilizes computer programs to run an attack that tries out every combination of, for example, username and password and runs through them as long as it takes until the key is found.⁴⁴ An RFI, or “remote file inclusion,” attack is when a malicious actor inserts language in a website’s URL to redirect a website’s request for a remote file to be used on the website to a malicious file.⁴⁵ Instead of connecting with the intended file, the website connects and runs the malicious file, and depending on what the file is meant to do, can allow the attacker to gain access to the website’s server and the data contained therein.⁴⁶ A Structured Query Language (SQL) Injection is an injection of SQL statements within user inputs to gain access to a database.⁴⁷ SQL is a coding language that is used in managing databases of information.⁴⁸ Often to access this database, a user is asked for a username and password

³⁹ *Threat Actions*, *supra* note 27; Ernest Sampera, *What You Need to Know About Logical Security vs Physical Security*, VXCHNGE, (Jan. 30, 2019), <https://www.vxchnge.com/blog/logical-security-vs-physical-security> [<https://perma.cc/2K2Y-SUNE>].

⁴⁰ *See The Playpen Cases: Frequently Asked Questions*, *supra* note 27; *Threat Actions*, *supra* note 27.

⁴¹ VERIZON, *supra* note 28, at 5.

⁴² VERIZON, *supra* note 28, 10.

⁴³ Chris Hoffman, *Brute Force Attacks Explained: How All Encryption is Vulnerable*, HOW-TO-GEEK (Jul. 6, 2013), <https://www.howtogeek.com/166832/brute-force-attacks-explained-how-all-encryption-is-vulnerable/> [<https://perma.cc/9V3W-9EVQ>].

⁴⁴ *Id.*; Jeff Peters, *What is a Brute Force Attack*, VARONIS (Oct. 16, 2018), <https://www.varonis.com/blog/brute-force-attack/> [<https://perma.cc/P3BA-TE58>].

⁴⁵ *Remote File Inclusion (RFI)*, CWATCH (Feb. 4, 2019), <https://cwatch.comodo.com/blog/cyber-attack/remote-file-inclusion-rfi/> [<https://perma.cc/RXQ9-YW9D>].

⁴⁶ *Id.*

⁴⁷ *What is SQL Injection (SQLi) and How to Prevent It*, ACUNETIX <https://www.acunetix.com/websitesecurity/sql-injection> (accessed Nov. 20, 2019) [<https://perma.cc/32TQ-V4BU>].

⁴⁸ *Understanding SQL Injection*, CISCO SECURITY, https://tools.cisco.com/security/center/resources/sql_injection (accessed Nov. 20, 2019) [<https://perma.cc/7N8A-9GE4>].

(user inputs).⁴⁹ The malicious actor inserts user inputs that are coded in the SQL that the database server reads as a correct input and allows the malicious user access without guessing the username or password.⁵⁰

3. *Social Engineering*

Social engineering utilizes “deception, manipulation, or intimidation” to manipulate humans into providing access to sensitive data. The most prevalent types are phishing and pretexting.⁵¹ Social engineering was a component in thirty-three percent of data breaches in 2018.⁵²

Phishing is a technique where nefarious actors send a designated recipient an email or text message that looks like it is sent from a trusted source.⁵³ They will then direct you to click on a link or open an attachment.⁵⁴ Once the recipient clicks on the link, often either malware is downloaded or a website pops up asking the user to submit sensitive information, which will unknowingly be given to nefarious actors.⁵⁵ Pretexting is a more targeted form of phishing where a cybercriminal attempts to create a dialogue with a designated target.⁵⁶ An example of pretexting would be if a cybercriminal sent an email to a company employee pretending to be the CEO of the company.⁵⁷ The cybercriminal would ask for sensitive information pretending to be the CEO with the aim of getting the employee to provide information that the cybercriminal would otherwise not have access to.⁵⁸

III. WHY MALICIOUS ACTORS CONDUCT DATA BREACHES

There are a number of reasons why an individual or group or organization would conduct a data breach. However, the two most common

⁴⁹ *What is SQL Injection (SQLi) and How to Prevent It*, *supra* note 47.

⁵⁰ *SQL Injection*, HACKSPLAINING, <https://www.hacksplaining.com/exercises/sql-injection/#/hack-complete> (accessed Nov. 20, 2019) [<https://perma.cc/K6X7-NTTJ>].

⁵¹ *Threat Actions*, *supra* note 27.

⁵² VERIZON, *supra* note 28, at 5.

⁵³ Fed. Trade Comm’n, *How to Recognize and Avoid Phishing Scams*, FTC: CONSUMER INFORMATION, <https://www.consumer.ftc.gov/articles/how-recognize-and-avoid-phishing-scams> (Oct. 19, 2019) [<https://perma.cc/C98U-RGGJ>].

⁵⁴ *Id.*

⁵⁵ Mark Bassingthwaighe, *Cybercrime and Social Engineering*, 61 *ADVOCATE* 42, 42 (2018); *Phishing vs Spear Phishing*, BARRACUDA NETWORKS, INC., <https://www.barracuda.com/glossary/phishing-spear-phishing> (accessed Oct. 19, 2019) [<https://perma.cc/Z69R-KLLE>].

⁵⁶ VERIZON, *supra* note 26, at 11.

⁵⁷ *Id.*

⁵⁸ *Id.*

motivations of data breaches are financial gain and espionage.⁵⁹ These two motivations are believed to make up the vast majority of data breaches.⁶⁰

A. Financial Gain

Often, hackers and cybercriminals will conduct data breaches for the sole purpose of financial gain. In fact, well over fifty percent of data breaches from 2010 to 2016 were for financial gain.⁶¹ Cybercriminals can profit off of data breaches through both goods (the stolen data itself) and services (conducting attacks that result in data breaches), both of which can be bought and sold on the online black market.⁶² On the black market earlier this year, the value of an individual Facebook account was \$9.12 and the value of individual debit card information was \$250.05.⁶³ Further, the market for a specific piece of data can vary depending on its potential access to new sources of data, especially if the data may be used to unlock other accounts.⁶⁴ These purchases of goods and services may be completed with traditional currency but are increasingly conducted using cryptocurrencies as both a source of security and anonymity for the parties involved.⁶⁵

B. Espionage

The second most common motivation behind data breaches is espionage.⁶⁶ These are conducted to gain insight into either company trade secrets or to spy into a country's classified information.⁶⁷ State-affiliated groups and countries are behind ninety-six percent of espionage motivated

⁵⁹ COUNCIL OF ECON. ADVISERS, EXEC. OFFICE OF THE PRESIDENT, *THE COST OF MALICIOUS CYBER ACTIVITY TO THE U.S. ECONOMY 6* (2018), <https://www.whitehouse.gov/wp-content/uploads/2018/02/The-Cost-of-Malicious-Cyber-Activity-to-the-U.S.-Economy.pdf> [https://perma.cc/5DFW-MK68].

⁶⁰ VERIZON, *supra* note 26, at 6, 7.

⁶¹ *Top 3 Causes of Data Breach Are Expensive*, *supra* note 15.

⁶² Lillian Ablon, *Data Thieves: The Motivations of Cyber Threat Monetization of Stolen Data*, RAND CORP. (Mar. 15, 2018), https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT490/RAND_CT490.pdf [https://perma.cc/26A6-FY8F].

⁶³ *Id.*, Simon Migliano, *Dark Web Market Price Index-2019 (US Edition)*, TOP10VPN.COM (Feb. 20, 2019), <https://www.top10vpn.com/news/privacy/dark-web-market-price-index-2019-us-edition/> [https://perma.cc/ZSM4-G5T6].

⁶⁴ Ablon, *supra* note 62.

⁶⁵ *Id.*

⁶⁶ VERIZON, *supra* note 26, at 6.

⁶⁷ NAT'L COUNTERINTELLIGENCE AND SEC. CTR., *FOREIGN ECONOMIC ESPIONAGE IN CYBERSPACE* (2018), <https://www.dni.gov/files/NCSC/documents/news/20180724-economic-espionage-pub.pdf> [https://perma.cc/ZH92-NY6Y]; Ablon, *supra* note 62.

breaches.⁶⁸ Phishing and backdoors were used in a high number of cyber-espionage incidents.⁶⁹ Attacks conducted by state-sponsored actors are unique in that they are conducted on behalf of a nation's interest and are widely considered "legitimate state activity."⁷⁰ A number of data breaches over the past few years have come to the front of public conversation due to state-sponsored subversion and espionage, with prominent examples being the North Korean attack on Sony Entertainment in response to the release of the movie *The Interview* and the United States Democratic National Committee breach in 2016.⁷¹

IV. CONCLUSION

Data breaches are a constant threat to information kept on computer systems and all sectors and industries within our society are targeted. The breaches occur in a number of ways ranging from innocent human mistakes to malicious and complex malware infections or hacking. When these breaches are deliberate, they are often conducted for financial gain or espionage purposes. The legal community is increasingly asked to respond to the persistent threat of data breaches. Understanding the problem of how data breaches happen and what the motivations are behind them is a baseline for informed decision-making to combat this threat in all aspects of our society.

⁶⁸ VERIZON, *supra* note 28, at 25.

⁶⁹ *Id.*

⁷⁰ Ablon, *supra* note 62.

⁷¹ Ablon, *supra* note 62; SYMANTEC, INTERNET SECURITY THREAT REPORT: VOLUME 22 at 7, 44(2017), <https://www.symantec.com/content/dam/symantec/docs/reports/istr-22-2017-en.pdf> [<https://perma.cc/4HAD-3QLL>].

5G WIRELESS CONNECTIVITY: THE NEXT STEP

Matthew Wells*

CITE AS: 4 GEO. L. TECH. REV. 325 (2019)

TABLE OF CONTENTS

I. INTRODUCTION	325
II. 5G DEFINED.....	326
III. 5G ENABLING TECHNOLOGIES.....	327
A. Small Cells.....	328
B. Massive MIMO.....	328
C. Beamforming	329
D. Full Duplex	329
IV. BENEFITS OF 5G.....	330
V. APPLICATIONS OF 5G.....	330
VI. CONCLUSION.....	332

I. INTRODUCTION

Wireless communication technology is about to enter a new era: 5G networks. “5G” stands for “fifth generation,” the next step up from the currently mainstream 4G networks.¹ From 2014 levels, mobile traffic is expected to increase by 1000 times by 2024.² The next generation of wireless network, known as the 5G network, must be able to support this massive increase.³ This paper will define 5G networks, explain the technologies that underly its creation and feasibility, and explore the network’s potential applications.

* Georgetown University Law Center, J.D. Candidate 2021; University of Michigan, B.A. in Philosophy and Cognitive Science 2018.

¹ *What is 5G?*, VERIZON, <https://www.verizon.com/about/our-company/5g/what-5g> (Oct. 27, 2019) [<https://perma.cc/7BBH-SAL3>].

² Qian Clara Li et al., *5G Network Capacity: Key Elements and Technologies*, IEEE VEHICULAR TECH. MAG. 71 (2014), <https://ieeexplore.ieee.org/document/6730679> [<https://perma.cc/XX2N-3VZ7>].

³ *Id.*

II. 5G DEFINED

The 5G network is the fifth generation of wireless technology. The first generation, capable of transmitting wireless phone calls, emerged with the invention of cell phones.⁴ The second generation paved the way for text messaging and voicemails.⁵ 3G brought web browsing to mobile devices, and 4G brought deep web functionality to those devices.⁶

Many people are familiar with the term “4G LTE.” LTE stands for “long term evolution,” an industry standard for 4G wireless technology that allowed for greater speed on 4G networks over its 3G predecessor.⁷ The equivalent of LTE for 5G wireless networks is “NR,” or “new radio,” a standard that allows for 5G technology to scale as needed.⁸ Think of the jump from 4G LTE to 5G NR as analogous to the jump from Blu-ray to Netflix; “it’s the way it is built rather than just the speed that makes it different.”⁹ The 3rd Generation Partnership Project (3GPP), a group of telecom organizations that create standards for wireless technology, set the 4G LTE and 5G NR standards.¹⁰

Wireless networks connect devices by utilizing a spectrum of radio frequency waves transmitted through the air to send and receive messages.¹¹ As smart phones and tablets have risen in popularity, existing cellular networks are expected to have a capacity shortage in the coming years.¹² In order to fight this shortage, 5G networks seek to expand the frequency spectrum that devices use to communicate.¹³ Millimeter wave technology

⁴ *What is 5G?*, *supra* note 1.

⁵ *Id.*

⁶ *Id.*

⁷ *In re Qualcomm Antitrust Litigation*, 292 F. Supp. 3d 948, 955 (N.D. Ca. 2017).

⁸ Simon Rockman, *Why 5G Isn't Just Faster 4G*, FORBES (May 25, 2019, 5:12 PM), <https://www.forbes.com/sites/simonrockman1/2019/05/25/why-5g-isnt-just-faster-4g/#361e45f143a6> [<https://perma.cc/P54Y-4YE5>].

⁹ *Id.*

¹⁰ *About 3GPP Home*, 3GPP, <https://www.3gpp.org/about-3gpp/about-3gpp> (accessed Oct. 27, 2019) [<https://perma.cc/2SB2-F7GP>].

¹¹ Maruerite Reardon, *Wireless Spectrum: What It Is, and Why You Should Care*, CNET (Aug. 13, 2012, 12:01 AM), <https://www.cnet.com/g00/news/wireless-spectrum-what-it-is-and-why-you-should-care/?i10c.ua=4&i10c.encReferrer=aHR0cHM6Ly9wZXJtYS5jYy9ZwJjXLU02UVA%3d> [<https://perma.cc/YZ2W-M6QP>].

¹² Kei Sakaguichi et al., *Millimeter-Wave Evolution for 5G Cellular Networks*, E98-B IEICE TRANSACTIONS ON COMM. 388 (2015), <http://home.deib.polimi.it/capone/papers/IEICE2015.pdf> [<https://perma.cc/5T7K-Y2MH>].

¹³ *Id.*

utilizes previously unused radio frequencies to transmit wireless signals.¹⁴ Previous generations of wireless networks used frequencies below 6 GHz, whereas millimeter waves broadcast between 30 and 300 GHz, meaning that 5G networks use waves that are much smaller compared to their 4G predecessors.¹⁵ By using previously untapped frequencies, 5G networks lower the traffic on the old networks, as well as open the door for low-traffic communications on the new networks. Millimeter waves have several other advantages as well. For example, millimeter waves reduce interference between neighboring connections.¹⁶ Additionally, utilizing millimeter waves could improve energy efficiency in wireless networks.¹⁷ Millimeter waves, however, have a glaring flaw; they can only travel short distances and are highly susceptible to interference from weather and physical obstacles.¹⁸

5G aims to not only improve network speeds and capacity, but also to improve spectrum efficiency. Spectrum efficiency measures how effectively the wireless frequency system is utilized, measured in bits per second per Hertz.¹⁹ Essentially, spectrum efficiency is transferring as much information as possible while utilizing the minimum amount of the spectrum necessary to do so.²⁰ By using less of the spectrum to send the same amount of information, spectrum efficiency frees up space on the network to send more information. Thus, 5G networks will be able to better support the large number of wireless devices projected to come into the market in the coming years.

III. 5G ENABLING TECHNOLOGIES

5G is often referred to as a single technology, but in reality, it is an amalgam of several different technologies which comprise the network. Some of those technologies are small cells, Massive MIMO, beamforming, and Full Duplex.²¹

¹⁴ Amy Nordrum et al., *Everything You Need to Know About 5G*, IEEE SPECTRUM (Jan. 27, 2017, 7:00 PM), <https://spectrum.ieee.org/video/telecom/wireless/everything-you-need-to-know-about-5g> [<https://perma.cc/KS6Z-3Q7A>].

¹⁵ Broadband Div, Fed. Comm. Comm'n, *Millimeter Wave 70/80/90 GHz Service*, FCC.GOV, <https://www.fcc.gov/wireless/bureau-divisions/broadband-division/microwave-services/millimeter-wave-708090-ghz-service> [<https://perma.cc/B4S6-T7YS>].

¹⁶ Sakaguchi et al., *supra* note 12.

¹⁷ *Id.*

¹⁸ Margaret Rouse, *5G*, SEARCHNETWORKING, <https://searchnetworking.techtarget.com/definition/5G> (accessed Oct. 20, 2019) [<https://perma.cc/M2LP-QNJQ>].

¹⁹ Hong et al., *Cognitive Radio in 5G: A Perspective on Energy-Spectral Efficiency Trade-off*, 52 IEEE COMM. MAG. 46, 48 (July 2014), <https://ieeexplore.ieee.org/document/6852082> [<https://perma.cc/5B86-3TX5>].

²⁰ FED. COMM. COMM'N, SPECTRUM POLICY TASK FORCE: REPORT OF THE SPECTRUM EFFICIENCY WORKING GROUP 5 (2002).

²¹ Nordrum et al., *supra* note 14.

A. Small Cells

Small cells are 5G wireless signal base stations that are placed around an area in large numbers.²² Small cells are necessary to make millimeter waves technology viable, as they can be more easily placed in a wide range of places around cities than traditional cell towers.²³ This solves two issues posed by utilizing millimeter waves. First, by having a greater number of access points, devices are less likely to experience interference from weather and physical objects. Second, it increases the network's efficiency by allowing the spectrum to be reused in the various cells.²⁴ However, the large number of small cells needed and the short distances between them pose two new problems: first, that it will be difficult to set up 5G networks in rural areas, and second, that it will require a significantly larger amount of antennas at base stations than current networks use.

B. Massive MIMO

Massive multiple-input, multiple-output (Massive MIMO) is central to the evolution of 5G networks due to its ability to boost spectral efficiency.²⁵ MIMO technology allows a network to transmit and receive multiple data signals over a singular frequency, thus multiplying the capacity of the network without having to expand the frequency band.²⁶ Put more simply, "If spectrum is a highway, then MIMO doesn't just add more lanes; it adds levels, creating multiple-decker highways that vastly increase the capacity of the network."²⁷ Current 4G base stations already implement MIMO technology, but they only have about twelve ports for antennas, whereas 5G base stations with massive MIMO can support a hundred ports.²⁸ Thus, implementing massive MIMO

²² CHRIS D. LINEBAUGH, CONG. RESEARCH SERV., LSB10265, OVERVIEW OF LEGAL CHALLENGES TO THE FCC'S 5G ORDER ON SMALL CELL SITING (Feb. 25, 2019), <https://crsreports.congress.gov/product/pdf/LSB/LSB10265> [<https://perma.cc/H2C9-4NGT>].
²³ *Id.*

²⁴ Theodore S. Rappaport et al., *Overview of Millimeter Wave Communications for Fifth-Generation (5G) Wireless Networks—With a Focus on Propagation Models*, 65 IEEE TRANSACTIONS ON ANTENNAS & PROPAGATION 6213, 6214–15 (2017), <https://ieeexplore.ieee.org/document/7999294> [<https://perma.cc/D2BL-UGK6>].

²⁵ Xiaochen Xia et al., *A 5G-Enabling Technology: Benefits, Feasibility, and Limitations of In-Band Full-Duplex mMIMO*, IEEE VEHICULAR TECH. MAG. 81, 82 (Sept. 2018), <https://ieeexplore.ieee.org/abstract/document/8396262> [<https://perma.cc/FU6A-3U2U>].

²⁶ Jon Mundy & Kevin Thomas, *What is Massive MIMO Technology?* 5G.CO.UK, <https://5g.co.uk/guides/what-is-massive-mimo-technology> (accessed Nov. 18, 2019) [<https://perma.cc/9NAZ-MWAB>].

²⁷ *A Closer Look at Massive MIMO*, SPRINT BUS. (Nov. 7, 2018), <https://business.sprint.com/blog/massive-mimo/> [<https://perma.cc/AD5X-K9KM>].

²⁸ Xia et al., *supra* note 25.

over MIMO in 5G base stations causes considerable gains in network efficiency by lowering the workload of each antenna. Unfortunately, however, the massive number of antennas in close proximity can cause significant interference in the network when the signals crash into one another.²⁹

C. Beamforming

Beamforming is a traffic-signaling system that reduces the interference inherent in Massive MIMO base stations.³⁰ Essentially, beamforming shapes a radio frequency signal in order to send that signal along the most efficient pathway to the wireless device.³¹ This allows for the network to control interference and more equitably allocate network resources by shaping the signal around objects and other signals.³² Due to the signal crashing that occurs with massive MIMO, beamforming is necessary to make sure the signals are shaped in such a way that they do not interfere with one another. Beamforming also assists millimeter waves in avoiding signal-blocking obstacles that normally interfere with smaller radio waves.³³

D. Full Duplex

Full Duplex is another technology used to increase efficiency on 5G networks. Current wireless networks require base stations and phones to take turns transmitting and receiving on the network.³⁴ Full Duplex allows transceivers on base stations and cell phones to simultaneously transmit and receive data on the same frequency.³⁵ It does so by utilizing silicon transistors which can act as switches on the frequency, temporarily rerouting signals to avoid a crash between two signals traveling on the same frequency.³⁶ Thus, Full Duplex technology can theoretically double the capacity of wireless networks.

²⁹ Nordrum et al., *supra* note 14.

³⁰ *Id.*

³¹ *Id.*

³² Faezeh Alavi et al., *Beamforming Techniques for Nonorthogonal Multiple Access in 5G Cellular Networks*, 67 IEEE TRANSACTIONS ON VEHICULAR TECH. 9474 (2018), <https://ieeexplore.ieee.org/document/8411153> [<https://perma.cc/WW7C-Q4ZS>].

³³ *Id.*

³⁴ Xia et al., *supra* note 25.

³⁵ Nordrum et al., *supra* note 14.

³⁶ Amy Nordrum et al., *5G Bytes: Full Duplex Explained*, IEEE SPECTRUM (Apr. 1, 2017 2:21 PM), <https://spectrum.ieee.org/video/telecom/wireless/5g-bytes-full-duplex-explained> [<https://perma.cc/WWF6-7L36>].

IV. BENEFITS OF 5G

5G NR technology boasts increased data transfer speeds, increased capacity, and decreased latency over its 4G LTE predecessor.³⁷ This means more devices will be able to download more information at greater speeds, with less lag time. According to Samsung, 5G will be around 100 times faster than current 4G networks,³⁸ and others claim that some parts of the 5G network will support up to a million devices per square kilometer.³⁹

Low latency is a key improvement of 5G networks that could enable applications that were impossible over 4G networks. Latency is the time it takes for data to be sent from a device until it reaches a receiver.⁴⁰ This is distinct from network speed, which is the amount of information that can be transmitted in a certain amount of time.⁴¹ Part of the reason that 5G networks will be able to handle all these new technologies is a feature called network slicing.⁴² Network slicing allows operators to create multiple networks embedded within the larger network, tailoring each device's connection to meet its specific needs.⁴³ For example, an autonomous vehicle may need a faster connection than a smart light bulb; network slicing can allocate the car a faster connection than the bulb.⁴⁴ This optimizes the network by making sure that the lower-usage devices are not eating up the bandwidth required to run higher-usage devices.

V. APPLICATIONS OF 5G

Some technologies that were initially developed for 4G networks were not viable for the mass market because of 4G's drawbacks. 5G networks' high speed, high capacity, and low latency characteristics allow it to improve technologies such as remote healthcare, autonomous vehicles, virtual reality, and the Internet of Things.⁴⁵

³⁷ Rouse, *supra* note 18.

³⁸ *How Fast Is 5G?* SAMSUNG, <https://www.samsung.com/global/galaxy/what-is/5g-speeds/> [https://perma.cc/8FTG-T685].

³⁹ Rockman, *supra* note 8.

⁴⁰ *Id.*

⁴¹ *Id.*

⁴² Rouse, *supra* note 18.

⁴³ *Id.*

⁴⁴ *Id.*

⁴⁵ Nordrum et al., *supra* note 14.

The reliable, low latency connections provided by 5G networks will help improve technology in industries such as healthcare services.⁴⁶ Two applications for 5G networks in healthcare are bio-connectivity and remote robotic surgery.⁴⁷ Bio-connectivity allows for the decentralization of hospitals so that medical care can be provided in a patient's home or in an emergency vehicle by providing access to electronic medical records, data analysis for predictive healthcare, and pharmaceutical analysis.⁴⁸ 5G networks in this context can allow the information to travel more quickly and efficiently, potentially saving lives. Remote interventions allow for complex medical interventions and surgeries using special equipment over an internet connection.⁴⁹ This can help doctors provide these services to people in areas where healthcare is less accessible.⁵⁰

Autonomous vehicles are another use case for 5G networks. Many of the devices necessary for operating an autonomous vehicle require low latency and high reliability.⁵¹ A low latency 5G network is required to ensure minimal communication times between the sensors and the car's computer to be as low as possible, which can allow the car to react quickly to dangerous situations.⁵² Additionally, autonomous vehicles must transfer an extremely large amount of data, which requires faster speeds than 4G networks.⁵³

5G networks also will help push media and entertainment technology forward. Recent trends in the entertainment industry have caused users to demand a more immersive experience.⁵⁴ 5G can allow for higher quality live and streamed content, offer live audiences an enhanced experience, and allow for faster information sharing within media production.⁵⁵ The gaming industry, in particular, could benefit from a 5G network. Virtual and Augmented Reality (VR and AR, respectively) is a growing trend in the gaming industry designed to enhance immersive experience.⁵⁶ VR allows

⁴⁶ Abdul Ahad et al., *5G-Based Smart Healthcare Network: Architecture, Taxonomy, Challenges and Future Research Directions*, 7 IEEE ACCESS 100747, 100752–53 (2019), <https://ieeexplore.ieee.org/document/8769822> [<https://perma.cc/V4PU-G665>].

⁴⁷ Maria A. Lema et al., *Business Case and Technology Analysis for 5G Low Latency Applications*, 5 IEEE ACCESS 5917, 5919 (2017), <https://arxiv.org/pdf/1703.09434.pdf> [<https://perma.cc/TN22-V8J8>].

⁴⁸ *Id.*

⁴⁹ *Id.*

⁵⁰ *Id.*

⁵¹ *Id.* at 5920.

⁵² Ralf Llanasas, *5G's Important Role in Autonomous Car Technology*, MACHINEDSIGN (Mar. 11, 2019), <https://www.machinedesign.com/motion-control/5g-s-important-role-autonomous-car-technology> [<https://perma.cc/QC78-YYLZ>].

⁵³ *Id.*

⁵⁴ Lema, *supra* note 47, at 5920.

⁵⁵ *Id.*

⁵⁶ *Id.*

users to enter a virtual world, whereas AR adds virtual elements to the real world.⁵⁷ Low latency is fundamentally necessary for VR and AR technologies to provide a truly immerse experience.⁵⁸ VR gaming is also demanding from a capacity standpoint,⁵⁹ and 5G increased bandwidth makes it more suitable for such applications.

5G networks can also help improve the capabilities of the Internet of Things. The Internet of Things is a network infrastructure that configures itself using standard communication protocols between devices, allowing for greater compatibility of connections across devices.⁶⁰ The information stream allows for easier automation for everyday tasks.⁶¹ For example, a connected alarm clock could check the traffic report before a morning commute, adjust the alarm accordingly, and also adjust the connected coffee maker's timing.⁶² The Internet of Things could also be useful in the a manufacturing context, where the connectivity of sensors and robots could lead to improvements in factory or warehouse maintenance.⁶³ This network requires a large number of network-connected devices, and 5G has the capacity to support significantly more devices than its predecessor 4G LTE networks.⁶⁴ Latency in 4G LTE networks have also limited Internet of Things technologies in the past, and 5G's low latency can help push past those limitations.⁶⁵

VI. CONCLUSION

4G LTE changed the way we live our everyday lives. It allowed us to access a previously untapped wealth of information, all from devices small enough to fit in our pocket. 5G wireless networks are the next step in that evolution. It is unclear at this early stage exactly how it will change our lives, but there is little doubt that these new technologies will have a profound impact on our history.

⁵⁷ Brian Cooley, *AR and VR made simple*, CNET (Aug. 6, 2018, 5:00 AM), <https://www.cnet.com/how-to/ar-and-vr-made-simple/> [https://perma.cc/ZA2Z-MFZE].

⁵⁸ Lema, *supra* note 47, at 5920–21.

⁵⁹ *Id.* at 5921

⁶⁰ Ahad, *supra* note 46, at 100748.

⁶¹ *Id.*

⁶² *Id.*

⁶³ Hatem Zeine, *What the Future of IoT and 5G May Look Like*, FORBES (Nov. 1, 2018, 7:45 AM), <https://www.forbes.com/sites/forbestechcouncil/2018/11/01/what-the-future-of-iot-and-5g-may-look-like/#6cb33251629b> [https://perma.cc/8HCJ-9H3C].

⁶⁴ Nicholas Shields, *Here's How 5G Will Revolutionize the Internet of Things*, BUS. INSIDER (June 15, 2017, 10:15 AM), <https://www.businessinsider.com/how-5g-will-revolutionize-the-internet-of-things-2017-6> [https://perma.cc/KF6Q-K9YG].

⁶⁵ *Id.*