

# **35:1 BERKELEY TECHNOLOGY LAW JOURNAL**

2020

**Pages  
1  
to  
366**

Berkeley Technology Law Journal  
Volume 35, Number 1

**Production:** Produced by members of the *Berkeley Technology Law Journal*.  
All editing and layout done using Microsoft Word.

**Printer:** Joe Christensen, Inc., Lincoln, Nebraska.  
Printed in the U.S.A.  
The paper used in this publication meets the minimum requirements  
of American National Standard for Information Sciences—  
Permanence of Paper for Library Materials, ANSI Z39.48—1984.

**Copyright © 2020 Regents of the University of California.**  
All Rights Reserved.



Berkeley Technology Law Journal  
University of California  
School of Law  
3 Law Building  
Berkeley, California 94720-7200  
editor@btlj.org  
<https://www.btlj.org>

# BERKELEY TECHNOLOGY LAW JOURNAL

VOLUME 35

NUMBER 1

2020

## TABLE OF CONTENTS

### ARTICLES

DIGITAL REMEDIES .....	1
<i>Maayan Perel</i>	
PATENT MARKETS AND INNOVATION IN THE ERA OF BIG PLATFORM COMPANIES.....	53
<i>Robert P. Merges</i>	
ARTIFICIAL INTELLIGENCE OPINION LIABILITY .....	113
<i>Yavar Bathaee</i>	
TARPTIS: THE STICKY CONSEQUENCES OF POORLY IMPLEMENTING TECHNOLOGY-ASSISTED REVIEW .....	171
<i>David Dowling</i>	
MEASURING AND PROTECTING PRIVACY IN THE ALWAYS-ON ERA.....	197
<i>Dan Feldman &amp; Eldar Haber</i>	
WHAT IS IT ABOUT LOCATION? .....	251
<i>Kirsten Martin &amp; Helen Nissenbaum</i>	
CAN YOU PAY FOR PRIVACY? CONSUMER EXPECTATIONS AND THE BEHAVIOR OF FREE AND PAID APPS.....	327
<i>Kenneth A. Bamberger, Serge Egelman, Catherine Han, Amit Elazari Bar On &amp; Irvin Reyes</i>	

## SUBSCRIBER INFORMATION

The *Berkeley Technology Law Journal* (ISSN1086-3818), a continuation of the *High Technology Law Journal* effective Volume 11, is edited by the students of the University of California, Berkeley, School of Law and is published in print three times each year (March, September, December), with a fourth issue published online only (July), by the Regents of the University of California, Berkeley. Periodicals Postage Rate Paid at Berkeley, CA 94704-9998, and at additional mailing offices. POSTMASTER: Send address changes to Journal Publications, University of California, Berkeley Law—Library, LL123 Boalt Hall—South Addition, Berkeley, CA 94720-7210.

**Correspondence.** Address all correspondence regarding subscriptions, address changes, claims for non-receipt, single copies, advertising, and permission to reprint to Journal Publications, University of California, Berkeley Law—Library, LL123 Boalt Hall—South Addition, Berkeley, CA 94705-7210; (510) 643-6600; [JournalPublications@law.berkeley.edu](mailto:JournalPublications@law.berkeley.edu). *Authors:* see section titled Information for Authors.

**Subscriptions.** Annual subscriptions are \$65.00 for individuals and \$85.00 for organizations. Single issues are \$30.00. Please allow two months for receipt of the first issue. Payment may be made by check, international money order, or credit card (MasterCard/Visa). Domestic claims for non-receipt of issues should be made within 90 days of the month of publication; overseas claims should be made within 180 days. Thereafter, the regular back issue rate (\$30.00) will be charged for replacement. Overseas delivery is not guaranteed.

**Form.** The text and citations in the *Journal* conform generally to the THE CHICAGO MANUAL OF STYLE (16th ed. 2010) and to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 20th ed. 2015). Please cite this issue of the *Berkeley Technology Law Journal* as 35 BERKELEY TECH. L.J. \_\_\_\_ (2020).

## BTLJ ONLINE

The full text and abstracts of many previously published *Berkeley Technology Law Journal* articles can be found at <https://www.btlj.org>. Our site also contains a cumulative index; general information about the *Journal*; the BTLJ Blog, a collection of short comments and updates about new developments in law and technology written by BTLJ members; and *BTLJ Commentaries*, an exclusively online publication for pieces that are especially time-sensitive and shorter than typical law review articles.

## INFORMATION FOR AUTHORS

The Editorial Board of the *Berkeley Technology Law Journal* invites the submission of unsolicited manuscripts. Submissions may include previously unpublished articles, essays, book reviews, case notes, or comments concerning any aspect of the relationship between technology and the law. If any portion of a manuscript has been previously published, the author should so indicate.

**Format.** Submissions are accepted in electronic format through Scholastica online submission system. Authors should include a curriculum vitae and resume when submitting articles, including his or her full name, credentials, degrees earned, academic or professional affiliations, and citations to all previously published legal articles. The Scholastica submission website can be found at <https://btlj.scholasticahq.com/for-authors>.

**Citations.** All citations should conform to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 20th ed. 2015).

**Copyrighted Material.** If a manuscript contains any copyrighted table, chart, graph, illustration, photograph, or more than eight lines of text, the author must obtain written permission from the copyright holder for use of the material.

# DONORS

The *Berkeley Technology Law Journal* and the Berkeley Center for Law & Technology acknowledge the following generous donors to Berkeley Law's Law and Technology Program:

## Partners

FENWICK & WEST LLP

ORRICK, HERRINGTON &  
SUTCLIFFE LLP

WHITE & CASE LLP

## Benefactors

BAKER BOTTS LLP

MORRISON & FOERSTER LLP

COOLEY LLP

PAUL HASTINGS LLP

COVINGTON & BURLING LLP

POLSINELLI LLP

FISH & RICHARDSON P.C.

SIDLEY AUSTIN LLP

JONES DAY

WEIL, GOTSHAL & MANGES LLP

KIRKLAND & ELLIS LLP

WILMER CUTLER PICKERING HALE  
AND DORR LLP

LATHAM & WATKINS LLP

WILSON SONSINI GOODRICH &  
ROSATI

MCDERMOTT WILL & EMERY

WINSTON & STRAWN LLP

## Corporate, Government, Individual, and Foundation Sponsors

ATLASSIAN

KILBURN & STRODE LLP

CORNERSTONE RESEARCH

LITINOMICS

DARTS IP

MARKS & CLERK LAW LLP

DORSEY & WHITNEY

MICROSOFT CORPORATION

FUTURE OF PRIVACY FORUM

NERA ECONOMIC CONSULTING

ROBERT GLUSHKO

PALANTIR

GOOGLE, INC.

PHARMACEUTICAL RESEARCH AND  
MANUFACTURERS OF AMERICA

H. WILLIAM HARLAN

PwC

HICKMAN PALERMO BECKER  
BINGHAM LLC

QUALCOMM

INTEL

VIA LICENSING CORP

INVENTIONSHARE

VYNYL

WESTERN DIGITAL

## Members

ANJIE LAW FIRM	KILPATRICK TOWNSEND & STOCKTON LLP
BAKER & MCKENZIE LLP	KNOBBE MARTENS OLSON & BEAR LLP
BEIJING EAST IP	MORGAN, LEWIS & BOCKIUS LLP
CROWELL & MORING	ROBINS KAPLAN LLP
DESMARAIS LLP	ROPES & GRAY LLP
DURIE TANGRI LLP	SIMPSON THACHER & BARTLETT LLP
GREENBERG TRAURIG	TENSEGRITY LAW GROUP LLP
GTC LAW GROUP LLP & AFFILIATES	TROUTMAN SANDERS LLP
HAYNES AND BOONE, LLP	VAN PELT, YI & JAMES LLP
HOGAN LOVELLS, LLP	WANHUIDA INTELLECTUAL PROPERTY
IRELL & MANELLA LLP	WEAVER AUSTIN VILLENEUVE & SAMPSON LLP
KEKER VAN NEST & PETERS LLP	WILLKIE FARR & GALLAGHER LLP
WOMBLE BOND DICKINSON LLP	



# BOARD OF EDITORS

# 2019–2020

---

## *Executive Board*

---

*Editor-in-Chief*  
CHELSEA ANDRE

*Senior Articles Editors*  
LESLIE DIAZ  
CRISTINA MORA  
COURTNEY REED

*Senior Executive Editor*  
SAVANNAH CARNES

*Managing Editor*  
AISLINN SMALLING

*Senior Production Editor*  
MEGAN MCKNELLY

*Senior Scholarship Editor*  
DANIEL CHASE

*Senior Annual Review Editors*  
JULEA LIPIZ  
MIRANDA RUTHERFORD

*Senior Online Content Editor*  
CONCORD CHEUNG

---

## *Editorial Board*

---

*Production Editors*  
ANGELA GRIGGS  
JANELLE LAMB  
EMILY ROBERTS  
HAILEY YOOK

*Symposium Editor*  
ARMBIEN SABILLO  
  
*Online Content Editor*  
GINETTA SAGAN

*Technical Editors*  
MADISON BOWER  
MIN JUNG “MJ” HAN  
ZACK JACOBS  
RACHEL WILSON

*Annual Review Editors*  
KRISTINA KRASNIKOVA  
KEVIN YANG

*Submissions Editors*  
CHRISTINA CROWLEY  
DAVID FANG  
MEHTAB KHAN

*Notes & Comments Editors*  
HARRISON GERON  
ALLAA MAGEID

*Podcast Editor*  
ALLAN HOLDER

*Web & Technology Editor*  
KARNIK HAJJAR

*LLM Editor*  
IGOR SILVA

*Alumni Relations Editor*  
NICK CALCATERRA

*External Relations Editor*  
ASHLEIGH LUSSENDEN

*Member Relations Editor*  
MICHELLE ZIPERSTEIN

*Articles Editors*  
MUHTADI CHOUDHURY  
MATTHEW CHUNG  
SHWETA DUGGAL  
JASON FRANCIS

*Articles Editors*  
KELLY GO  
EMMA LEE  
WALTER MOSTOWY

*Articles Editors*  
JOSH SEDGWICK  
CARMEN SOBCZAK  
MARTA STUDNICKA  
MEI XUAN

# MEMBERSHIP

Vol. 35 No. 1

---

## *Associate Editors*

---

ELIZABETH FU  
LOC HO

NOAHLANI  
LITWINSELLA

JENNY QUANG  
ANDY ZACHRICH

---

## *Members*

---

LIAM AARTASH	CALVIN HANNAGAN	MAXIMIN ORSERO
SHAHAD ALFAWAZ	ALEXANDRA HARVEY	JOSHUA PARZIVAND
BADER ALSHABANAT	ELIZABETH HECKMANN	NINA POUGET
TAIT ANDERSON	SOONYOUNG HEO	GAYATRI RAGHUNANDAN
CHRISTOPHER BARCLAY	JENNIFER HEWITT	EMILY ROBERTS
JOHN BATOHA	THOMAS HORN	FABIOLA ROSSY SEPTYA
MAYA BAUMER	JEFFREY JACOBSEN	CHRISTINA
MARGERITE BLASE	TOM JAMES	EILEEN SANFORD
RASMUS BLOM	CARTER JANSEN	SHEETAL SARAN
CONNOR BOEHM	ANJANAYE JARIWALA	YEMAJ SHEIK
VERONICA BOGNOT	GIA JUNG	ZIYU SHI
JONATHAN CHACON	PHILIP KATZ	DAKOTA SNEED
SUSIE CHEN	IAN KELLY	TAYLOR TAM
KEVIN J. CHEN	YEJI KIM	THERESA TAN
JENNIFER CHUNG	GRACE HO JUNG KIM	AMREEN TANEJA
HENDRIK COPPOOLSE	JOSEPH KROON	RACHEL TERRELL-PERICA
NATALIE CRAWFORD	TZU-I LEE	ELHITA THAMPURAN
KATHARINE CURRAULT	YI SHYUAN LEE	RACHEL THOMPSON
JAMESON DAVIS	GASPARÉ LODERER	MICKEALA TU
CHENXI DUAN	ANNE LUQUETTE	BLAINE VALENCIA
EVAN ENZER	MARGARET LYNCH	ARPAWAN WAEN VEJJAJIVA
ANUJ EZEKIEL	CHRISTIAN MCFALL	DEREK WEST
JOSHUA FRANK	GRACE MCFEE	CRISTINA WHITIE
LIZ FREEMAN ROSENZWEIG	MEET MEHTA	MELODY WONG
ADITI GHATLIA	MEHREEN MIR	BILLY WU
BEN GOLDFEIN	ERIN MOORE	YEXI XU
EDWIN JESUS GONZALEZ	JACKSON MORAWSKI	MEI XUAN
JAKE GORHAM	DEBBIE MOSLEY	JOSHUA YOO
ISHA GULATI	GODHULI NANDA	CHENZHAO YU
SALONI GUPTA	DAN NOEL	MICHELLE ZHOU

# BTLJ ADVISORY BOARD

JIM DEMPSEY  
*Executive Director of the  
Berkeley Center for Law & Technology*  
U.C. Berkeley School of Law

ROBERT C. BERRING, JR.  
*Walter Perry Johnson Professor of Law, Emeritus*  
U.C. Berkeley School of Law

MATTHEW D. POWERS  
Tensegrity Law Group, LLP

JESSE H. CHOPER  
*Earl Warren Professor of Public Law, Emeritus*  
U.C. Berkeley School of Law

PAMELA SAMUELSON  
*Richard M. Sherman Distinguished Professor of  
Law & Information and Faculty Director of the  
Berkeley Center for Law & Technology*  
U.C. Berkeley School of Law

REGIS MCKENNA  
*Chairman and CEO*  
Regis McKenna, Inc.

LIONEL S. SOBEL  
*Professor of Law, Emeritus and Director of the  
International Entertainment & Media Law  
Summer Program in London*  
Southwestern University School of Law

PETER S. MENELL  
*Koret Professor of Law and Faculty  
Director of the Berkeley Center  
for Law & Technology*  
U.C. Berkeley School of Law

LARRY W. SONSINI  
Wilson Sonsini Goodrich & Rosati

ROBERT P. MERGES  
*Wilson Sonsini Goodrich & Rosati Professor of  
Law and Faculty  
Director of the Berkeley Center  
for Law & Technology*  
U.C. Berkeley School of Law

MICHAEL STERN  
Cooley LLP

DEIRDRE K. MULLIGAN  
*Associate Professor and Faculty Director of the  
Berkeley Center for Law & Technology*  
U.C. Berkeley School of Information

MICHAEL TRAYNOR  
Cobalt LLP

JAMES POOLEY  
James Pooley, PLC

THOMAS F. VILLENEUVE  
Gunderson Dettmer Stough Villeneuve  
Franklin & Hachigian LLP

# BERKELEY CENTER FOR LAW & TECHNOLOGY 2019–2020

---

---

## *Executive Director*

---

JIM DEMPSEY

---

## *Faculty Directors*

---

KENNETH A.  
BAMBERGER  
CATHERINE CRUMP  
CATHERINE FISK  
CHRIS HOOFNAGLE  
SONIA KATYAL

ORIN KERR  
PETER S. MENELL  
ROBERT P. MERGES  
DEIRDRE K. MULLIGAN  
TEJAS NARECHANIA  
ANDREA ROTH  
PAMELA SAMUELSON

PAUL SCHWARTZ  
ERIK STALLMAN  
JENNIFER M. URBAN  
MOLLY S. VAN  
HOUWELING  
REBECCA WEXLER

---

## *Fellows*

---

KATHRYN HASHIMOTO

CHRISTINA KONINGISOR

---

## *Staff*

---

MARK COHEN  
NATHALIE COLETTA  
JANN DUDLEY

RICHARD FISK  
MATTHEW RAY  
IRYS SCHENKER

# DIGITAL REMEDIES

*Maayan Perel*<sup>†</sup>

## ABSTRACT

Legal disputes increasingly arise on digital grounds in relation to an array of subjects such as online enforcement of intellectual property, the First Amendment and online speech, and the right to privacy in personal data stored on digital devices. When courts are called upon to resolve disputes relating to cyberspace, many of the reliefs they grant are executed by digital means, such as technologies that restrict access to unwarranted content or technical solutions that enable or disable access to digital devices. The essence of digital remedies is their profound technological details, some of which may elude judicial review. Like equitable remedies directed to the physical world, digital remedies are usually open-ended, affording their executors broad discretion on how to implement them. However, unlike physical remedies, the implementation of digital remedies is embedded in inherently non-transparent technologies designed and executed privately outside the courthouse and has a robust, dynamic, and ongoing impact on third-party stakeholders. Digital remedies' technical details may far surpass what the court defines, converting compliance from a technical matter of law enforcement into a substantial matter of law making. Although equitable remedies generally create greater difficulties for courts in ascertaining and ensuring compliance, digital remedies take these concerns to the next level, presenting serious challenges to the rule of law.

This Article argues that the issuance and execution of digital remedies challenges the court's ability to fulfill its longstanding duty to exercise its adjudication power in accordance with rule of law, to competently prescribe remedies that are fit to redress the violation of rights, and to assure these remedies are enforced properly. Using the example of website-blocking injunctions, this Article demonstrates that the devil is in the details of implementing digital remedies. These details play a crucial role in shaping the meaning of digital remedies, and consequently the definition of the rights they purport to vindicate. Overall, the Article recommends several mechanisms that courts can exploit in order to extend their oversight and retain more control over the critical implementation stage of digital remedies. This Article builds on the system of equitable remedies, which includes, in addition to the remedy itself, equitable managerial devices that allow courts to manage the parties and ensure compliance, as well as special equitable restraints. This Article aims to empower judges who resolve cyber-related disputes with a broader and a more accurate understanding of the meaning of their digital solutions.

---

DOI: <https://doi.org/10.15779/Z38RX93D8V>

© 2020 Maayan Perel.

<sup>†</sup> Assistant Professor, Netanya Academic College; Senior Researcher, Center for Cyber Law and Policy, University of Haifa; S.J.D, University of Pennsylvania School of Law. I would like to thank Eran Bareket, Daniel Benoliel, Dan Burk, Karni Chagal, Peter Drahos, Amit Elazari, Niva Elkin-Koren, Orit Fischman-afori, Nissan Franco, Ellen Goodman, Eldar Haber, Jacob Assaf and Sharon Sandeen for their excellent comments. Special thanks are also due to the participants of the 2018 GIF Young Scientists Meeting at Potsdam, the participants of the 2018 Internet Law Scholars Conference at New York Law School and the participants of the ICIL 2018 Conference at Antwerp University for fruitful brainstorming. This research was supported by the Center for Cyber Law and Policy, University of Haifa. Any mistakes or omissions are the author's.

## TABLE OF CONTENTS

<b>I.</b>	<b>INTRODUCTION .....</b>	<b>2</b>
<b>II.</b>	<b>THE RISE OF DIGITAL RELIEFS.....</b>	<b>8</b>
A.	JUDICIAL REMEDIES .....	8
B.	THE SYSTEM OF EQUITABLE REMEDIES .....	11
C.	CLASSIFYING DIGITAL RELIEFS .....	12
<b>III.</b>	<b>DIGITAL REMEDIES: WHEN MEANS DEFINE MEANING .....</b>	<b>16</b>
A.	WEBSITE-BLOCKING INJUNCTIONS—BASIC INTRODUCTION .....	17
B.	THE SCI-HUB CASE.....	21
C.	VARIED BLOCKING MEASURES.....	23
1.	<i>IP Blocking</i> .....	24
2.	<i>Blocking Based on Deep Packet Inspection</i> .....	25
3.	<i>URL-Based Blocking</i> .....	25
4.	<i>Platform Filtering</i> .....	26
5.	<i>DNS-Based Blocking</i> .....	27
<b>IV.</b>	<b>DIGITAL REMEDIES, JUDICIAL DECISION MAKING AND THE RULE OF LAW .....</b>	<b>29</b>
A.	ROBUST IMPACT ON NUMEROUS STAKEHOLDERS.....	30
B.	DYNAMIC AND ONGOING IMPACT .....	35
C.	NON-TRANSPARENT IMPLEMENTATION ON PRIVATE GROUNDS.....	37
<b>V.</b>	<b>OVERSEEING DIGITAL REMEDIES .....</b>	<b>42</b>
A.	MANAGERIAL DEVICES .....	43
1.	<i>Ex-Post Revision</i> .....	43
2.	<i>Advising Technical Experts</i> .....	44
3.	<i>Imposing Duration Limitations</i> .....	46
4.	<i>Contempt</i> .....	47
5.	<i>Encourage Ongoing Participation of Various Stakeholders</i> .....	48
B.	EQUITABLE CONSTRAINTS.....	50
<b>VI.</b>	<b>CONCLUSION.....</b>	<b>51</b>

### I. INTRODUCTION

The impact of digital technology on regulation, law enforcement, and compliance has been investigated extensively.<sup>1</sup> Law and technology

---

1. See generally Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules through Technology*, 76 TEX. L. REV. 553 (1998) (showing that the creation and implementation of information policy are embedded in network designs and standards as well

scholarship explores how governance with the aid of technology challenges fundamental rights and democratic values, such as due process and the rule of law.<sup>2</sup> Prior work argues that the delegation of public powers to private actors using proprietary technology is black-boxed and thus difficult to oversee.<sup>3</sup> Specifically, current literature focuses on *out-of-court* delegations of public powers held by administrative actors, such as credit score providers,<sup>4</sup> regulated firms,<sup>5</sup> police,<sup>6</sup> municipal cities,<sup>7</sup> or online platforms that regulate online

---

as in system configurations); LAWRENCE LESSIG, CODE: VERSION 2.0 (2006); Kenneth A. Bamberger, *Technologies of Compliance: Risk and Regulation in a Digital Age*, 88 TEX. L. REV. 669 (2010) (describing private automated law systems that failed to recognize risks to bank capital reported, leading into global financial crisis); Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1256 (2008) (describing the Colorado Benefits Management System, which generates welfare eligibility decisions); Maayan Perel & Niva Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, 19 STAN. TECH. L. REV. 473, 477 (2016) (describing internet service provider algorithms); Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 673 (2016) (showing how algorithmic techniques like data mining challenge the prohibition of discrimination in employment).

2. See, e.g., Daniel J. Steinbock, *Data Matching, Data Mining, and Due Process*, 40 GA. L. REV. 1 (2005); TARLETON GILLESPIE, WIRED SHUT: COPYRIGHT AND THE SHAPE OF DIGITAL CULTURE 240–42 (2007); Citron, *Technological Due Process*, *supra* note 1, at 1252; Frank Pasquale, *Restoring Transparency to Automated Authority*, 9 J. TELECOMM. & HIGH TECH. L. 235, 235–36 (2011); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward A Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93 (2014); Perel & Elkin-Koren, *Accountability in Algorithmic Enforcement*, *supra* note 1; Maayan Perel & Niva Elkin-Koren, *Black Box Tinkering: Beyond Transparency in Algorithmic Enforcement*, 69 FLA. L. REV. 181 (2017); Robert Brauneis & Ellen P. Goodman, Note, *Algorithmic Transparency for the Smart City*, 20 YALE J.L. & TECH. 103, 114–15 (2018); Nicholas Diakopoulos, *We Need to Know the Algorithms the Government Uses to Make Important Decisions About Us*, CONVERSATION (May 23, 2016), <https://theconversation.com/we-need-to-know-the-algorithms-the-government-uses-to-make-important-decisions-about-us-57869> [<https://perma.cc/U37E-HHKD>].

3. FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 8 (2015); Perel & Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, *supra* note 1, at 482; Perel & Elkin-Koren, *Black Box Tinkering*, *supra* note 2, at 183.

4. See, e.g., Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 89 (2014).

5. See, e.g., Bamberger, *Technologies of Compliance*, *supra* note 1, at 673 (contending that government regulators encourage compliance through automation).

6. Walter L. Perry et al., *Predictive Policing: Forecasting Crime for Law Enforcement*, RAND (2013), [https://www.rand.org/pubs/research\\_briefs/RB9735.html](https://www.rand.org/pubs/research_briefs/RB9735.html) [<https://perma.cc/T4WZ-NJHK>].

7. See, e.g., Robert Brauneis & Ellen P. Goodman, *Algorithmic Transparency for the Smart City*, 20 YALE J.L. & TECH. 103, 107 (2018) (describing how the “smart city” movement impresses on local governments the importance of collecting and analyzing data more effectively).

speech,<sup>8</sup> to privately designed systems of automated decision-making.<sup>9</sup> Left largely unaddressed by this work, however, is the privatization of remedial powers ordinarily held by courts. These powers are often outsourced to private parties who employ digital means for compliance purposes.

Judicial remedies increasingly encompass a crucial aspect of algorithmic compliance by private actors. When courts are called upon to resolve disputes relating to cyberspace (everything that relies on interconnected technologies, such as online content or digital devices), many of the reliefs they grant depend on digital implementation by private actors. Restricting access to online content or fixing security flaws in digital devices, for instance, are all done by digital means. Nevertheless, as this Article contemplates, implementation of digital remedies is far from being solely a procedural matter of compliance. It essentially shapes the scope and breadth of the remedy and defines the practical balance between various rights and interests.

The interplay between rights and remedies has been widely explored before.<sup>10</sup> Most notable is the notion that remedies determine the efficacy of rights.<sup>11</sup> But remedies are also known for shaping the meaning of substantive law. Indeed, recent scholarship in public law highlights the importance of thinking carefully about the remedial environments from which substantive law emerges. Though varied in their evaluative approaches and prescriptive contributions, remedies law scholars agree that “remedy-related variables affect not just the intensity with which substantive rights get enforced, but also the defining of substantive rights themselves.”<sup>12</sup> This Article contributes to this discourse, contending that the technological details of implementation are a crucial variable in defining the meaning of digital remedies and the rights they vindicate. Therefore, digital remedies demand the close attention of the judiciary.

---

8. Perel & Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, *supra* note 1, at 480–81 (explaining that online intermediaries currently manage and police the usage of online content pursuant to different laws).

9. See Kenneth A. Bamberger, *Regulation as Delegation: Private Firms, Decisionmaking, and Accountability in the Administrative State*, 56 DUKE L.J. 377, 380 (2006).

10. Within individual fields, commentators have drawn attention to the linkage between remedial context and substantive law, and some commentators have proposed targeted responses to particular instances of the phenomenon. See, e.g., Douglas Laycock, *How Remedies Became a Field: A History*, 27 REV. LITIG. 161, 165 (2008); Samuel L. Bray, *The Myth of the Mild Declaratory Judgment*, 63 DUKE L.J. 1091, 1110–13 (2014); Daryl J. Levinson, *Rights Essentialism and Remedial Equilibration*, 99 COLUM. L. REV. 857, 887 (1999); Nancy Leong, *Making Rights*, 92 B.U. L. REV. 405, 421–75 (2012); Jennifer E. Laurin, *Rights Translation and Remedial Disequilibration in Constitutional Criminal Procedure*, 110 COLUM. L. REV. 1002, 1007 (2010).

11. Michael Coenen, *Spillover Across Remedies*, 98 MINN. L. REV. 1211, 1213 (2014).

12. *Id.* at 1216.



A prime example concerns the cryptographic legal battle between the FBI and Apple regarding the FBI's access to the locked iPhone of one of the San Bernardino terrorists. The FBI requested that the court force Apple to create software to help them defeat the phone's encryption by creating a technological "backdoor" that would allow the government access to the data stored not just on the suspect's device, but also on millions of Apple devices.<sup>13</sup> While the FBI eventually withdrew its motion, choosing instead to use the services of a private third party to break into the phone, a decree forcing Apple to redesign its digital devices could have had dramatic implications for U.S. residents, dissidents, and especially individuals in countries with repressive governments.<sup>14</sup> Indeed, how Apple would have practically designed this "backdoor" would affect the vulnerability of national security networks to penetration by malicious hackers, including ones from other nations.<sup>15</sup> It would have also redefined the scope of freedom of expression.<sup>16</sup> Far-reaching ramifications for collective safety and security would have also resulted from a remedy "weakening cryptography through the creation of mandatory backdoors."<sup>17</sup>

Digital remedies can have a robust impact on the rights of numerous stakeholders. In particular, the *details* of implementing digital remedies shape their substance, transforming compliance from a technical matter of law

---

13. Ron Wyden, *This Isn't About One iPhone. It's About Millions of Them*, WIRED (Feb. 19, 2016, 12:00 AM), <https://www.wired.com/2016/02/this-isnt-about-one-iphone-its-about-millions-of-them> [<https://perma.cc/5E5L-RQJT>].

14. See *Amicus Briefs in Support of Apple*, APPLE (Mar. 2, 2016), <https://www.apple.com/newsroom/2016/03/03Amicus-Briefs-in-Support-of-Apple> [<https://perma.cc/V6NM-8SAG>]; Brief of American Civil Liberties Union et al. as Amici Curiae Supporting Apple, Inc., *In re Search of an Apple iPhone*, No. CM 16-10 (C.D. Cal. Mar. 3, 2016); Brief of Privacy International and Human Rights Watch as Amici Curiae Supporting Apple, Inc., *In re Search of an Apple iPhone*, No. CM 16-10 (C.D. Cal. Mar. 3, 2016); Brief of the Center for Democracy & Technology as Amicus Curiae Supporting Apple Inc., *In re Search of an Apple iPhone*, No. CM 16-10 (C.D. Cal. Mar. 3, 2016); Letter from David Kaye, Special Rapporteur on the Promotion & Prot. of the Right to Freedom of Op. & Expression, United Nations Human Rights Council, to Hon. Sheri Pym (Mar. 2, 2016), [https://freedex.org/wp-content/blogs.dir/2015/files/2017/08/Letter\\_from\\_David\\_Kaye\\_UN\\_Special\\_Rapporteur\\_on\\_the\\_promotion\\_and\\_protection\\_of\\_the\\_right\\_to\\_freedom\\_of\\_opinion\\_and\\_expression.pdf](https://freedex.org/wp-content/blogs.dir/2015/files/2017/08/Letter_from_David_Kaye_UN_Special_Rapporteur_on_the_promotion_and_protection_of_the_right_to_freedom_of_opinion_and_expression.pdf) [<https://perma.cc/8FC5-RHVY>].

15. Apple Inc.'s Motion to Vacate Order Compelling Apple Inc. to Assist Agents in Search, and Opposition to Government's Motion to Compel Assistance, *In re Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300*, Cal. License Plate 35KGD203, No. CM 16-10 (C.D. Cal. Feb. 25, 2016).

16. Brief for International and Human Rights Watch as Amici Curiae Supporting Apple Inc., *In re Search of an Apple iPhone*, No. CM 16-10 (C.D. Cal. Mar. 3, 2016).

17. Deirdre K. Mulligan & Kenneth A. Bamberger, *Saving Governance by Design*, 106 CALIF. L. REV. 697, 725 (2018).

enforcement into a substantive matter of law making. The implementation of digital remedies requires defendants to act as both judge and executor and perform functions that are normally reserved for authorized governmental bodies.<sup>18</sup> Website-blocking injunctions demonstrate this idea perfectly, as they show how focal the technical details of the blocking technique could turn out to be. Such injunctions have been used widely in various jurisdictions throughout Europe.<sup>19</sup> The United States has also recently implemented such an injunction in a default judgment against Sci-Hub, a popular online platform for unauthorized dissemination of scientific scholarship.<sup>20</sup>

Technically, website blocking can be achieved by different means, each of which having its own special attributes. The particular implementation technique applied (whether through Internet Protocol (IP) blocking or URL blocking) ultimately shapes the boundaries of enforcement; it can actually surpass settled law, resetting the effective balance between copyright on the one hand, and free speech, privacy, and access to information on the other, while affecting the rights and interests of numerous internet users.

Overseeing how digital remedies unfold and anticipating their ultimate impact is nonetheless challenging. Remedies that compel action or inaction—that is, equitable remedies—generally create great difficulties for courts in ascertaining and ensuring compliance;<sup>21</sup> digital remedies heighten these issues. Like equitable remedies directed to the physical world, such as ordering a defendant to restore the plaintiff's property to its undamaged condition,<sup>22</sup> digital remedies leave room for flexible implementations.<sup>23</sup> Nevertheless, contrary to the evident, real-world implementation of remedies directed to the physical world, the implementation of digital remedies is generally embedded in proprietary, inherently non-transparent technologies. Additionally, predicting the ultimate reach of digital remedies in advance is extremely challenging, as their efficacy often depends on their ability to adjust promptly to the changing digital landscape. Oftentimes, they are directed to resolve an

---

18. Perel & Elkin-Koren, *Accountability in Algorithmic Enforcement*, *supra* note 1, at 485.

19. *See generally* MARTIN HUSOVEC, INJUNCTIONS AGAINST INTERMEDIARIES IN THE EUROPEAN UNION (2017).

20. *See* Am. Chem. Soc'y v. Sci-Hub, No. 1:17-cv-726 (E.D. Va. June 23, 2017).

21. Samuel L. Bray, *The System of Equitable Remedies*, 63 UCLA L. REV. 530, 564 (2016) (explaining that defendants might be recalcitrant, unsure how to comply or slow to react; that circumstances may change, and that court orders might be mistaken).

22. *See, e.g.,* Barngrover v. City of Columbus, 739 S.E.2d 377 (Ga. 2013) (describing history of an equitable remedy order in a nuisance case).

23. Bray, *supra* note 21, at 562 (“In contemporary American law the remedies that compel action or inaction are paradigmatically equitable ones. And the remedies that not only compel action or inaction, but also do so in an open-ended and less determinate fashion, are wholly equitable.”).

ongoing problem. Blocking injunctions, for instance, can soon become outdated if users and content providers conceal their online conduct by using virtual private networks (VPNs), proxy services, etc. Their efficacy largely depends on their ability to adapt to changing digital circumstances, and this further complicates the ability of courts to oversee how they evolve. But if courts cannot anticipate how digital reliefs unfold, they cannot ensure that they are actually fit to redress specific violations of rights. Ultimately, this challenges the rule of law.

The meaning of digital remedies is defined by their profound technical details which are determined and implemented outside the courthouse, on private grounds, under the veil of algorithmic opaqueness and private considerations. The execution of digital remedies, however, must not be left unchecked. Proper safeguards are necessary to preserve the rule of law and ensure that digital remedies effectively achieve their intended purpose. Otherwise, potential distortions of settled law will avoid judicial review.

Accordingly, the Article proceeds as follows. Part II provides a basic introduction of remedies law. To probe why digital remedies introduce new and intricate challenges for the judiciary, this Part describes the various goals of remedies and describes their fundamental distinctions. Following several examples, it proceeds to classify digital reliefs as specific, prospective, and equitable remedies. Part III uses the example of website-blocking injunctions to demonstrate why the digital details of implementation play such a crucial role in shaping the meaning of digital remedies, and consequently the definition of the rights they purport to vindicate. Next, whether this shift in adjudication power could be adequately dominated by the judiciary is considered in Part IV. Specifically, Part IV addresses how digital remedies challenge the ability of the court to fulfill its longstanding duty to exercise its adjudication power in accordance with the rule of law, to competently prescribe remedies that are fit to redress the violation of rights, and to ensure these remedies are enforced properly. Overall, this Part points at three attributes of digital remedy that impede their predictability. First, their ultimate meaning evolves outside the courthouse. Second, their implementation details are dynamic in their implications, costs, and capabilities of adjusting to the changing digital landscape. And third, these details are embedded in privately-developed, non-transparent codes. Finally, Part V recommends several mechanisms that courts can exploit in order to extend their oversight and retain more control over the critical implementation stage of digital remedies. Particularly, it builds on the system of equitable remedies, which includes, in

addition to the remedy itself, equitable managerial devices that allow courts to ensure compliance, as well as special equitable restraints.<sup>24</sup>

## II. THE RISE OF DIGITAL RELIEFS

To probe why digital remedies introduce new and intricate challenges for the judiciary, it is helpful first to gain a general understanding of the law of remedies. This Part explains the various goals of remedies and describes their fundamental distinctions. Following several examples, it proceeds to classify digital reliefs as specific, prospective, or equitable remedies.

### A. JUDICIAL REMEDIES

Court decisions end by either granting the plaintiff a relief or otherwise rejecting her request. “Remedies are the means by which substantive law is given its actual effect.”<sup>25</sup> Indeed, there is “no right without a remedy.”<sup>26</sup> The goals of remedies law are varied. Compensatory damages purport to restore the “plaintiff’s rightful position” through monetary transfers between plaintiff and defendant.<sup>27</sup> Preventive remedies, on the other hand, seek to avoid harm, for instance, by enjoining individuals from acting or ordering them to take affirmative steps to thwart the violation of the law.<sup>28</sup> Equitable remedies promote restitution: they are designed to deprive defendants of the benefit of wrongful acts. Remedies could also promote deterrence and morality, for instance, when courts “enhance damages beyond what is necessary to compensate plaintiffs or deprive defendants of profits in order to punish” culpable behaviors.<sup>29</sup>

A core distinction in the law of remedies is the difference between specific and substitutionary relief.<sup>30</sup> While specific reliefs afford the plaintiff the original thing to which she was entitled, substitutionary reliefs afford the plaintiff

---

24. For additional reasons, see Bray, *supra* note 21, at 534.

25. Mark A. Lemley & Bryan Casey, *Remedies for Robots*, 86 U. CHI. L. REV. 1311, 1343 (2019).

26. Frederick Pollock, *The Continuity of the Common Law*, 11 HARV. L. REV. 423, 424 (1898) (noting the phrase already functioned as a “maxim” in the 19th century).

27. DOUGLAS LAYCOCK, *MODERN AMERICAN REMEDIES* 11–15 (4th ed. 2011).

28. *Id.*

29. Lemley & Casey, *supra* note 25, at 3.

30. See, e.g., DAN B. DOBBS, *LAW OF REMEDIES: DAMAGES, EQUITY, RESTITUTION* 209 (2d ed. 1993) (distinguishing between substitutionary and specific remedies); JAMES M. FISCHER, *UNDERSTANDING REMEDIES* 4 (1999) (discussing the distinction between specific and substitutional remedies in section on “Types of Remedies”); DOUGLAS LAYCOCK, *THE DEATH OF THE IRREPARABLE INJURY RULE* 12–13 (1991) (“The most fundamental remedial choice is between substitutionary and specific remedies.”).

something that substitutes for the original thing to which she was entitled.<sup>31</sup> Money is a typical example of the latter.<sup>32</sup> Injunctions are a typical example of the former;<sup>33</sup> they are considered specific reliefs because they either direct or restrain the defendant's actions.<sup>34</sup> The idea is that an injunction, such as one ordering the defendant to stop selling counterfeit goods, intends to prevent ongoing or future violations of the plaintiff's legal entitlement (ownership of intellectual property, in this example). Similarly, mandamus, ejectment, replevin, and specific performance are also considered specific remedies because they purport to give the plaintiff the original thing or condition to which she was entitled.<sup>35</sup>

To grant a specific relief, the court must first define the borderline of the plaintiff's entitlement, or in other words, the scope of the legal right that was violated. This depends on the court's specific approach regarding the nature of the substantive law. A "normative" approach, which is consistent with laws enforced by property rules,<sup>36</sup> views the substantive law as a prohibition against certain conduct, and thus seeks to stop the wrongful act or to compensate the plaintiff for the damage done.<sup>37</sup> An "economic" approach, which is consistent with laws enforced by liability rules, holds that the substantive law "merely specifies the foreseeable consequences of various choices."<sup>38</sup> Under this approach, remedies essentially signal the costs of doing business.<sup>39</sup> When granting a substitutionary relief, courts have to evaluate the plaintiff's loss and then design a substitute equal to the value of her original entitlement.<sup>40</sup>

---

31. Colleen P. Murphy, *Money as "Specific" Remedy*, 58 ALA. L. REV. 119, 120 (2006).

32. *Id.* ("[T]he defendant has violated a legal entitlement belonging to the plaintiff—such as a personal, proprietary, dignitary, or economic entitlement—and the court awards money for the resulting harm."). Of course, money might also be a specific remedy; for instance, when the plaintiff's original entitlement is monetary (and the defendant fails to pay what he owes to the plaintiff).

33. Although injunctions could arguably be also substitutionary (for instance, when they provide a thing or condition other than the plaintiff's original entitlement). See Charles Alan Wright, *The Law of Remedies as a Social Institution*, 18 U. DETROIT L.J. 376, 378 (1955).

34. See, e.g., *Larson v. Domestic & Foreign Commerce Corp.*, 337 U.S. 682, 688 (1949).

35. Murphy, *supra* note 31, at 123.

36. See generally Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1092 (1972) ("An entitlement is protected by a property rule to the extent that someone who wishes to remove the entitlement from its holder must buy it from him in a voluntary transaction in which the value of the entitlement is agreed upon by the seller.").

37. Lemley & Casey, *supra* note 25, at 44.

38. *Id.*

39. See Ian Ayres & Eric Talley, *Solomonic Bargaining: Dividing a Legal Entitlement to Facilitate Coasean Trade*, 104 YALE L.J. 1027, 1033 (1995); see also Louis Kaplow & Steven Shavell, *Do Liability Rules Facilitate Bargaining? A Reply to Ayres and Talley*, 105 YALE L.J. 221, 222 (1995).

40. Laycock, *supra* note 30, at 13.

Another remedial distinction is between prospective and retrospective relief. Prospective relief refers to “remedies that prevent wrongful conduct or that prevent the post-judgment accrual of harms flowing from the defendant’s pre-judgment conduct.”<sup>41</sup> Retrospective relief refers to “remedies for harms that have accrued up to the date of judgment.”<sup>42</sup> Oftentimes (but not always), prospective remedies will be specific reliefs because they will usually afford the plaintiff the original thing to which she is entitled.<sup>43</sup> Retrospective reliefs, on the other hand, will usually (but again, not always) be substitutionary, namely awarding money for physical harm caused by the defendant.<sup>44</sup>

Finally, a longstanding dichotomy in remedies law, which is also the most suitable to address digital remedies as explained henceforth, is the one that differentiates between legal and equitable remedies. This historical classification is essentially evaluated by asking whether a given remedy was available in courts of law or courts of equity.<sup>45</sup> The most common remedy in the courts of law was money, whereas the most common remedy in the courts of equity was the personal order to act in a specific manner or refrain from acting in some way, such as with orders of specific performance or injunctions.<sup>46</sup> Accordingly, equitable remedies are granted “to compel action (or inaction), especially when that action may be continuing or iterative and not easily measured.”<sup>47</sup> The available equitable remedies are the injunction, specific performance, reformation, quiet title, and various “restitutionary remedies: accounting for profits, constructive trust, equitable lien, subrogation, and equitable rescission.”<sup>48</sup> The legal remedies mainly include “damages, mandamus, habeas, replevin, ejectment, and certain restitutionary remedies.”<sup>49</sup>

A standard view among American scholars is that the distinction between legal and equitable remedies is outmoded.<sup>50</sup> Modern courts treat equitable remedies as specific remedies and legal remedies as substitutionary ones,

---

41. Murphy, *supra* note 31, at 137.

42. *Id.*

43. *Id.*

44. *Id.*

45. *Id.* at 134–35.

46. *Id.* at 135.

47. Bray, *supra* note 21, at 533.

48. *Id.* at 541–42.

49. *Id.* at 542.

50. Doug Rendleman, *The Trial Judge’s Equitable Discretion Following eBay v. MercExchange*, 27 REV. LITIG. 63, 97 (2007); Caprice L. Roberts, *The Restitution Revival and the Ghosts of Equity*, 68 WASH. & LEE L. REV. 1027, 1033, 1060 (2011); *see also* James Steven Rogers, *Restitution for Wrongs and the Restatement (Third) of the Law of Restitution and Unjust Enrichment*, 42 WAKE FOREST L. REV. 55, 56 (2007) (calling distinctions between legal and equitable restitution “little short of gibberish”).

although some remedies law scholars contest that such a treatment is inaccurate.<sup>51</sup> Classifying digital remedies as equitable ones is nonetheless important because equitable remedies afford courts with special managerial tools which enable them to better manage the enforcement of equitable remedies.

#### B. THE SYSTEM OF EQUITABLE REMEDIES

Equitable remedies are not just about compelling action or inaction. A core distinction of equitable remedies relates to their open-ended and ongoing nature. This makes them far less determinate than other outcome-specific, one-shot remedies, such as damages. Consequently, equitable remedies may give rise to a serious problem of compliance.<sup>52</sup> Specifically,

[s]ome defendants will be recalcitrant, refusing to comply. Others will be ignorant or unsure exactly how to comply. Still others may slow their pace, dragging things out, even if they would not refuse a clear order. Nor does the fault always lie with the defendant. There will be circumstances that the court could not foresee, or at least did not foresee, when it gave the order compelling action or inaction. There will be judicial mistakes, impossibilities, and absurdities.<sup>53</sup>

While assessing compliance with legal remedies is rather straightforward—the actual payment of damages, the moment a prisoner is released from custody, or when property is being replevied and returned—it could be relatively challenging to determine full compliance with equitable remedies. For example, prohibiting a former employee of a pizza parlor from “using, divulging, and communicating to anyone else any of the trade secrets or confidential information” about the pizza parlor’s sauce requires ongoing avoidance from the part of the former employee.<sup>54</sup> Whether this injunction is fully complied with or not largely depends on the degree and scope of the employee’s cooperation.

The law of equitable remedies, hence, offers a mechanism for managing compliance. This mechanism includes several managerial doctrines that improve the courts’ ability to ensure better enforcement of equitable remedies. Part V discusses these doctrines in breadth; thus, for now it is sufficient to

---

51. Murphy, *supra* note 31, at 135.

52. Bray, *supra* note 21, at 563.

53. *Id.*

54. 205 Corp. v. Brandow, 517 N.W.2d 548, 552 (Iowa 1994). For more examples, see Bray, *supra* note 21, at 563–64.

mention them generally: (1) ex-post revision; (2) contempt; (3) equitable helpers; (4) flexibility; and (5) judicial decision-making.<sup>55</sup>

The exploitation of these managerial devices, especially ex-post revision, contempt, and equitable helpers, can be notably costly. Indeed, “the direct and indirect costs of complying with the court’s command and the possibility of an afterlife in which that command is clarified, modified, enforced, or dissolved” could be substantial.<sup>56</sup> Therefore, the system of equitable remedies also provides safety valves that purport to prevent their misapplication. These include the doctrine of ripeness, requirements for specificity, and the equitable defenses of laches and unclean hands that are available for defendants.<sup>57</sup> A detailed discussion of these constraining measures and their application to digital remedies is provided in Part VI.

### C. CLASSIFYING DIGITAL RELIEFS

Aspects of our everyday conduct are increasingly becoming digital.<sup>58</sup> Technology is embedded so deeply in human lives that in many cases, there is no other way to govern human behavior than to interact with the technologies that shape it.<sup>59</sup> Criminal enforcement, for example, often depends on the police having access to digitally stored data;<sup>60</sup> preventing terrorists from unleashing terror depends heavily on online intermediaries monitoring inciting content;<sup>61</sup> data security builds on applications’ developers addressing security flows in their smart devices.<sup>62</sup>

---

55. See *infra* Part V.

56. Bray, *supra* note 21, at 577.

57. *Id.* at 578–86.

58. Rob Kitchin, *Thinking Critically About and Researching Algorithms* 7 (The Programmable City, Working Paper No. 5, 2014), <http://ssrn.com/abstract=2515786> [<https://perma.cc/4ZAW-U55G>]; Jeff Fuhrman, *The Personalization and Optimization of the Internet of Things*, ADOBE BLOG (July 14, 2015), <https://theblog.adobe.com/the-personalization-and-optimization-of-the-internet-of-things/> [<https://perma.cc/FA7J-J222>].

59. Mulligan & Bamberger, *Saving Governance by Design*, *supra* note 17, at 701.

60. For a comparative analysis about government’s access to personal data, see generally Ira S. Rubinstein, *Systematic Government Access to Personal Data: a Comparative Analysis*, 4 INT’L DATA PRIVACY L. 96 (2014).

61. Jen Kirby, *Zuckerberg: Facebook Has Systems to Stop Hate Speech. Myanmar Groups: No, it Doesn’t.*, VOX (Apr. 6, 2018), <https://www.vox.com/2018/4/6/17204324/zuckerberg-facebook-myanmar-rohingya-hate-speech-open-letter> [<https://perma.cc/CK9K-GC93>].

62. Fed. Trade Comm’n v. D-Link Corp., No. 3:17-CV-00039 (N.D. Cal. Sept. 19, 2017) (bringing request for permanent injunction and other equitable relief). The Federal Trade Commission (FTC) brought this complaint against a Taiwanese corporation, D-Link, which develops and sells, among other things, IP cameras that enable customers to monitor private areas of their homes or business. FTC’s basic argument is that D-Link has failed to take reasonable steps to protect their routers and IP cameras from widely known and reasonably



With technology playing such a central role in our lives, it is not surprising that many legal disputes, in various legal contexts, including intellectual property, First Amendment and online speech,<sup>63</sup> the right to privacy in personal data,<sup>64</sup> and the right to non-discrimination,<sup>65</sup> are cyber-related, and hence largely dependent on digital resolution. As such, digital reliefs can only be enforced by digital means, although their implications may extend to the physical world, as well.<sup>66</sup> Digital reliefs typically take the form of injunctions and therefore they could be generally characterized as specific, prospective remedies. Most often they are open-ended, setting an ongoing outcome which may be achieved through various digital means. Thus, they could also fall neatly into the category of equitable remedies. What is it, then, that makes them different? To answer this question, let's explore two examples of digital remedies.

TickBox TV, LLC was a distributor of a small Roku-style device that allows users to perform many computer functions on their television set or other monitor, including browsing the internet and streaming media content through various applications that are preloaded by TickBox or later downloaded by users. In a complaint filed by prominent copyright holders in the motion picture industry, plaintiffs alleged that the device's user interface contained

---

foreseeable software security flaws, and by failing to do so it violated section 5(a) of the FTC Act, which prohibits "unfair or deceptive acts or practices in or affecting commerce." *Id.* The FTC requires the court to enter a permanent injunction to prevent future violations of the FTC Act by D-Link. This case is still standing in front of a district court in California, but to the extent that the court will grant the order requested, it is possible that it will require D-Link to take technological steps to address the security flaws identified by the FTC.

63. *Sandvig v. Sessions*, 315 F. Supp. 3d 1, 8 (D.D.C. 2018) (arguing that the access section of the Computer Fraud and Abuse Act (CFAA) would criminalize a group of researchers' research activities, which are conducted as a response to new trends in real estate, finance, and employment transactions, which increasingly have been initiated on the internet. As part of their research activities, they wish to find out if automated transactions in these fields are discriminatory. One way to determine whether members of protected classes are being discriminated against is to engage in "outcomes-based audit testing," which involves accessing a website or other network service repeatedly, generally by creating false or artificial user profiles, to see how websites respond to users who display characteristics attributed to certain classes. These activities will violate certain website Terms of Service, and hence could amount to unauthorized access to a computer, violating the CFAA).

64. *hiQ Labs, Inc. v. LinkedIn Corp.*, 938 F.3d 985 (9th Cir. 2019). This case, which was recently decided in favor of hiQ, involved the use of bots by hiQ to scrape data from LinkedIn website in order to create services that alerts employers about their employees' online activity. LinkedIn argued that hiQ violated the privacy of its users, but the court of appeals affirmed the district court's preliminary injunction prohibiting LinkedIn from blocking hiQ from accessing its website.

65. *See Sandvig*, 315 F. Supp. 3d at 8–9.

66. For instance, a court order requiring a supplier of digital home cameras to address specific security flaws and make his camera more protected may reduce house break-ins.

links to applications that provided access to unauthorized streaming versions of their copyrighted works.<sup>67</sup>

In its initial order issued on January 30, 2018, the California Central District Court ordered TickBox to maintain the current version of its software, which had the pre-loaded infringing applications removed. Additionally, the court refused to order TickBox to remove the already-downloaded offending applications from its users' devices, explaining that such an order raised outstanding questions that had to be answered by the parties. Interestingly, the court directed its outstanding questions to the parties, ordering them to "negotiate and attempt to reach agreement upon a stipulated preliminary injunction that will supersede the Court's initial preliminary injunction order."<sup>68</sup>

Subsequently, on February 13, 2018, the court granted another order in the case:

TickBox shall issue an update to the TickBox launcher software to be automatically downloaded and installed onto any previously distributed TickBox TV device and to be launched when such device connects to the internet. Upon being launched, the update will delete the Subject Software downloaded onto the device prior to the update, or otherwise cause the TickBox TV device to be unable to access any Subject Software downloaded onto or accessed via that device prior to the update.<sup>69</sup>

Ordering TickBox to perform a software update that removes all pre-loaded applications from its users' devices is a digital remedy. It is an open-ended injunction which sets a specific, prospective outcome to be achieved—that TickBox's launcher software will not include or provide applications that link to copyright-infringing websites—but without imposing limitations on the digital means for achieving this outcome. As stressed in the second order, a software update that achieves the desired outcome may either *delete* the problematic apps or *block* the devices' access to these apps. As expanded in Parts III and IV, restricting users' access to content can be accomplished by varied technological means that differ in their cost, scope, and accuracy. Placing such broad discretion to choose how to block the devices' access to allegedly infringing apps in the hands of a private, profit-maximizing defendant

---

67. Universal City Studios Prods. L.L.P. v. TickBox TV L.L.C., No. CV 17-7496-MWF (ASX) (C.D. Cal. Oct. 13, 2017).

68. Universal City Studios Prods. L.L.P. v. TickBox TV L.L.C., No. CV 17-7496-MWF (ASX) (C.D. Cal. Jan. 30, 2018) [hereinafter TickBox 1].

69. Universal City Studios Prods. L.L.P. v. TickBox TV L.L.C., No. 2:17-cv-07496-MWF (AS) (C.D. Cal. Feb. 13, 2018) [hereinafter Tickbox 2].

makes it difficult for the court to ensure that the relief, as it ultimately unfolds, is adequately tailored to redress the infringement of plaintiff's rights.

The famous battle between the FBI and Apple presents another interesting example of digital remedies. Following the massacre of fourteen people in California at San Bernardino's Inland Regional Center in December 2015, the FBI sought access to the murderer's iPhone. Apple refused to assist the FBI in breaking into the locked phone, so the FBI sought the court's intervention.<sup>70</sup> Relying on the ancient All Writs Act,<sup>71</sup> Magistrate Judge Sheri Pym of the Central District of California issued an order compelling Apple to assist law enforcement agents in decrypting the locked phone.<sup>72</sup> Interestingly, Judge Pym also set forth a recommended technological roadmap describing the specific steps to be taken in order to achieve this outcome.<sup>73</sup> At the same time, the judge allowed Apple to use "alternate technological means from that recommended by the government," as long as the government concurred and these means achieved the functions designated in the order, as well as the functionality described in the technological roadmap provided by the court.<sup>74</sup> In the end, the FBI did not have to enforce this order because a private, external technology company successfully circumvented Apple's security lock and enabled access into the iPhone.<sup>75</sup> Nevertheless, this order remains an excellent example of a digital remedy that is far more specific in its language, although it still remains open-ended in its nature.

These two examples demonstrate that digital reliefs are essentially remedies that compel a specified digital outcome, and therefore they could be generally classified as specific, prospective, and equitable remedies. Digital reliefs are open-ended in varied degrees, leaving the issue of implementation to the defendant's discretion. But this is not new in the realm of equitable remedies. In fact, employing privately-developed technology to redress violations of individual rights is quite prevalent, especially in the areas of environmental law

---

70. Government's Motion to Compel Apple Inc. to Comply with this Court's February 16, 2016 Order Compelling Assistance in Search at 16–18, *In re Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, Cal. License Plate 35KGD203*, No. CM 10-16 (C.D. Cal. Feb. 19, 2016).

71. All Writs Act, 28 U.S.C. § 1651(a) (2012).

72. Order Compelling Apple Inc. to Assist Agents in Search, No. ED 15-0451M, *In re Search of an Apple iPhone* (C.D. Cal. Feb. 16, 2016).

73. *Id.* at 3, 4.

74. *Id.*

75. Laura Hautala & Shara Tibken, *FBI to Apple: We Don't Need your iPhone Hack*, CNET (Mar. 21, 2016), <https://www.cnet.com/news/fbi-v-apple-we-dont-need-your-iphone-hack/> [<https://perma.cc/UYG9-PT8K>].

and consumer protection law.<sup>76</sup> What makes digital reliefs different relates to the characteristics and merits of the *digital means* that implement them: as explained and demonstrated henceforth, the digital means executed to implement digital remedies effectively shape their substantial meaning. With digital remedies, *implementation* defines the scope and breadth of the remedies in an incomparable way. Using the example of website-blocking injunctions, the following Part shows how the implementation of digital remedies is far beyond a technical issue of compliance, and therefore should not be left to out-of-court, unchecked management.

### III. DIGITAL REMEDIES: WHEN MEANS DEFINE MEANING

In adjudicating claims for relief, courts often proceed in two stages. First, they determine whether a violation of the law has occurred. If so, they next decide whether to grant the requested relief.<sup>77</sup> Formally, these stages are separated. That is, the law of remedies operates independently of the substantive law.<sup>78</sup> Practically, however, remedial law often interacts with rights-based law in many respects. One important interaction relates to enforcement: a right without a remedy is “existent and identifiable, but of limited practical use to its purported beneficiaries.”<sup>79</sup> Furthermore, remedial law may shape the meaning of the substantive law: for instance, when a court’s ruling on the merits stems from the way it anticipates “the remedial consequences of a legal violation.”<sup>80</sup> Additionally, remedies may affect the incentives of litigants to advance particular substantive claims, or “trigger cognitive biases within the judges evaluating these claims.”<sup>81</sup>

It is not surprising, then, that this right-remedy interdependence attracts the attention of public law scholars.<sup>82</sup> Underlying this scholarship is “the basic premise that remedy-related variables affect not just the intensity with which substantive rights get enforced, but also the defining of substantive rights themselves.”<sup>83</sup> Digital remedies take this premise several steps forward: they

---

76. For instance, where defendants are required to reduce their polluting disposals; or where product developers are compelled to make their products safer.

77. See, e.g., *Marbury v. Madison*, 5 U.S. 137, 154 (1803) (asking first, “[h]as the applicant a right to the commission he demands?” and asking second, “[i]f he has a right, and that right has been violated, do the laws of his country afford him a remedy?”).

78. Coenen, *supra* note 11, at 1213.

79. *Id.*

80. *Id.* at 1213–14.

81. *Id.* at 1215.

82. See *supra* note 10.

83. Coenen, *supra* note 11, at 1216.

show that it is not only the prescription of remedies that impacts the substantive law, it is also—and often more so—the subsequent, out-of-court implementation of digital remedies. Digital reliefs can be implemented through various means that differ very substantially from one another: the *details* of implementation shape the *substance* of the remedy and determine its impact on numerous stakeholders in an unprecedented way.

The following discussion uses website-blocking injunctions to demonstrate why the digital details of implementation play such a crucial role in shaping the meaning of digital remedies, and consequently the definition of the rights they purport to vindicate. In fact, the need to choose between significantly different enforcement methods transforms implementation into an issue of law-making. Whether this shift in adjudication power could be adequately overseen by the judiciary demands careful consideration, as subsequently explained in Part IV.

#### A. WEBSITE-BLOCKING INJUNCTIONS—BASIC INTRODUCTION

A major objective of cyberlaw is to regulate illegal content online.<sup>84</sup> One of the greatest challenges in this respect is to ensure prompt and efficient enforcement where acting directly against the primary speakers “has proven to be ‘heavy-handed, disproportionate, and ineffective.’”<sup>85</sup> Indeed, direct users often conceal their identity behind anonymous user names, complicating the ability to act directly against them.<sup>86</sup> Additionally, illegal content may originate from places outside of the jurisdiction’s reach, further complicating enforcement.<sup>87</sup> While “bringing actions against individual users is expensive . . . regulating access via intermediaries is more cost-effective.”<sup>88</sup> Therefore, “the liability of [i]nternet intermediaries, particularly Internet Service Providers

---

84. See, e.g., HANNIBAL TRAVIS, *CYBERSPACE LAW: CENSORSHIP AND REGULATION OF THE INTERNET* (2013).

85. Christophe Geiger & Elena Izyumenko, *The Role of Human Rights in Copyright Enforcement Online*, 32 AM. U. INT’L L. REV. 43, 44 (2016).

86. *Id.*

87. See, e.g., *Discussion Paper: Roles and Responsibilities of Intermediaries: Fighting Counterfeiting and Piracy in the Supply Chain*, ICC BUSINESS ACTION TO STOP COUNTERFEITING AND PIRACY 74 (Mar. 2015), <https://iccwbo.org/content/uploads/sites/3/2015/03/ICC-BASCAP-Roles-and-Responsibilities-of-Intermediaries.pdf> [<https://perma.cc/T376-AUME>] (noting that “[o]ne of the main challenges is addressing both counterfeiting and piracy from websites based outside the jurisdiction in which the infringement takes place”).

88. David Lindsay, *Website Blocking Injunctions to Prevent Copyright Infringement: Proportionality and Effectiveness*, 40 U. NEW S. WALES L.J. 1507, 1507 (2017); Jonathan Zittrain, *Internet Points of Control*, 44 B.C. L. REV. 653, 662 (2003).

(“ISPs”), for the unlawful online actions of third party users is a persistent theme” of the content moderation discourse.<sup>89</sup>

Indeed, online intermediaries are becoming a focal point of content moderation.<sup>90</sup> They may enable or disable access by removing or blocking controversial content, or by terminating users’ accounts altogether. “Imposing liability on intermediaries can, however, have significant unwelcome effects, or ‘collateral damage,’ especially on the rights to freedom of expression and privacy of end-users.”<sup>91</sup> Indeed, making platforms legally liable for content posted by users could chill free speech and stifle the development of the internet industry.<sup>92</sup>

The most recent addition to intermediary liability law is the prerogative to award injunctions against intermediaries to block internet access (that is, use digital means) in order to prevent online infringements of intellectual property rights. Such injunctions are directed to private intermediaries that are not direct parties to the legal dispute, but presumably have the technological ability to resolve it.<sup>93</sup> They have been used, quite extensively, in Europe.<sup>94</sup> In *Google Inc.*

89. *Id.* at 1507.

90. Niva Elkin-Koren & Maayan Perel, *Guarding the Guardians: Content Moderation by Online Intermediaries and the Rule of Law*, OXFORD HANDBOOK OF INTERMEDIARY LIABILITY ONLINE (Apr. 2019), <https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780190900571.001.0001/oxfordhb-9780190900571-e-9> [https://perma.cc/U6BK-HRN8].

91. Lindsay, *supra* note 88, at 1507.

92. Zeran v. AOL, Inc., 129 F.3d 327, 331, 335 (4th Cir. 1997); Niva Elkin-Koren, *After Twenty Years: Revisiting Copyright Liability of Online Intermediaries*, in THE EVOLUTION AND EQUILIBRIUM OF COPYRIGHT IN THE DIGITAL AGE 29 (Susy Frankel & Daniel J Gervais eds., 2014).

93. MARTIN HUSOVEC, INJUNCTIONS AGAINST INTERMEDIARIES IN THE EUROPEAN UNION (2017).

94. Geiger & Izyumenko, *supra* note 85, at n. 65; see, e.g., Althaf Marsoof, *The Blocking Injunction—A Critical Review of Its Implementation in the United Kingdom within the Legal Framework of the European Union*, 46 INT’L REV. IP & COMPETITION L. 632, 656 (2015). See, for example, in the UK: Twentieth Century Fox Film Corp. & Ors v. British Telecomms. Plc [2011] EWHC 1981 (Ch); EMI Records Ltd. & Ors v. British Sky Broad. Ltd. & Ors [2013] EWHC 379 (Ch); Cartier Int’l AG & Ors v. British Sky Broad. Ltd. & Ors [2014] EWHC 3354 (Ch); Cartier Int’l Ltd. & Anor v. British Telecomms. Plc & Ors [2016] EWHC 339 (Ch). See, for example, in Denmark: Maritime and Commercial Court in Copenhagen, Fritz Hansen A/S and Others v. Telia Danmark, no. A-38-14, transcript from the record of judgments, p. 10 (Dec. 11, 2014), <http://kluwercopyrightblog.com/wp-content/uploads/sites/49/2015/01/IA11122014EN.pdf> [https://perma.cc/9GMC-BRQ3]. See, for example, in Germany: German Federal Supreme Court of Justice (Bundesgerichtshof), I ZR 3/14, 26 November 2015, DE:BGH:2015:261115UIZR3.14.0. For examples in France, see SCPP v. Orange, High Court of Paris (Tribunal de Grande Instance de Paris), 3rd chamber, Free, SFR et Bouygues Télécom, no. 14/03236, at 7 (Dec. 4, 2014), [http://www.legalis.net/spip.php?page=jurisprudence\\_decision&id\\_article=4386](http://www.legalis.net/spip.php?page=jurisprudence_decision&id_article=4386) [https://perma.cc/ZD7C-AX4V] [French]; CJEU, Judgment in

*v Equustek Solutions Inc.*, the Canadian Supreme Court held that it had power, under its general equitable jurisdiction, to grant an injunction against Google, a non-party to the underlying action, to cease indexing or referencing search results that would provide access to a website involved in intellectual property infringement.<sup>95</sup>

In the United States, however, website blocking seems to clash with the deeply rooted regime of safe harbor. In the early days of the internet, online companies and policymakers feared that making platforms legally liable for content posted by users would chill free speech and stifle the development of the internet. Hence, to mitigate such a threat, legislatures limited the liability of sites that hosted digital content for harm caused by their users (safe harbor). The safe harbor provisions of the Digital Millennium Copyright Act (DMCA)<sup>96</sup> and Section 230 of the Communications Decency Act<sup>97</sup> are intended to protect the democratic nature of the internet and prompt diversity and participation in the online sphere. They are still considered by many as “the most influential law[s] to protect the kind of innovation that has allowed the [i]nternet to thrive . . . .”<sup>98</sup> Accordingly, intermediaries are free to facilitate users’ exchange

---

UPC Telekabel Wien, C-314/12, EU:C:2014:192 (Mar. 27, 2014); ECtHR, *Akdeniz v. Turkey* (dec.), no. 20877/10 (Mar. 11, 2014).

95. *Google Inc. v. Equustek Solutions Inc.*, [2017] 1 S.C.R. 824 (Can.). In this landmark decision released recently by the Supreme Court of Canada, the court upheld the lower courts’ decision ordering Google to de-index all websites selling goods that violated a Canadian company’s trade secrets worldwide. *Id.* Equustek is a small Canadian technology company whose intellectual property was infringed by Datalink, a former distributor of Equustek’s products. *Id.* Equustek brought an action against Datalink and obtained court orders prohibiting the sale of inventory and the use of Equustek’s intellectual property. *Id.* Nevertheless, Datalink left Canada and continued offering the infringing products from an unknown location. *Id.* Google had subsequently de-indexed 345 specific webpages associated with Datalink; however, since it did not de-index entire websites and it limited the de-indexing to searches conducted on google.ca, this voluntary step was ineffective. *Id.* Datalink simply moved the objectionable content to new pages within its websites, circumventing the court orders. *Id.* As a result, Equustek obtained an interlocutory injunction to enjoin Google from displaying any part of Datalink’s websites on any of its search results worldwide. *Id.* Subsequently, the U.S. District Court of Northern California granted Google a temporary injunction blocking the enforceability of the Supreme Court of Canada’s order in the United States, reasoning that Google was protected as a neutral intermediary under Section 230 of the Communications Decency Act 1996. *Google L.L.C. v. Equustek Sols. Inc.*, No. 5:17-CV-04207-EJD (N.D. Cal. Nov. 2, 2017).

96. 17 U.S.C. § 512(a)–(d), (i).

97. 47 U.S.C. § 230.

98. *CDA 230: The Most Important Law Protecting Internet Speech*, ELECTRONIC FRONTIER FOUND., <https://www EFF.ORG/issues/cda230> [<https://perma.cc/69S2-MD6S>] (last visited Jan. 4, 2020); accord Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2313 (2014) (“Section 230 immunity and, to a lesser extent, § 512 safe harbors have been

of information without worrying about exposing themselves and their investors to legal risks, and this might include content-blocking obligations.<sup>99</sup>

In relation to intellectual property-related blockings, two anti-piracy bills introduced in 2011, the Stop Online Piracy Act (SOPA)<sup>100</sup> and its Senate counterpart, the Protect IP Act (PIPA),<sup>101</sup> which would purportedly enable courts to issue blocking orders against blacklisted pirate websites, were successfully defeated, following a powerful public protest.<sup>102</sup> The core argument raised by the bills' opponents was that affording law enforcement agents with unprecedented power to create blacklists of illegitimate websites and request the court to compel various internet services to censor them, even though no court had previously found that these services infringed copyright, would disproportionately chill protected speech, given that laws and procedures are already in place for taking down infringing websites.<sup>103</sup>

Nevertheless, a recent case decided by a Virginia district court, *ACS v. Sci-Hub*,<sup>104</sup> may signal a shift in the judiciary's attitude to website blocking.<sup>105</sup> The next Section provides a brief description of the dispute, followed by a discussion of the digital relief granted.

---

among the most important protections of free expression in the United States in the digital age.”); David Post, *A Bit of Internet History, or How Two Members of Congress Helped Create a Trillion or So Dollars of Value*, WASH. POST: VOLOKH CONSPIRACY (Aug. 27, 2015), <http://wapo.st/1K9AmTh> [<https://perma.cc/8253-LYWL>].

99. *Hassell v. Bird*, 420 P.3d 776, 778 (Cal. 2018) (ruling that Yelp cannot be forced to remove a review posted on its website since such a removal order improperly treats Yelp as the publisher or speaker of information provided by another information content provider).

100. Stop Online Piracy Act of 2011, H.R. 3261, 112th Cong. (2011).

101. Preventing Real Online Threats to Economic Creativity and Theft of Intellectual Property (Protect IP) Act of 2011, S. 968, 112th Cong. (2011).

102. Yafit Lev-Aretz, *Copyright Lawmaking and Public Choice: From Legislative Battles to Private Ordering*, 27 HARV. J.L. & TECH. 203, 204–07 (2013).

103. *SOPA/PIPA: Internet Blacklist Legislation*, ELECTRONIC FRONTIER FOUND., <https://www.eff.org/issues/coica-internet-censorship-and-copyright-bill> [<https://perma.cc/PS9P-D5KU>] (last visited Jan. 4, 2020).

104. Proposed Findings of Fact and Recommendations, *Am. Chem. Soc’y v. Sci Hub*, No. 1:17-cv-0726-LMB-JFA (E.D. Va. Sept. 28, 2017) [hereinafter Magistrate Judge’s Proposed Findings].

105. See Mitch Stoltz, *Another Court Overreaches With Site-Blocking Order Targeting Sci-Hub*, ELECTRONIC FRONTIER FOUND. (Nov. 10, 2017), <https://www.eff.org/deeplinks/2017/11/another-court-overreaches-site-blocking-order-targeting-sci-hub> [<https://perma.cc/BR9K-CUBX>].



## B. THE SCI-HUB CASE

Sci-Hub is a well-known website that makes research papers that are normally behind paywalls free to access.<sup>106</sup> Sci-Hub states that its mission is to provide “free access to scientific literature,” hosting “more than 58 million peer-reviewed scientific articles for free download.”<sup>107</sup> According to a recent study, Sci-Hub provides greater coverage of toll access scholarly articles than the University of Pennsylvania.<sup>108</sup> On June 23, 2017 the American Chemical Society (ACS) sued Sci-Hub for copyright and trademark infringement. ACS contended that “in order to lure users to its illegitimate sources of the Society’s stolen content, Sci-Hub conspirators most recently created ‘spoofed’ websites that mirror the look and feel of the Society’s own scientific publishing website.”<sup>109</sup>

As happened in a previous copyright suit brought against Sci-Hub,<sup>110</sup> the person behind the website, Alexandra Elbakyan, who operated the site out of Russia using various domain names and IP addresses, did not appear to defend Sci-Hub in court.<sup>111</sup> The Computer & Communications Industry Association (CCIA),<sup>112</sup> however, submitted a brief as amicus curiae, objecting to some portion of the injunction sought by ACS.<sup>113</sup> On November 3, 2017, the court

---

106. SCIENCE HUB, <https://sci-hub.tw/> [<https://perma.cc/VB8P-R3LJ>] (last visited Jan. 4, 2020).

107. Magistrate Judge’s Proposed Findings, *supra* note 104.

108. Daniel S. Himmelstein et al., *Research: Sci-Hub Provides Access to Nearly All Scholarly Literature*, ELIFE (Feb. 9, 2018), <https://doi.org/10.7554/eLife.32822> [<https://perma.cc/AHX6-A25R>].

109. *American Chemical Society Files Suit Against Sci-Hub*, AM. CHEMICAL SOC’Y (June 28, 2017), <https://www.acs.org/content/acs/en/pressroom/newsreleases/2017/june/acs-files-suit-against-sci-hub.html> [<https://perma.cc/C3GL-DZXH>].

110. Quirin Schiermeier, *US Court Grants Elsevier Millions in Damages from Sci-Hub*, NATURE (June 22, 2017), <https://www.nature.com/news/us-court-grants-elsevier-millions-in-damages-from-sci-hub-1.22196> [<https://perma.cc/WST8-CVL2>].

111. Diana Kwon, *American Chemical Society Wins Lawsuit Against Sci-Hub*, SCIENTIST (Nov. 7, 2017), <https://www.the-scientist.com/news-opinion/american-chemical-society-wins-law-suit-against-sci-hub-30648> [<https://perma.cc/A8JM-S9D3>].

112. The CCIA represents more than twenty large, medium, and small companies in the high technology products and services sectors, including computer hardware and software, electronic commerce, telecommunications, and internet products and services—companies that provide online services to billions of people around the world.

113. CCIA urged the court to reject the Magistrate Judge’s recommendation, insofar as it would extend a permanent injunction in this case to online intermediaries that are not direct parties to the dispute, including internet search engines, web hosting services, and ISPs and require them to “cease facilitating access to any or all domain names and websites through which Defendants engage in unlawful access to, use, reproduction, and distribution of the ACS Marks or ACS’s Copyrighted Works.” Brief of CCIA as Amicus Curiae Supporting

issued a default judgment ordering Sci-Hub to stop distributing ACS content and imitating its trademark. Furthermore, the court also ruled that

any person or entity in privity with Sci-Hub and with notice of the injunction, including any Internet search engines, web hosting and Internet service providers, domain name registrars, and domain name registries, cease facilitating access to any or all domain names and websites through which Sci-Hub engages in unlawful access to, use, reproduction, and distribution of the ACS's trademarks or copyrighted works.<sup>114</sup>

Additionally, ACS was awarded \$4.8 million in damages.<sup>115</sup>

Such a broad, open-ended injunction is most exceptional in the landscape of remedies law.<sup>116</sup> Opponents of this injunction argued that requiring third parties to censor a pirate website may over-burden innocent actors, who merely provide basic services without encouraging illegal activity.<sup>117</sup> On a procedural level, this may overstep the limits of Rule 65 of the Federal Rules of Civil Procedure, which is extremely strict regarding the specific circumstances under which non-parties to a legal dispute may be enjoined.<sup>118</sup> Indeed, the main argument of CCIA in its amicus brief was that the broad language of the injunction could “sweep in various Neutral Service Providers, despite their having violated no laws and having no connection to this case,”<sup>119</sup> without giving them an opportunity to be heard as required under due process.<sup>120</sup>

---

Objections to Magistrate Judge's Proposed Findings of Fact and Recommendations at 1, *Am. Chem. Soc'y v. Sci-Hub*, No. 1:17-cv-0726-LMB-JFA (E.D. Va. Oct. 12, 2017) [hereinafter *CCIA Amicus Brief*].

114. Magistrate Judge's Proposed Findings, *supra* note 104, at 14–15.

115. *Am. Chem. Soc'y v. Sci-Hub*, No. 1:17-cv-726-LMB-JFA (E.D. Va. Oct. 12, 2017).

116. Diana Kwon, *Judge Recommends Ruling to Block Internet Access to Sci-Hub*, *SCIENTIST* (Oct. 4, 2017), <https://www.the-scientist.com/daily-news/judge-recommends-ruling-to-block-internet-access-to-sci-hub-30793> [<https://perma.cc/S5QR-5BYM>].

117. *See Stoltz, supra* note 105.

118. According to FED. R. CIV. P. 65(d)(2), “The order binds only the following who receive actual notice of it by personal service or otherwise: . . . (c) other persons who are in active concert or participation with anyone described in Rule 65(d)(2)(A) or (B).” *See CCIA Amicus Brief, supra* note 113, at 1.

119. *Id.* at 2.

120. *Id.* at 4. Courts have long interpreted this rule narrowly, explaining that “the only occasion when a person not a party may be punished, is when he has helped to bring about, not merely what the decree has forbidden, because it may have gone too far, but what it has power to forbid, an act of a party.” *Alemite Mfg. Corp. v. Staff*, 42 F.2d 832, 833 (2d Cir. 1930); *New York v. Operation Rescue Nat'l*, 80 F.3d 64, 70 (2d Cir. 1996); *Haizlip v. Alston*, No. 1:14CV770, 2015 WL 8668230, at \*1 (M.D.N.C. Dec. 11, 2015). In other words, an

However, the uncertainty surrounding this order is not just about to *whom* it applies. *How* this order will be effectively implemented (insofar as the ACS specifically enforces it) and *what* its actual impact on ACS's intellectual property and the public interest in access to knowledge is, also remain unknown.<sup>121</sup> As demonstrated henceforth, different digital measures could be applied to disable access to allegedly infringing websites. These means vary substantially in their costs of implementation, accuracy, and efficiency (potency against circumvention). However, these differences between the varied blocking measures effectively define the scope and breadth of blocking: the more accurate and potent the blocking is, the narrower is the remedy, and vice versa. Of course, the scope and breadth of the remedy, which stem from the specific blocking measure applied, further define the ultimate balancing between the competing rights and interests. These are the rights-holders' intellectual property rights, on the one hand, and third parties' free speech and access to information, on the other.<sup>122</sup> The following discussion briefly explains the differences between major blocking techniques to elaborate this point.

### C. VARIED BLOCKING MEASURES

Access to websites may be blocked by various technological means that differ in their technical and policy limitations, as well as in their consequences. In March 2017, the Internet Society—an international organization whose vision is “to promote the development of the Internet as a global technical infrastructure,”<sup>123</sup> published an overview of internet content blocking, which relies on public policy considerations.<sup>124</sup> The overview offers “a technical assessment of the benefits and drawbacks of the most common blocking techniques used to prevent access to content deemed illegal,” in order “to help readers understand what each technique can, and cannot, block, along with the

---

injunction may not “make punishable the conduct of persons who act independently and whose rights have not been adjudged according to law.” *Regal Knitwear Co. v. NLRB*, 324 U.S. 9, 13 (1945).

121. Andrew Silver, *Sci-Hub Domains Inactive Following Court Order: 'Free science'/Pirate Site Operator 'working on solving DNS issue'*, REGISTER (Nov. 23, 2017), [https://www.theregister.co.uk/2017/11/23/sci\\_hubs\\_become\\_inactive\\_following\\_court\\_order/](https://www.theregister.co.uk/2017/11/23/sci_hubs_become_inactive_following_court_order/) [<https://perma.cc/L7FB-LPXT>].

122. See *infra* Part IV.

123. *Our Mission*, INTERNET SOC'Y, <https://www.internetsociety.org/mission/> [<https://perma.cc/5DTR-4ZBN>] (last visited Jan. 4, 2020).

124. *Internet Society Perspectives on Internet Content Blocking: An Overview*, INTERNET SOC'Y (Mar. 24, 2017), <https://www.internetsociety.org/wp-content/uploads/2017/03/ContentBlockingOverview.pdf> [<https://perma.cc/Z78P-YQ8K>] (explaining that there are other motivations for blocking content, such as preventing or responding to network security threats or managing network usage) [hereinafter INTERNET SOC'Y].

side effects, pitfalls, trade-offs, and associated costs.”<sup>125</sup> According to this overview (and other similar reports<sup>126</sup>), approximately five main content-blocking methods exist that target the elements of a typical end-user sequence of searching, retrieving, and viewing content with a web browser or similar tool. Note that while these methods may be applied at different points of access—national,<sup>127</sup> individual telecommunication carriers,<sup>128</sup> local network,<sup>129</sup> or endpoint<sup>130</sup>—blockings based on public policy, such as blocking of pirate websites, occur on the national or carrier level.

### 1. IP Blocking

The simplest website blocking method is based on IP addresses, and its essential goal is to block all traffic to the IP address associated with the designated website. This means that any attempt to connect to a server with that IP address will be interrupted.

In terms of accuracy, this blocking method ranks poorly. To the extent that legitimate content shares the same IP address with the illegitimate content, legitimate content will be inevitably blocked too.<sup>131</sup> In legal terms, this equates to over-enforcement of copyrights, which tilts the balance between free speech and copyright protection to the benefit of the latter. Moreover, the fact that only the hosting provider knows exactly how many websites share the same IP address suggests that IP-based blocking could be quite arbitrary.<sup>132</sup>

Furthermore, the efficiency of this blocking method is also doubtful. IP-based blocking is implemented by devices located between the end-user and the pirate website.<sup>133</sup> Hence, users who are not “behind” the blocking device, because they use the services of an internet provider that has not inserted a

---

125. *Id.* at 5.

126. See “*Site Blocking*” to Reduce Online Copyright Infringement: A Review of Sections 17 and 18 of the Digital Economy Act, OFCOM 26 (May 27, 2010), [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/78095/Ofcom\\_Site-Blocking\\_-\\_report\\_with\\_redactions\\_vs2.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/78095/Ofcom_Site-Blocking_-_report_with_redactions_vs2.pdf) [<https://perma.cc/7PHT-4G7M>] [hereinafter OFCOM, “*Site Blocking*” to Reduce Online Copyright Infringement].

127. When all traffic entering or leaving a country may be subject to content blocking.

128. When mobile carriers and traditional ISPs install content blocking tools.

129. When local networks, such as home or school networks, install blocking tools, usually for the purpose of network management or security policy.

130. When software is installed directly on end-user computers, usually for security reasons but also for network management or parental control reasons.

131. INTERNET SOC’Y, *supra* note 124, at 12 (providing a diagram showing how IP blockings could easily result in over-enforcement).

132. Lukas Feiler, *Website Blocking Injunctions under EU and U.S. Copyright Law—Slow Death of the Global Internet or Emergence of the Rule of National Copyright Law?* 9–10 (ITIF Working Paper, No. 13, 2012).

133. INTERNET SOC’Y, *supra* note 124, at 13.

blocking device, as well as users who use technology that conceals the true destination of their traffic (such as VPN), can bypass the blocking.<sup>134</sup> Additionally, the effectiveness of IP-based blocking diminishes when website owners use content delivery networks (CDNs) that constantly change the infringing content's IP addresses.<sup>135</sup>

### 2. *Blocking Based on Deep Packet Inspection*

Another website-blocking method is based on Deep Packet Inspection (DPI). Unlike IP-based blocking, with deep packet inspect, sophisticated software filters all content according to specific blocking rules.<sup>136</sup>

This method also raises a number of issues. Privacy is particularly threatened because all users' actions that are not encrypted are being inspected.<sup>137</sup> Meanwhile, there are several questions regarding the effectiveness of this blocking method since it cannot inspect encrypted content, even though more than half of internet traffic is encrypted.<sup>138</sup> In terms of costs, this blocking method is considered quite expensive to apply because it depends on the development of filtering software. Since its success rests on the software's ability to identify particular content (according to keywords, traffic characteristics, or filenames), it is more efficient for network management and security enforcement, but not for policy-based blocking, which is far more flexible.<sup>139</sup>

### 3. *URL-Based Blocking*

A third website-blocking method is based on the URL. This blocking device may be located on the end-user's computer or in a network between

---

134. *Id.*

135. *Id.*

136. *Id.* at 14.

137. *See infra* Part IV.

138. Cam Cullen, *The Global Internet Phenomena Report*, SANDVINE (Oct. 2018), <https://www.sandvine.com/hubfs/downloads/phenomena/2018-phenomena-report.pdf> [<https://perma.cc/RW8W-D65M>].

139. For instance, some uses of copyrighted material constitute fair use for various policy reasons, such as promoting criticism, enabling research, and supporting education. Yet, fair use is a flexible standard, whose application depends on the specific circumstances of the particular use: (1) the purpose and character of the use; (2) the nature of the copyrighted work; (3) the amount taken; (4) the effect of the use on the market for the copyrighted work. While designing a software that would meet this standard now seems more possible than ever, given the recent developments in big data and machine learning, it is definitely much more complicated than designing a software that meets more rigid black line security rules. *See* Niva Elkin-Koren, *Fair Use by Design*, 64 UCLA L. REV. 1082 (2017). For a skeptical view on this issue, see Dan Burk, *Algorithmic Fair Use*, 86 U. CHI. L. REV. 283, (2019).

the end-user and the rest of the internet.<sup>140</sup> URL is the global address of documents and resources on the World Wide Web; therefore, URL-based blocking is not suitable for blocking non-web applications (such as Voice over Internet Protocol).<sup>141</sup> URL-based blocking can be implemented by proxies, as well as by firewalls and routers that block the connection to the web server requested by the end-user (as indicated by the Hypertext Transfer Protocol request), or otherwise direct web traffic to a different webpage. The blocking device intercepts the flow of web traffic and filters URLs that appear in the blocking list.

This, too, raises concerns. Based on the infrastructure, this method depends on the blocking party's questionable ability to control traffic between the end-user and the internet. Designing such a filter can be quite costly.<sup>142</sup> In terms of accuracy, URL-based blocking may suffer from false positives and false negatives alike. On one hand, it may block legitimate content that resides on a blocked web page (take the Wikipedia model, for instance, where blocking a single web page could block access to additional hyperlinks that are embedded in that page and that may link to legitimate content). On the other hand, content providers can quite easily evade the blocking by changing their file's name or using a different server.<sup>143</sup> Additionally, URL-based blocking monitors web traffic while intervening with users' privacy.<sup>144</sup>

#### 4. *Platform Filtering*

The fourth blocking method depends on platform filtering implemented by major online services such as search engines, social media platforms, or mobile application stores (such as Apple's App Store or the Google Play store). This blocking method depends on cooperation on the part of platforms that filter out objectionable content, either due to local regulation and government requirements or to the platforms' own terms of service (regarding pornography, for instance).

This method results in inconsistency and ineffectiveness. With regards to inconsistency, users of different search engines, as well as users accessing the internet from different countries (for instance, using the U.S. as opposed to the German version of Google) may be able to retrieve different content.<sup>145</sup> Furthermore, since this blocking method only filters out pointers to illegitimate content, but not actual content—which remains available online

---

140. INTERNET SOC'Y, *supra* note 124, at 15.

141. *Id.*

142. *Id.* at 16.

143. *Id.*

144. *Id.* at 17.

145. *Id.* at 18.

and accessible through other means of retrieving content—it is considered extremely ineffective.<sup>146</sup>

However, it is still very popular both at the national level, especially in online copyright enforcement,<sup>147</sup> and on a private individual level, because it enforces the right to be forgotten.<sup>148</sup>

### 5. DNS-Based Blocking

A fifth website blocking method is based on Domain Name Systems (DNS). DNS is an easy, user-friendly system for looking up and retrieving content. Users enter their queries in words, separated by dots (for instance, [www.haifa.ac.il](http://www.haifa.ac.il)), or otherwise enter a specific URL (for instance, <https://www.haifa.ac.il/index.php/he/>), and the domain name lookup result directs them to the matching IP address (for instance, 132.74.189.243).

The major advantage of DNS-based content blocking over other blocking methods is that it does not rely on designing a complicated filter which intercepts all web traffic—hence it is both privacy-friendly and less expensive to implement.<sup>149</sup> With DNS-based blocking, the DNS resolver validates specific search names against a list of illegitimate names, and whenever there is a match the DNS resolver returns incorrect information, or else declares that the name does not exist, so users' access to content using certain domain names is disabled. To be effective, DNS-based blocking depends on the blocking party having complete control over the end-user's network connection, since both users and content providers can easily avoid this blocking technique by using different internet connections or using an alternative set of DNS servers.<sup>150</sup> Like IP-based blocking, DNS-based blocking may also result in blocking legitimate content which resides in the same server

---

146. *Id.*

147. See *Government Requests to Remove Content*, GOOGLE TRANSPARENCY REP., [https://transparencyreport.google.com/government-removals/overview?removal\\_requests=group\\_by:totals;period:&lu=removal\\_requests](https://transparencyreport.google.com/government-removals/overview?removal_requests=group_by:totals;period:&lu=removal_requests) [<https://perma.cc/95L3-DBFZ>] (last visited Jan. 4, 2020).

148. Case C-131/12, *Google Spain SL, Google Inc. v. Agencia Española de Protección de Datos (AEPD)*, 2014 EUR-Lex CELEX LEXIS 62012CJ0131 (May 13, 2014) (acknowledging users' right to request search engines to remove links to personal data unless a strong public interest suggests otherwise). Google has received more than 3.4 million requests to remove URLs. See GOOGLE TRANSPARENCY REP., <https://transparencyreport.google.com/eu-privacy/overview> [<https://perma.cc/RLB3-KUGX>] (last visited Nov. 16, 2019).

149. INTERNET SOC'Y, *supra* note 124, at 19.

150. See Chris Hoffman, *5 Ways to Bypass Internet Censorship and Filtering*, HOW-TO GEEK (Aug. 2, 2016), <https://www.howtogeek.com/167418/5-ways-to-bypass-internet-censorship-and-filtering/> [<https://perma.cc/7YWU-BX52>].

using the same domain name (for instance, management.Haifa.ac.il).<sup>151</sup> Indeed, DNS blocking usually targets the uppermost level of the infringing domain.<sup>152</sup> However, compared to IP-based blocking it is slightly more accurate because it is easier regularly to update lists of domain names. However, it is less effective than IP-based blocking because bypassing DNS-based blocking is even easier than bypassing IP-based blocking.<sup>153</sup>

To summarize, injunctions directing third parties to block users' access to pirate websites could be achieved through various content-blocking means, which diverge in terms of accuracy, effectiveness, and cost. Common to all blocking methods are their robust collateral effects,<sup>154</sup> which impact human rights and shape the balance between clashing rights and interests.<sup>155</sup> Specifically, the particular technical details which underline a specific blocking ultimately define the scope and breadth of the blocking remedy itself: how substantially it will burden the financial interests of the ISP; to what extent it could harm legitimate content; and whether it is expected to work efficiently in preventing piracy.

The example of blocking injunctions is imperative for expressing how central the details of digital remedies' implementations could turn out to be. As explained in the following Part, the significant meaning of the remedy's technical implementation details raises a serious compatibility question, which challenges the ability of the court to fulfill its longstanding duty to exercise its adjudication power in accordance with the rule of law, to competently prescribe remedies that are expected to redress the violation of rights, and to assure these remedies are enforced properly. Since the implementation details of digital remedies are defined and executed outside the courthouse, on private grounds, and considering their ample meaning, the fact that they could surpass the court's dominion calls for special attention.

---

151. INTERNET SOC'Y, *supra* note 124, at 19.

152. OFCOM, "Site Blocking" to Reduce Online Copyright Infringement, *supra* note 126, at 34. In the domain hierarchy, the top-level domains are represented by extensions such as ".com," ".eu," ".edu," etc.

153. *Id.*

154. *Id.*; Geiger & Izyumenko, *supra* note 85, at 11–16.

155. *See infra* Part IV.



#### IV. DIGITAL REMEDIES, JUDICIAL DECISION MAKING AND THE RULE OF LAW

Most, if not all, remedy law scholars would agree that “the available remedy influences the content of the right that courts articulate in a given case.”<sup>156</sup> This close remedy-right interdependence suggests that prescribing remedies is a fundamental stage in judicial decision making.<sup>157</sup> Specifically, in relation to equitable remedies, courts enjoy relatively broad discretion to fashion remedies that are appropriate to the justice of the particular case.<sup>158</sup>

But this discretion is limited.<sup>159</sup> Like any other exercise of judicial decision making, when judges apply their remedial power, they must preserve the rule of law and exercise their discretion competently, fairly, and transparently.<sup>160</sup> Generally, the rule of law has long been interpreted as comprising two basic ideas: first, that individuals should be governed by law rather than by the arbitrary will of others;<sup>161</sup> and second, that no person is above the law.<sup>162</sup> The law must be clear, so people can develop reliable expectations and make autonomous choices accordingly. Judges are hence “expected to give a reasoned explanation of the process by which they reach their conclusions.”<sup>163</sup> In application to the prescription of judicial remedies, it is fair to posit that courts are expected to delineate a clear and precise redress, which expresses a delicate balance between the various rights and interests of those who might

---

156. Leong, *Making Rights*, *supra* note 10, at 416; Kermit Roosevelt III, *Aspiration and Under Enforcement*, 119 HARV. L. REV. F. 193, 194 (2006) (arguing that remedial considerations exert an important influence over the shape of the standards courts adopt to implement constitutional rights).

157. See, e.g., Mitchell N. Berman, *Constitutional Decision Rules*, 90 VA. L. REV. 1, 43–50 (2004); Laurin, *Rights Translation*, *supra* note 10, at 1007–08.

158. Doug Rendleman, *The Triumph of Equity Revisited: The Stages of Equitable Discretion*, 15 NEV. L.J. 1397, 1402–03 (2015) (providing two examples of equitable discretion in equity areas: one in family law and one in property law).

159. See *Heine v. Levee Comm’rs*, 86 U.S. 655, 658 (1873) (rejecting the notion that a court of equity may “depart from all precedent and assume an unregulated power of administering abstract justice at the expense of well-settled principles”); PHILIP HAMBURGER, *LAW AND THE JUDICIAL DUTY* 142–43 (2008) (describing “equitable discretion” in the eighteenth century as “a discernment of circumstances” sometimes “beyond reconsideration on error, but this was not to say it was necessarily beyond rules of either equity or law”).

160. Guri Ademi, Comment, *Legal Intimations: Michael Oakeshott and the Rule of Law*, 1993 WIS. L. REV. 839, 845 (1993).

161. ALBERT V. DICEY, *INTRODUCTION TO THE STUDY OF THE LAW OF THE CONSTITUTION* 189–90 (10th ed. 1959).

162. *Id.* at 193.

163. Maria L. Marcus, *Judicial Overload: The Reasons and the Remedies*, 28 BUFF. L. REV. 111, 114 (1979).

be affected from the remedy granted. In short, we expect judges to dominate the scope and reach of the remedies they grant.

Some remedies, however, make it difficult for judges to exercise complete control over the remedies they grant and anticipate their ultimate impact. As explained earlier, “remedies compelling either action or inaction,” for instance, often present a problem of “specifying, measuring, and ensuring compliance.”<sup>164</sup> In particular, equitable remedies “are costly to administer because they do more than transfer a lump sum from defendant to plaintiff, the standard ‘legal’ remedy.”<sup>165</sup> Digital remedies, as a sub-category of equitable reliefs, take these concerns to the next level. Not only are there substantial underlying digital details determined outside the courthouse, these details are very hard to appreciate and control.

First, digital remedies have a robust impact on the rights and interests of numerous stakeholders. Second, the implementation details of digital remedies are dynamic in their implications, costs, and capabilities of adjusting to the changing digital landscape. And third, the implementation details are embedded in privately-developed, non-transparent codes. The following discussion describes these unique attributes of the means used to implement digital remedies and explains how they challenge the ability of courts to engage in responsible decision making.

#### A. ROBUST IMPACT ON NUMEROUS STAKEHOLDERS

Digital reliefs are directed to cyberspace and, therefore, they are inherently widespread in their impact.<sup>166</sup> Whether sought to interfere with the operation of digital devices, such as a streaming device (e.g., TickBox<sup>167</sup>) or a smartphone (e.g., iPhone<sup>168</sup>), or otherwise to manage online content (e.g., block online copyright infringement<sup>169</sup>), digital reliefs have a robust effect on numerous actors, far exceeding their direct impact on the parties to the legal dispute. Even when it appears that courts narrowly tailor digital reliefs—for instance, when courts order to disable access to specific websites or to decrypt a particular iPhone—digital reliefs unfold in a wide-reaching fashion.

In particular, the implementation of digital remedies could have a substantial impact over the fundamental rights of numerous internet users who are not direct parties to the legal dispute and, thus, whose interests are not

---

164. Bray, *supra* note 21, at 563.

165. *Avitia v. Metro. Club of Chi., Inc.*, 49 F.3d 1219, 1231 (7th Cir. 1995).

166. Mulligan & Bamberger, *Saving Governance by Design*, *supra* note 17, at 739.

167. *See supra* note 69 and accompanying text.

168. *See supra* note 72 and accompanying text.

169. *See supra* note 116 and accompanying text.

necessarily adequately represented. For instance, blocking users' access to legitimate online content could curtail their First Amendment rights to freely consume information in the marketplace of ideas.<sup>170</sup> When the operators of TickBox issued a software update to delete all infringing applications from their devices to comply with the court's digital injunction, they essentially diminished their users' ability to consume non-infringing content through these apps or otherwise make non-infringing uses of the content, while also limiting their users' freedom of expression. The same is true in relation to the implementation of website-blocking injunctions, which obviously limit users' right to receive information.<sup>171</sup>

Besides the right to freedom of expression, the implementation of digital remedies may also affect users' privacy. The order which compelled Apple to develop a technological "backdoor" to allow law enforcement agents to break into the locked iPhone of the deceased shooter in San Bernardino is a prominent example.<sup>172</sup> If Apple had complied with the order and written a code to unlock its strong security system, it would have put the data of millions of individuals, inside and outside the United States, at serious risk of unwarranted surveillance, potentially making them victims of crime.<sup>173</sup> Moreover, had the FBI won this legal battle, other technology companies might have followed suit, redesigning their security features to accommodate what they might have interpreted as a judge-made requirement: to design technological backdoors to their digital devices.<sup>174</sup> The order could have had a worrying impact on both national and international security, particularly

whether used by a black hat hacker who might infiltrate Apple systems, a future FBI investigation emboldened by [the court's] order to apply the precedent in other less compelling settings, or a dictatorship looking for new ways to oppress people that might cite

---

170. See Jerome A. Barron, *Access to the Press—A New First Amendment Right*, 80 HARV. L. REV. 1641, 1666–78 (1967); Jamie Kennedy, Comment, *The Right to Receive Information: The Current State of the Doctrine and the Best Application for the Future*, 35 SETON HALL L. REV. 789, 789–90 (2005); Susan Nevelow Mart, *The Right to Receive Information*, 95 LAW LIBR. J. 175, 175 (2003).

171. See Geiger & Izyumenko, *supra* note 85, at 49.

172. See Mulligan & Bamberger, *Saving Governance by Design*, *supra* note 17, at 722–26.

173. See Apple Inc.'s Motion to Vacate Order Compelling Apple Inc. to Assist Agents in Search, and Opposition to Government's Motion to Compel Assistance at 4, *In re Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300*, Cal. License Plate 35KGD203, (No. CM 16-10 (SP)) (C.D. Cal. Feb. 25, 2016).

174. See Mulligan & Bamberger, *Saving Governance by Design*, *supra* note 17, at 726.

the company's compliance with this FBI demand as a reason to comply with those of its own intelligence agencies.<sup>175</sup>

Indeed, while the FBI framed its demand as addressing a single phone, in practice, the implementation of the order would necessarily place the security of millions of other devices and the people who use them at risk.<sup>176</sup>

Similarly, the implementation of the TickBox injunction discussed earlier could also affect the privacy of numerous end users. As the court noted, deleting applications that are independently downloaded by users because they induce copyright infringement may require TickBox operators to hack into their users' devices.<sup>177</sup>

An additional circle of stakeholders which might be significantly affected by the grant of some digital reliefs are those acting "in concert or active participation" with the defendants, who might be compelled to abide by the court injunction even though they are not party to the action brought by plaintiffs.<sup>178</sup> The *Sci-Hub* injunction, for instance, required that

any person or entity in privity with Sci-Hub and with notice of the injunction, including any Internet search engines, web hosting and Internet service providers, domain name registrars, and domain name registries, cease facilitating access to any or all domain names and websites through which Sci-Hub engages in unlawful access to, use, reproduction, and distribution of ACS's trademarks or copyrighted works.<sup>179</sup>

Holding such a broad spectrum of actors accountable for pursuing the open-ended outcome of restricting access to particular websites may exceed the boundaries of Rule 65 of the Federal Rules of Civil Procedure.<sup>180</sup> However, this debate is beyond the scope of this paper.

Yet, even assuming that such an injunction is procedurally permitted, requiring distinct intermediaries to actively cooperate in its implementation may affect both their free speech and business interests. First, it interferes with distinct online intermediaries in setting and employing "their own content

---

175. Shahid Buttar, *Apple, Americans, and Security vs. FBI*, ELECTRONIC FRONTIER FOUND. (Feb. 20, 2016), <https://www EFF.ORG/deep links/2016/02/apple-americans-and-security-vs-fbi> [<https://perma.cc/94B8-HTSW>].

176. *Id.*

177. See CCIA Amicus Brief, *supra* note 113.

178. See Husovec, *supra* note 93, at 12.

179. Magistrate Judge's Proposed Findings, *supra* note 104, at 12.

180. See FED. R. CIV. P. 65(d)(2)(C).

standards.”<sup>181</sup> Second, it inflicts high compliance costs on nonparties<sup>182</sup> without affording them the opportunity to object, raising serious due process concerns.<sup>183</sup> To satisfy procedural due process, sufficient evidence showing that remote intermediaries have aided and abetted the defendants in circumventing the injunction issued, or are likely to do so, should be presented in a proceeding where those entities are given an opportunity to be heard.<sup>184</sup> Nonetheless, at least in the *Sci-Hub* case, none of these entities had their day in court. Hence their interests remained largely unrepresented.<sup>185</sup>

Moreover, as demonstrated in Part III above, blocking injunctions, for instance, could affect providers of legitimate content that might be unintentionally blocked due to over-enforcement.<sup>186</sup> This depends on the accuracy of the blocking method applied: the less accurate the method is, the more likely it is to block non-infringing content, as well. Such restrictions of legitimate speech would potentially harm the rights and interests of content providers. Regarding the *TickBox* injunction, for example, software deleted in response to the court’s injunction may include applications that link to legitimate content, such as CBS, WatchESPN, The Weather Channel, or Cartoon Network.<sup>187</sup> This means that in addition to impairing the rights of users to access non-infringing content, the implementation of the injunction could also violate the rights and interests of various speakers.

These examples suggest that the overall impact of digital remedies could far exceed the particular rights and interests of the direct parties to the legal dispute. Fair and appropriate prescription of digital remedies requires a thorough consideration of the fundamental rights held by numerous stakeholders, which must be balanced against other important interests such as public safety and security, access to information, or various business

---

181. Rebecca Tushnet, *Power Without Responsibility: Intermediaries and the First Amendment*, 76 GEO. WASH. L. REV. 986, 1002 (2008).

182. See, e.g., Case C-70/10, *Scarlet Extended SA v. Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)*, 2011 E.C.R. I-12006; Case C-360/10, *Belgische Vereniging van Auteurs, Componisten en Uitgevers CVBA (SABAM) v. Netlog NV*, 2012 EUR-Lex CELEX LEXIS 62010CJ0360 (refusing to grant a website-blocking order, reasoning that its high implementation costs as well as its complexity would overburden the service provider).

183. See, e.g., Feiler, *supra* note 132.

184. See *Mullane v. Cent. Hanover Bank & Tr. Co.*, 339 U.S. 306, 313 (1950).

185. See Magistrate Judge’s Proposed Findings, *supra* note 104, at 4.

186. See *supra* notes 131–155 and accompanying text.

187. Response in Opposition to Motion for Preliminary Injunction at 3, *Universal City Studios v. TickBox TV*, No. 17-7496 (C.D. Cal. Dec. 28, 2017).

interests,<sup>188</sup> especially given that these rights and interests are not necessarily voiced during the regular legal process.<sup>189</sup> The problem, however, is that it is not enough to make this consideration in advance because translating legal balances into digital processes may result in alterations of meaning.<sup>190</sup> Indeed, digital remedies are open to different implementations, and these are subsequently interpreted and embedded in proprietary codes.<sup>191</sup>

Alteration of meaning may occur twice: First, when the private operator who executes the order decides which digital measure to apply in order to achieve the desired outcome. Second, when program developers create the code which applies this measure. Thus, even if digital remedies could reflect a broad and inclusive deliberation of diverse rights and interests, their practical, out-of-court implementation could effectively reshape settled legal balances. But if courts cannot anticipate how digital reliefs unfold, they cannot ensure that they are actually fit to redress specific violations of rights, and this further challenges the rule of law.

Consider, for instance, the implementation of content-blocking injunctions. Normally, under settled copyright doctrine, content is allowed unless it is found to be infringing,<sup>192</sup> creating a delicate balance between the property rights of current creators and the freedom of expression of future ones.<sup>193</sup> Nevertheless, as shown in Part III, content-blocking techniques may over-enforce copyrights and block legitimate content, at the expense of the rights of creators of legitimate content and the public at large. Similarly, if the FBI had not withdrawn its motion to compel Apple to develop a technological

---

188. Geiger & Izyumenko, *supra* note 85, at 77–82 (discussing the economic impact of copyright website blockings on ISPs, which are not only complex but also quite expensive to implement).

189. *Ex parte* demands, such as the FBI's demand in the Apple v. FBI dispute, normally completely deprive the court of defendant's perspective altogether. Bamberger & Mulligan, *Saving Governance by Design*, *supra* note 17, at 723. In the Apple v. FBI dispute, however, Apple and numerous organizations did receive an opportunity to raise their concerns because the FBI filed a motion to compel Apple to comply with the assistance order. *Id.*

190. See Austl. Admin. Rev. Council, *Automated Assistance in Administrative Decision Making*, Issues Paper No. 35, 18–19 (2003), <https://www.ag.gov.au/LegalSystem/AdministrativeLaw/Documents/practice-guides-and-other-publications/automated-assistance.pdf> [<https://perma.cc/GQK4-JPTL>]; James Grimmelman, *Regulation by Software*, 114 YALE L.J. 1719, 1727–28 (2005).

191. See *infra* Section IV.C.

192. Jennifer M. Urban & Laura Quilter, *Efficient Process or “Chilling effects”? Takedown Notices Under Section 512 of the Digital Millennium Copyright Act*, 22 SANTA CLARA COMPUTER & HIGH TECH. L.J. 683 (2006).

193. As stated by James Madison, the framer of the Constitution's Copyright Clause, “the public good fully coincides . . . with the claims of individuals.” See THE FEDERALIST NO. 43 (James Madison).

backdoor to its iPhone security system, and other technology companies had followed suit, adjusting their devices' security features so as to make them breakable, the balancing of rights and interests initially set by the court could have been skewed. Even if, originally, surveillance was to be allowed only in this particular case to protect public security, by now accessing personal data without the owner's consent could have become generally easier, while imposing a serious threat to users' privacy.<sup>194</sup>

To sum up, the ways in which digital remedies unfold have a widespread affect over innumerable right holders. Even if judges could potentially afford adequate consideration to all the rights and interests on the table, the problem remains unresolved: the out-of-court, digital implementation of digital remedies could practically redefine judicial balances and have dramatic impacts on settled law.

#### B. DYNAMIC AND ONGOING IMPACT

Another major problem with digital remedies which further complicates courts' capacity to control and anticipate how they evolve relates to the dynamic nature which surrounds their implementation. Unlike judicial reliefs that provide a "one-shot" solution to a legal dispute, any application of structured technological solutions to resolve legal disputes arising in the digital ecosystem must be able to adjust to a rapidly changing technological environment.<sup>195</sup> For instance, blocking access to pirate websites could be easily circumvented if users and content providers conceal their online conduct by using VPNs, proxy services, and the like.<sup>196</sup> History has taught us that the circumvention of digital locks is only a matter of time and persistence.<sup>197</sup> This suggests that the efficacy of content blocking is, at most, temporary. But if their effectiveness decreases, what is left to compensate for the censorship of legitimate content? To address this issue, digital remedies must allow for timely adaptations.

Moreover, digital remedies are often directed to resolving an ongoing problem, which further blurs their anticipated limits. The *Sci-Hub* injunction, for instance, was amended soon after it was initially signed by the court according to a magistrate judge's proposed findings in order to expand ACS's

---

194. See Mulligan & Bamberger, *Saving Governance by Design*, *supra* note 17, at 726.

195. See *id.* at 739.

196. *Supra* Section III.C.

197. The circumvention of Digital Rights Management systems (DRMs) which are supposed to restrict users' use of and access to copyrighted protected works is one example. See Brandon Widder, *DRM Getting You Down? Here's How to Strip Your Music and Movies of Restrictions*, DIGITAL TRENDS (Feb. 22, 2015), <https://www.digitaltrends.com/home-theater/how-to-remove-drm-from-music-and-movie-files/> [<https://perma.cc/TS4G-SWV5>].

ability to act against newly registered domain names as well as the domain names already registered when the initial injunction was issued.<sup>198</sup> Without this amendment, ACS would have been “forced to engage in a game of whac-a-mole whereby new sci-hub domain names emerge” rapidly.<sup>199</sup>

Furthermore, unexpected dynamics in the technological environment which surround the implementation of digital remedies could also affect innovation in different and unpredicted ways. The German “free Wi-Fi” experience is an excellent example. In 2010, the German Supreme Court held that a private operator of an open Wi-Fi network should help rights-holders enforce their rights by sufficiently password-locking the network’s connectivity in order to prevent possible misuse.<sup>200</sup> Consequently, password-protected Wi-Fi connections became the *de facto* standard in Germany.<sup>201</sup> When subsequent technological solutions became dependent on open Wi-Fi, Germany suffered a serious innovative setback.<sup>202</sup> Hence, the court’s failure to anticipate the full impact of the digital remedy it had granted eventually slowed down progress and innovation.

But it is not only the technological environment which surrounds the implementation of digital remedies that is dynamic—it is also the means of implementation themselves, and their potential costs. Consider, for instance, blocking injunctions. European courts have acknowledged that the cost of implementing blocking measures might be quite substantial.<sup>203</sup> As shown previously, these costs vary with the specific blocking technique implemented.<sup>204</sup> However, since the manner of implementation is determined on private grounds, or outside the courthouse, courts cannot really anticipate what would be the total compliance costs, presenting further challenges their

---

198. Ernesto, *Publisher Gets Carte Blanche to Seize New Sci-Hub Domains*, TORRENTFREAK (Apr. 10, 2018), <https://torrentfreak.com/publisher-gets-carte-blanche-to-seize-new-sci-hub-domains-180410/> [<https://perma.cc/J6SF-BVUS>].

199. *Id.*

200. Husovec, *supra* note 93, at 4–5.

201. *Id.*

202. See Loveday Wright, *Germany’s Wi-Fi Problem*, DW (Nov. 13, 2014), <https://www.dw.com/en/germanys-wi-fi-problem/a-18060000> [<https://perma.cc/6UMZ-NLKS>].

203. See, e.g., Case C-314/12, *UPC Telekabel Wien GmbH v. Constantin Film Verleih GmbH* 2014 EUR-Lex CELEX LEXIS 62012CA0314 (May 19, 2014) Bus LR 541; Case C-70/10, *Scarlet Extended SA v. Société Belge des Auteurs, Compositeurs et Éditeurs SCRL (SABAM)*, 2011 E.C.R. I-12006, ¶ 50; Case C-360/10, *Belgische Vereniging van Auteurs, Componisten en Uitgevers CVBA (SABAM) v. Netlog NV*, 2012 EUR-Lex CELEX LEXIS 62010CJ0360 (Feb. 16, 2012) (holding that the application of content filtering technology is too expensive and therefore ISPs cannot be obliged to include filtering in their services).

204. See *supra* Section III.C.



ability to exercise their remedial power in a fair and competent manner. How could they grant a relief of which its economic burden is unknown?

Additionally, to the extent that digital remedies are implemented through evolving measures, such as machine learning algorithms, the ability to anticipate their final reach becomes even more complicated. Thanks to recent developments in big data, some digital remedies may rely on advanced capabilities of machine learning to pursue their objectives more efficiently. For instance, different content-blocking methods, especially platform, URL, or DPI-based blocking, depend on filtering technologies that monitor all content that is available in the network.<sup>205</sup> These content filters could be designed to identify trends, relationships, and hidden patterns in disparate sources of content, which are then used to shape users' experience.<sup>206</sup> Yet, while shaping performance based on experience could be particularly valuable for implementing flexible policy-based blocking of copyright-infringing content, it is very hard to follow and predict its potential impact.

### C. NON-TRANSPARENT IMPLEMENTATION ON PRIVATE GROUNDS

Real-world compliance with judicial remedies is generally clear-cut and its underlying objectives are self-evident. This is because physical actions (or inactions) are generally easy to check: selling goods, erecting a fence, or avoiding trespassing. The implementation of digital remedies, on the other hand, is often embedded in proprietary black-box codes, which could be very difficult to evaluate.<sup>207</sup> Consider again, for example, the *TickBox* injunction, which essentially compelled *TickBox* to issue a software update that would delete all software that enabled users to access copyright-infringing content.<sup>208</sup> The practical breadth of this proprietary software update is unknown and largely unknowable.<sup>209</sup> One theory posits that *TickBox* released a software

---

205. *Id.*

206. Perel & Elkin-Koren, *Black Box Tinkering*, *supra* note 2, at 189.

207. Rob Kitchin, *Thinking Critically About and Researching Algorithms* 7 (The Programmable City, Working Paper No. 5, 2014), <http://ssrn.com/abstract=2515786> [<https://perma.cc/UYB9-KHCK>]; Perel & Elkin-Koren, *Accountability in Algorithmic Enforcement*, *supra* note 1, at 476; Citron, *Technological Due Process*, *supra* note 1, at 1261–62; Tal Z. Zarsky, *Governmental Data Mining and Its Alternatives*, 116 PA. ST. L. REV. 285, 293 (2011); Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1552 (2013).

208. *See supra* Section III.A.

209. *See, e.g.*, *Viacom Int'l Inc. v. YouTube Inc.*, 253 F.R.D. 256, 259–60 (S.D.N.Y. 2008) (refusing to force YouTube to provide Viacom with the computer source code which controls both YouTube.com's search function and Google's internet search tool "Google.com," explaining that "[t]he search code is the product of over a thousand person-years of work" and "[t]here is no dispute that its secrecy is of enormous commercial value." Earlier cases invoked trade secrets in Google's ranking algorithm); *see also* *Kinderstart.com LLC v. Google*,

update that removed copyright-infringing addons from previously shipped devices.<sup>210</sup> Since it must block access to “any ‘build,’ ‘theme,’ ‘app,’ ‘addon[.]’ or other software program that TickBox knows or has reason to know links directly or indirectly to third-party cyberlockers or streaming sites that transmit unauthorized performances of copyrighted motion pictures or television shows,”<sup>211</sup> it might also block access to additional, non-infringing, content. As rigorously contended by TickBox, many software programs designated by the plaintiffs in their complaint had substantial non-infringing uses, allowing users to access legitimate content.<sup>212</sup> Deleting these software programs would thus inevitably result in restricting even lawful content.<sup>213</sup> Hardly apparent, however, is precisely *which* pieces of content would be affected.

The same applies to the technological backdoor feature Apple was requested to design. Again, if Apple had designed a code enabling the FBI to access the terrorist’s locked iPhone, it would have probably been impossible to work out how it functioned.<sup>214</sup> To begin with, such a code would have probably been protected under trade secret law.<sup>215</sup> In fact, when the FBI dropped its case against Apple after a private tech firm managed to break into

---

Inc., No. C 06-2057 JF (RS), 2006 WL 3246596, at \*1–2 (N.D. Cal. July 13, 2006) (granting Google’s motion to dismiss and holding that Google’s use of secret methods to compile search results does not amount to anticompetitive conduct); MICHAEL J. MADISON, OPEN SECRETS IN THE LAW AND THEORY OF TRADE SECRECY 222, 241 (Rochelle C. Dreyfuss & Katherine J. Strandburg eds., 2011).

210. See *Tickbox: Customers: They’re About to Remotely Wipe Your Devices Without Your Consent*, TVADDONS (Feb. 14, 2018), <https://www.tvaddons.co/tickbox-remote-wipe/> [<https://perma.cc/6428-3GVG>].

211. TickBox 2, *supra* note 69, at 1.

212. TickBox Response in Opposition to Motion for Preliminary Injunction, Universal City Studios Prods. L.L.L.P. v. TickBox TV L.L.C., No. 2:17-cv-07496-MWF (ASX), at \*3 (C.D. Cal. Feb. 13, 2018) (“[T]he Box is simply a small computer which performs common and non-infringing functions of any smartphone, tablet, or desktop computer, and allows its users the ability to download a number of third-party applications that provide users access to authorized streaming content directly from content providers.”).

213. See Annemarie Bridy, *A New Front in the Set-Top Box Piracy Wars: Can SONY’S Safe Harbor Save TICKBOX TV?*, CTR. INTERNET & SOC’Y STAN. L. SCH. BLOG (Nov. 26, 2017), <http://cyberlaw.stanford.edu/blog/2017/11/new-front-set-top-box-piracy-wars-can-sony%E2%80%99s-safe-harbor-save-tickbox-tv> [<https://perma.cc/9LGW-8L6V>].

214. See Pasquale, *Restoring Transparency to Automated Authority*, *supra* note 2, at 237 (explaining how “[t]rade secrecy law also makes it all the more important to keep algorithms secret”).

215. Perel & Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, *supra* note 1, at 522–23.

the terrorist's phone, the FBI refused to reveal the identity of that third party or to disclose the method it had developed in order to access the iPhone.<sup>216</sup>

Indeed, with digital remedies, choosing between various implementation possibilities and applying them is done on private grounds, outside the courthouse, notwithstanding its important implications for the rule of law, human rights, and innovation. The defendants effectively operate as law makers, only without the safeguards which normally restrain traditional law making. To some extent, they "act as both a judge and an executioner, performing functions of great importance to the public which are normally reserved [for] authorized governmental bodies."<sup>217</sup> Nevertheless, as private actors, defendants are generally free to manage their own business in an undisturbed fashion.<sup>218</sup> While they arguably hold the necessary expertise to develop and implement the proper technology which will fit the digital remedy, they lack the responsibility to take into account broad and inclusive considerations that go beyond the defendants' obvious economic interest. Delegating the power to shape the ultimate scope and reach of digital remedies to private parties, hence, risks privileging their own economic interests.<sup>219</sup> For instance, leaving service providers with broad discretion to elect how to implement a blocking injunction may result in encouraging them to apply the cheapest blocking techniques, regardless of their efficacy or accuracy.

One possible way to address this issue of privatization is to grant technology-specific remedies. Particularly, courts could arguably point at specific digital measures that must be applied in order for the defendant to comply with the injunction. For example, Apple was required to accomplish three functions: (1) bypass or disable the self-destruct function on the phone; (2) allow the FBI to submit passcodes to the phone through electronic testing; and (3) ensure that software running on the phone would not introduce

---

216. Romain Dillet, *Justice Department Drops Lawsuit Against Apple as FBI Has Now Unlocked Farook's iPhone*, TECHCRUNCH (May 29, 2016), <https://techcrunch.com/2016/03/28/justice-department-drops-lawsuit-against-apple-over-iphone-unlocking-case/> [https://perma.cc/USF5-D3KA]. Note that a district judge had subsequently approved the FBI's refusal, ruling that it was not required to provide records relating to vendor identity under Exemptions 1, 3, and 7(E) of the Freedom of Information Act. See *Associated Press v. FBI*, 265 F. Supp. 3d 82 (D.C. Cir. 2017).

217. Perel & Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, *supra* note 1, at 485.

218. See John Eden, *Why Apple is Right to Resist the FBI*, TECHCRUNCH (Mar. 13, 2016), <https://techcrunch.com/2016/03/13/why-apple-is-right-to-resist-the-fbi/> [https://perma.cc/V4PY-STQJ] ("The FBI has no underlying right to compel Apple to create new software products.").

219. Bamberger & Mulligan, *Saving Governance by Design*, *supra* note 17, at 742.

additional delays between passcode attempts.<sup>220</sup> Judge Pym further advised Apple with regards to what it should actually do to reasonably pursue these functions, providing a recommended map of the specific technological actions that should be taken.<sup>221</sup> At the same time, however, since a privately developed code could be a form of protected speech,<sup>222</sup> the judge also allowed Apple to use alternate technological means as long the government concurred and these means achieved the functions designated in the order, as well as the functionality described in the technological map provided by the court.<sup>223</sup>

Technology-specific remedies arguably restrain the private executor's discretion in choosing the technological means to implement the injunction; still, they remain just as vague as open-ended injunctions. Indeed, as it is a private executor who eventually implements the injunction outside the courthouse, it remains difficult to check how far she applies the technological steps that the court has initially set forth. After all, these steps would be later embedded in proprietary technology, which is inherently non-transparent.

Moreover, the allegedly increased predictability of technology-specific injunctions may come at the price of hindering innovation and encumbering the accumulation of new technologies. This is because a particular technological map for achieving a specific legal outcome can only consider known technologies and their known pros and cons. However, technology changes rapidly. New technologies replace old ones, newly discovered attributes of old technologies may improve or negate their capabilities, and new combinations of technologies may expand their individualized effect.

---

220. Order Compelling Apple Inc. to Assist Agents in Search at 8, *In re Search of an Apple iPhone*, No. ED 15-0451M (C.D. Cal. Feb. 16, 2016).

221. *Id.* ¶ 3.

222. Kim Zetter & Brian Barrett, *Apple to FBI: You Can't Force us to Hack the San Bernardino iPhone*, WIRED (Feb. 25, 2016), <https://www.wired.com/2016/02/apple-brief-fbi-response-iphone/> [<https://perma.cc/J2L8-NT7G>] (referencing *Bernstein v. U.S. Dep't of Justice*, 176 F.3d 1132, 1141 (vacated) which held that "software, in its source code form . . . must be viewed as expressive for First Amendment purposes"); Apple Inc.'s Motion to Vacate Order Compelling Apple Inc. to Assist Agents in Search, and Opposition to Government's Motion to Compel Assistance at 32, *In re Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300*, Cal. License Plate 35KGD203, No. CM 16-10 (C.D. Cal. Feb. 25, 2016) ("The government asks this Court to command Apple to write software that will neutralize safety features that Apple has built into the iPhone in response to consumer privacy concerns. . . . This amounts to compelled speech and viewpoint discrimination in violation of the First Amendment."). *But see* Neil Richards, *Apple's "Code=Speech" Mistake*, MIT TECH. REV. (Mar. 1, 2016) (explaining that "[t]he Supreme Court has never accepted that code is protected like speech").

223. Order Compelling Apple Inc. to Assist Agents in Search, No. ED 15-0451M at ¶ 4 (Feb. 16, 2016).

Considering the ongoing nature of digital remedies discussed earlier, the need to adjust them from time to time is clear. However, confining executors' discretion to the technological standards applicable "back then," or when the court first issued its injunction, could hinder the development of better digital solutions going forward. To illustrate, curbing Apple's technological discretion might have forced it to follow the technological map provided by the court, which may not necessarily always be the most appropriate way to gain access to a locked iPhone. Presumably, Apple is in the best position to intervene with its own private technology in the least harmful way, and a technology-specific injunction could encumber that expertise, impeding the development of better, innovative solutions.

Apple, for instance, could have followed the technological map provided by the court in its order, but it could also have used alternative technological means to achieve this outcome while still being in compliance with the order. The government would likely have been satisfied either way, as long as it got access to the specific type of data it presumably sought: data indicating whether the shooter was acting independently or on behalf of a terror organization.<sup>224</sup> But successfully breaking into an iPhone is not the only thing that matters.

Equally as important are the *means* applied to achieve the outcome, especially when these may differ in their capabilities and costs, which in turn may directly impact human rights. GrayKey, for instance, is a small device which law enforcement agents use to access locked iPhones.<sup>225</sup> It takes GrayKey anywhere from an hour or two to a few days to guess an iPhone's password and give its operator full access to the phone's file system, including messages, photos, call logs, browsing history, and passwords.<sup>226</sup> However, an alternative device could be developed that would provide restricted access to data stored on locked iPhones which would be less intrusive to users' privacy. Such a device, for example, could restrict data portability, limiting law

---

224. Ann Kristin Glenser, *Decrypting Apple: Making Technology Companies the Referees of Law Enforcement on Privacy*, JOLT DIG. (June 7, 2017), <https://jolt.law.harvard.edu/digest/decrypting-apple-making-technology-companies-the-referees-of-law-enforcement-on-privacy> [<https://perma.cc/U5VK-YJFL>].

225. Zack Whittaker, *For \$15,000, GrayKey Promises to Crack iPhone Passcodes for Police*, ZDNET (Mar. 19, 2018), <https://www.zdnet.com/article/graykey-box-promises-to-unlock-iphones-for-police/> [<https://perma.cc/SD47-2BHX>]. However, Apple had very recently released a new feature, iOS 11.4.1, to address this security loophole. This feature requires users to unlock their device after an hour of inactivity to connect a USB accessory to make it more difficult for police to use GrayKey to unlock iPhones. See Isobel Asher Hamilton, *Apple is Reportedly Closing a Security Loophole that will Prevent Police from Accessing iPhones*, BUS. INSIDER (July 14, 2018), <http://www.businessinsider.com/apple-will-make-it-harder-for-police-to-access-locked-iphones-2018-6> [<https://perma.cc/A4CZ-YS4Z>].

226. Whittaker, *supra* note 225.

enforcement agents' ability to transfer the data they access to other devices. While such a hypothetical alternative might be more expensive, and even less effective for law enforcement purposes, it would better preserve privacy.

Overall, the essence of digital remedies is their profound technical details, and these are designed and executed outside the courthouse during implementation. Yet these details are far from being merely procedural; they effectively shape the balance between competing rights and interests held by numerous stakeholders. Given their opaque nature, and considering the dynamic environment in which digital remedies unfold, it becomes rather challenging to appreciate their scope and assure they constitute a fit redress. Therefore, the next and final Part explores how the toolkit of equitable managerial devices and constraints could assist courts in preserving their dominance over digital remedies.

## V. OVERSEEING DIGITAL REMEDIES

Overseeing how digital remedies unfold is vital to safeguard the rule of law, to protect human rights, and to ensure they are compatible with the changing digital reality. Although the grant of digital remedies is subject to traditional ex-ante judicial review, this is not enough to ensure courts exercise full and ongoing control of digital remedies. Accordingly, this last Part of the Article recommends several mechanisms that courts could exploit in order to extend their oversight and retain more control over the critical implementation stage of digital remedies. In essence, these tools purport to empower judges who resolve cyber-related disputes with a broader and a more accurate understanding of the meaning of their digital solutions.

This is where the system of equitable remedies comes into play. Recall that previously in Part II, digital reliefs were classified as specific, prospective, and equitable remedies, yet their equitable nature was especially emphasized given that they generally “compel action (or inaction), especially when that action may be continuing or iterative and not easily measured.”<sup>227</sup> Stressing the equitable nature of digital remedies is constructive because the system of equitable remedies includes, in addition to the remedy itself, equitable managerial devices that allow courts to manage the parties and ensure compliance, as well as special equitable restraints.<sup>228</sup>

---

227. Bray, *supra* note 21, at 533.

228. *See id.* at 534.

### A. MANAGERIAL DEVICES

Managerial devices generally purport to “enhance the court’s ability to manage the parties” and, thus, ascertain compliance.<sup>229</sup> In application to digital remedies, these devices could further enhance the court’s overseeing capabilities, allowing them to control the breadth and scope of the reliefs as they evolve. In particular, these devices could mitigate the problem of anticipating what would be the overall impact of particular digital remedies in advance.

#### 1. *Ex-Post Revision*

The dynamics which surround the implementation of digital remedies, and the rapidly changing ecosystem in which they operate, may warrant ex-post revision. When necessary, courts should exploit their power to revise their remedies in keeping with changing circumstances.<sup>230</sup> The example of the Wi-Fi problem in Germany, mentioned earlier, neatly illustrates the critical need for flexibility.<sup>231</sup> If the German courts had promptly considered adapting their original orders, which compelled private Wi-Fi providers to password-lock their services, when the new Wi-Fi-based technologies began blossoming outside Germany, they might have prevented the innovative setback that Germany suffered as a result of their technological remedies.<sup>232</sup> Indeed, ex-post revision of equitable remedies is tailored to meet the need for flexibility in remedies of injunction or specific performance.<sup>233</sup> This power enables courts to respond to events that were unforeseen when the remedy was first granted, because of changes in law or changes in fact, which typically occur in the digital ecosystem.<sup>234</sup>

---

229. *Id.* at 564.

230. *See id.* at 564–65.

231. *See generally* Mike Masnik, *German Court Says you Must Secure your WiFi or you may Get Fined*, TECHDIRT (May 12, 2012), <https://www.techdirt.com/articles/20100512/1116409394.shtml> [<https://perma.cc/95SU-ZBBS>].

232. *See, e.g.*, Loveday Wright, *Germany’s Wi-Fi Problem*, DW (Nov. 13, 2014), <https://www.dw.com/en/germanys-wi-fi-problem/a-18060000> [<https://perma.cc/6UMZ-NLKS>].

233. *See* Bray, *supra* note 21, at 564–65.

234. *See* Salazar v. Buono, 559 U.S. 700, 714–15 (2010) (plurality opinion) (“Because injunctive relief is drafted in light of what the court believes will be the future course of events, . . . a court must never ignore significant changes in the law or circumstances underlying an injunction lest the decree be turned into an instrument of wrong.”) (internal emphasis removed); King-Seeley Thermos Co. v. Aladdin Indus., Inc., 418 F.2d 31, 35 (2d Cir. 1969) (“While changes in fact or in law afford the clearest bases for altering an injunction, the power of equity has repeatedly been recognized as extending also to cases where a better appreciation of the facts in light of experience indicates that the decree is not properly adapted to accomplishing its purposes.”).

The *Sci-Hub* injunction, for instance, was amended soon after it was first issued, following the plaintiff's request to be given the authority to seize any and all Sci-Hub domain names, including those to be registered in the future.<sup>235</sup> In fact, the ease with which Sci-Hub could close existing domains and open new ones made the original order that targeted specific domains worthless. At the same time, however, content blocking may over-enforce plaintiffs' rights and block legitimate content, while restricting the fundamental rights of third parties that are not direct parties to the dispute and hence do not necessarily have standing to request injunction updates from the court.<sup>236</sup> For this reason, it is critical that courts independently invoke their power to modify remedies whose practical implementation is later found to exceed their original scope.

## 2. *Advising Technical Experts*

Furthermore, to subject digital remedies to meaningful oversight, it is vital that courts struggling to resolve cyber-related disputes understand the technological meaning of the relief they consider to grant. In its preliminary ruling in the Motion for Preliminary Injunction filed against TickBox, for instance, the court raised a handful of complex technological questions:

What is the best way to address the issue of themes (such as Paradox or Lodi Black) and/or addons (such as Covenant) that provide access to unauthorized versions of Plaintiffs' copyrighted work but that Device users have already installed? Is there a way to address this issue? Plaintiffs frame the solution as a simple software update whereby TickBox removes these previously-downloaded themes from its customers' Devices . . . . Is it possible to perform a similar software update whereby all Devices are reset, previously downloaded themes and addons are deleted, and TickBox's customers start anew with an offending-theme-free user interface?<sup>237</sup>

The court, however, did not attempt to answer these critical questions, but rather preferred to maintain the status quo and leave these questions for the parties to address.<sup>238</sup>

But the parties' technological expertise should not negate the need to empower courts with competent and professional capabilities. Out-of-court,

---

235. Ernesto, *Publisher Gets Carte Blanche to Seize New Sci-Hub Domains*, *supra* note 198.

236. Geiger & Izyumenko, *The Role of Human Rights in Copyright Enforcement Online*, *supra* note 85.

237. TickBox 1, *supra* note 68, at 1.

238. *Id.* at 2 ("Keeping these questions and the discussion that follows in mind, counsel for Plaintiffs and TickBox, working with others who possess relevant technical expertise as necessary, shall negotiate and attempt to reach agreement upon a stipulated preliminary injunction that will supersede the Court's initial preliminary injunction order.").



private negotiations about the qualities of a specific relief should not replace responsible decision making, which takes into account the full range of values and interests held by various stakeholders that might be affected by the relief. Specifically, counting on private, out-of-court settlements to reach the most appropriate solution ignores the robustness of digital remedies, the implications of which may far exceed the particular rights and interests of the direct parties to the legal dispute. As demonstrated in Part III, alternative technological solutions may vary in terms of cost, accuracy, and efficiency,<sup>239</sup> and these must be considered and assessed in an unbiased manner. In the United Kingdom, for instance, where blocking injunctions had become a very popular relief against online copyright infringement, Judge Richard Arnold, the undisputed authority when it comes to ordering ISPs to disable access to pirate websites, has been rolling up his sleeves to explore the practical meaning of each blocking alternative and ensure its overall proportionality.<sup>240</sup>

Enhancement of courts' oversight capacity could be achieved by appointing "equitable helpers" with the necessary technical expertise.<sup>241</sup> Particularly, Rule 53 of the Federal Rules of Civil Procedure authorizes judges to appoint special advisors<sup>242</sup> to aid them in handling pretrial matters tried without a jury that cannot be addressed effectively and promptly by available district or magistrate judges.<sup>243</sup> Accordingly, and despite the costs,<sup>244</sup> special masters have been called upon for their expertise in specific fields "such as

---

239. See *supra* Section III.C.

240. See Lindsay, *supra* note 88, at 1534–35. For instance, in *Twentieth Century Fox Film Corporation v. British Telecommunication P.L.C.*, Judge Arnold considered:

[T]he terms of an order requiring [the ISP] to implement [a] hybrid blocking system, concluding that it would be best to frame the injunction as requiring IP address re-routing (to the URL blocking) rather than IP address blocking, as the latter could be disproportionate in that it could result in over blocking [of legitimate speech].

*Id.* In *Dramatico Entertainment Ltd. v. British Sky Broadcasting*, on the other hand, Judge Arnold held that, "as IP address blocking might prevent circumvention, . . . it could be appropriate for [blocking] to be mandated, provided that the IP address was not shared with non-infringing websites." *Id.* at 1535.

241. See Bray, *supra* note 21, at 567–68.

242. *Id.* at 567. There are other authorities for appointing special masters. See, e.g., David I. Levine, *The Authority for the Appointment of Remedial Special Masters in Federal Institutional Reform Litigation: The History Reconsidered*, 17 U.C. DAVIS L. REV. 753 (1984); Wayne D. Brazil, *Authority to Refer Discovery Tasks to Special Masters: Limitations on Existing Sources and the Need for a New Federal Rule*, in MANAGING COMPLEX LITIGATION: A PRACTICAL GUIDE TO THE USE OF SPECIAL MASTERS 305 (W. Braz et al. eds., 1983).

243. FED. R. CIV. P. 53(a)(1)(C).

244. See Bray, *supra* note 21, at 574.

accounting, finance, science, and technology.”<sup>245</sup> Similarly, under Rule 706 of the Federal Rules of Evidence, “trial courts have wide discretion to appoint experts . . . to clarify issues under consideration,”<sup>246</sup> and there is also the “inherent authority [of federal courts] to appoint technical advisors.”<sup>247</sup> Hence, if legal disputes which “raise problems of unusual difficulty, sophistication, and complexity, or involve issues well beyond the regular questions of fact and law which judges routinely face” justify the appointment of technical experts and advisors,<sup>248</sup> then complicated and dynamic cyber-related disputes should also warrant such appointment.

### 3. *Imposing Duration Limitations*

Constructing equitable remedies in a flexible fashion is considered another equitable managerial device.<sup>249</sup> Specifically, courts could enhance their ability to supervise the implementation of digital remedies by limiting their duration in accordance with their relevance. Because the surrounding digital circumstances change rapidly, as do the technological capabilities to resolve digital problems, courts should regularly consider accompanying digital remedies with proper sunset clauses. Consider, for instance, a blocking order that blocks users’ access to a website providing unauthorized live streaming of the NBA finals. Such an order should be limited in time and not exceed the duration of the finals. Otherwise, the risk of over-enforcement and blocking of legitimate content will outweigh the benefit of decreasing copyright infringement.<sup>250</sup>

Limiting the duration of digital remedies will further facilitate their periodic review, which is necessary to allow courts to exploit their ex-post revision power in a timely manner and in light of experience.<sup>251</sup>

---

245. MANUAL FOR COMPLEX LITIGATION (FOURTH) § 11.52 (2004). For instance, the appointment of a special master in a case involving intellectual property claims by a manufacturer of medical devices against an inventor and his company who was requested to “mak[e] decisions with regard to search terms; oversee[ ] the design of searches and the scheduling of searches and production; coordinat[e] deliveries between the parties and their vendors; and advis[e] both parties, at either’s request, on cost estimates and technical issues.” *Medtronic Sofamor Danek, Inc. v. Michelson*, 229 F.R.D. 550, 559 (W.D. Tenn. 2003).

246. Maayan Filmar, *A Critique of In Re Bilski*, 20 DEPAUL J. ART, TECH. & INTELL. PROP. L. 11, 47 (2009).

247. *Id.* at 48.

248. *Id.*

249. See Bray, *supra* note 21, at 568; *supra* note 209.

250. In his blocking injunctions, Judge Arnold, for instance, has “recently imposed a sunset clause, which has time limitation [*sic*] of two years.” See Husovec, *supra* note 93, at 28.

251. See *Richemont Int’l SA v. British Sky Broad. Ltd.* [2014] EWHC (Ch) 3354, (Eng.) at 373.

For instance, restricting the TickBox injunction to a specified time limit could have enabled the court to promptly find out whether the applications that TickBox effectively deleted were indeed applications that “link[ed] directly . . . to third-party cyberlockers or streaming sites that transmit[ed] unauthorized performances of copyrighted motion pictures or television shows.”<sup>252</sup> Indeed, the court had originally raised its concern as to whether prior to deleting any software from TickBox’s current user interface, the parties ensured that it actually contained links to the apps or websites that provided access to unauthorized streaming versions of plaintiffs’ copyrighted works.<sup>253</sup> Yet, independently reviewing which software was deleted and deciding whether it induced copyright infringement, the court would have to essentially outsource their judicial discretion to private parties’ whose judgment might be mistaken or biased. Given the dramatic implications of erroneously restricting free speech, such restrictions should be addressed promptly.

#### 4. *Contempt*

“Equitable remedies may be enforced by contempt proceedings, through which a court may impose a range of highly discretionary punishments—including a new injunction, the payment of money to the plaintiff, the payment of fines to the state, or, less commonly, imprisonment.”<sup>254</sup> While this equitable device is not commonly used, it could nonetheless “allow the court to respond to new circumstances.”<sup>255</sup> Effectively, it allows the judge to direct, learn, respond, manage, or substitute for an alternative solution, “all with the goal of achieving the plaintiff’s rightful position.”<sup>256</sup>

Contempt proceedings could actually have a double effect. From an ex-ante perspective, they require courts to be as clear and precise as possible in defining the remedy,<sup>257</sup> and at the same time, encourage defendants to accurately follow the court’s instructions. From an ex-post perspective, like ex-post revision, contempt allows courts to adjust the relief if its practical implementation is found to exceed or override its intended reach. Note that since courts retain the power to review and adjust the remedies they grant,

---

252. TickBox 2, *supra* note 69, at 1.

253. TickBox 1, *supra* note 68, at 1.

254. Bray, *supra* note 21, at 565–66.

255. *Id.* at 566.

256. *Id.* at 567; *see also* DOUG RENDLEMAN, COMPLEX LITIGATION: INJUNCTIONS, STRUCTURAL REMEDIES, AND CONTEMPT 691–833 (2010).

257. *See* Schmidt v. Lessard, 414 U.S. 473, 476–77 (1974) (per curiam).

detailed architecture of remedies should not diminish their necessary flexibility.<sup>258</sup>

#### 5. *Encourage Ongoing Participation of Various Stakeholders*

Finally, another mechanism that could facilitate better oversight of the implementation of digital remedies is to give voice to affected users—not only during the initial legal procedure, but also during the subsequent ex-post revision procedures.<sup>259</sup> To avoid lengthy litigation, such participation of interested parties should only be allowed during strict time windows. Since the private, out-of-court implementation of digital remedies may unfold in an unexpected fashion, it is important to allow those whose rights are being affected, as well as those representing various public interests, including non-profits, human rights organizations, law enforcement agencies, and government representatives, to express their concerns before the court and demand the revision of digital remedies that are inefficient or disproportionate (e.g., restricting users' access to legitimate online content). This is especially important in cases where the specific procedural process governing the case negates the possibility of public participation during the early, ex-ante stage of in-court proceedings.

---

258. One example of detailed digital remedy is the blocking order, which was granted in *Richemont Int'l SA v. British Sky Broad. Ltd.* [2014] EWHC (Ch) 3354, (Eng.), at 319 which reads as follows:

In respect of its residential fixed line broadband customers [ . . . ], the [ . . . ] defendant [ISP] shall within 15 working days in relation to the initial notification (and thereafter, within ten working days of receiving any subsequent notification) adopt the following technical means to block or attempt to block access to the target websites, their domains and sub-domains and any other IP address or URL notified to the . . . defendant whose sole or predominant purpose is to enable or facilitate access to a target website. The technology to be adopted is:

- (i) IP blocking in respect of each and every IP address from which each of the target websites operate and which is [ . . . ] notified in writing to the . . . defendant by the applicants or their agents [ . . . ]
- (ii) IP address re-routing in respect of all IP addresses that provide access to each and every URL available from each of the target websites and their domains and sub-domains and which URL is notified in writing to the . . . defendant by the claimants or their agents; and
- (iii) URL blocking in respect of each and every URL available from each of the target websites and their domains and sub-domains and which is notified in writing to the . . . defendant by the [applicants] or their agents.

*Id.*

259. Mulligan & Bamberger, *Saving Governance by Design*, *supra* note 17, 772–73.

For instance, the FBI's request to the court to force Apple to create software to help them circumvent the phone's encryption was initially submitted as an ex-parte demand.<sup>260</sup> If the FBI had not submitted a subsequent motion to compel Apple to comply with the assistance order,<sup>261</sup> the ex-parte demand would have deprived the court of the perspectives of Apple and numerous organizations that raised diverse concerns about the FBI's request.<sup>262</sup> Similarly, the *Sci-Hub* injunction was ultimately granted as a default judgment, without the defenses of the allegedly direct infringer (i.e., the operator of the Sci-Hub site) or the ultimate enforcers (i.e., various service providers) being heard.<sup>263</sup>

Moreover, encouraging ex-post participation of affected users is especially important for digital injunctions that are directed to non-parties to the legal dispute (e.g., the *Sci-Hub* injunction).<sup>264</sup> Third parties that are required to implement a court order, even though they did not actively represent their interests during the ex-ante judicial procedures,<sup>265</sup> should at least be allowed to deliver their concerns during the stage of ex-post revision considerations. Firstly, because they are not regular non-parties whose interests are affected from the injunction, but are the long hand of the defendants that are effectively expected to obtain the resolution of the case, sometimes even on behalf of the defendants. Secondly, and relatedly, because the economic expenses of executing the remedy could be quite substantial.<sup>266</sup> Thirdly, because when digital remedies delegate adjudication powers to these third parties, directing them not only to choose *which* technological means to apply, but also to decide *how* to implement these means, it is important to provide them with an open

---

260. Government's Ex Parte Application for Order Compelling Apple Inc. to Assist Agents in Search at 1–2, *In re Search of an Apple iPhone*, <https://epic.org/amicus/crypto/apple/In-re-Apple-FBI-AWA-Application.pdf> [<https://perma.cc/T2SK-JP74>].

261. See Government's Motion to Compel Apple Inc., *supra* note 72.

262. See Mulligan & Bamberger, *Saving Governance by Design*, *supra* note 17, at 723.

263. See Diana Kwon, *American Chemical Society Wins Lawsuit Against Sci-Hub*, SCIENTIST (Nov. 7, 2017), <https://www.the-scientist.com/news-opinion/american-chemical-society-wins-lawsuit-against-sci-hub-30648> [<https://perma.cc/A8JM-S9D3>].

264. See *supra* Section III.B.

265. Generally, many courts apply a four-factored test for issuing preliminary injunctions, which inquire into: (1) whether the plaintiff will suffer irreparable harm absent the issuance of an injunction; (2) how the harm suffered by the plaintiff absent an injunction balances against the harm that an injunction would cause to the defendant; (3) the plaintiff's likelihood of success on the merits; and (4) the public interest. See CHARLES ALAN WRIGHT ET AL., FEDERAL PRACTICE AND PROCEDURE § 2948, 133 (2d ed. 1995). “Many courts interpret the public interest factor as a license to consider the impact that granting or denying injunctive relief will have on non-parties.” Laura W. Stein, *The Court and the Community: Why Non-Party Interests Should Count in Preliminary Injunction Actions*, 16 REV. LITIG. 27, 29 (1997).

266. See *supra* Section III.C.

judicial venue where they can seek technical advice and obtain feedback about their specific compliance. Otherwise, they might be left alone in the battlefield of compliance, which might encourage them to prefer robust technological means with a higher risk of over-enforcement,<sup>267</sup> over specifically tailored reliefs that are more accurate, but might result in under-enforcement.<sup>268</sup>

#### B. EQUITABLE CONSTRAINTS

The exploitation of the various managerial devices discussed above can be costly, both on the part of the court, especially when nominating technical advisors,<sup>269</sup> and on the part of defendants, especially when required to adjust their compliance in accordance with the changing digital circumstances.<sup>270</sup> Indeed, “equitable remedies have certain characteristic costs, especially the direct and indirect costs of complying with the court’s command and the possibility of an afterlife in which that command is clarified, modified, enforced, or dissolved.”<sup>271</sup> This is why equitable enforcement tools are subject to various limits.

For instance, there is the doctrine of ripeness, which ensures “the appropriateness of judicial review” in a given case.<sup>272</sup> “Ripeness is especially important for equitable remedies because they can depend on facts that are changing and contingent, and [they] can entangle the courts in the relationship of the parties, not just at the moment of decision but . . . on an ongoing basis.”<sup>273</sup> In particular, with regards to digital remedies, it is important to ensure the recourse they provide remains relevant. Additionally, there is the requirement for specificity, “which requires that an equitable decree be precisely worded and give clear notice of what is prohibited and required.”<sup>274</sup> Another limit on the use of equitable managerial devices relates to equitable defenses that prevent “the power of these remedies to be used on behalf of a

---

267. Such as IP blocking. *See supra* Section III.C.

268. Such as the DNS blocking technique of content blocking orders. *See* Feiler, *supra* note 132 and accompanying text.

269. Bray, *supra* note 21, at 573–74.

270. Gene R. Shreve, *Federal Injunctions and the Public Interest*, 51 GEO. WASH. L. REV. 382, 389 (1983) (“[An injunction] poses the threat of adjusting more aspects of the defendant’s behavior than those that would wrong the plaintiff if the injunction were not issued. It is difficult if not impossible to so finely adjust an order that it protects plaintiff without impairing defendant’s harmless activities or the rights of those who are not represented before the court.”).

271. Bray, *supra* note 21, at 577.

272. *See, e.g.,* G. Joseph Vining, *Direct Judicial Review and the Doctrine of Ripeness in Administrative Law*, 69 MICH. L. REV. 1443, 1446 (1971).

273. Bray, *supra* note 21, at 579.

274. *Id.*; *see also* FED. R. CIV. P. 65(d).

plaintiff who acts unjustly.”<sup>275</sup> For example, plaintiffs cannot bring their claims with unreasonable delay or with unclean hands.<sup>276</sup> Overall, these discretionary limits “focus[] judges’ attention on certain situations where equitable remedies and enforcement mechanisms are most likely to be misused.”<sup>277</sup>

## VI. CONCLUSION

“The devil is in the details,” or its predecessor “God is in the details,” means that “[t]he details of a plan, while seeming insignificant, may contain hidden problems that threaten its overall feasibility.”<sup>278</sup> This phrase captures the precise implication of using technological fixes as solutions for legal disputes: the details underlying such fixes are far from merely procedural. They are actually material, shaping the crux of the technological plan for resolving a concrete legal dispute. Digital remedies change the traditional dichotomy between adjudication and compliance in remedies. They blur the borderline between law making and law enforcement, depositing both powers in the hands of private executors who design and implement the remedy outside the courthouse.

As explained in this Article, digital remedies can be implemented through various means, which differ in their error rate, costs, and circumvention potential. These differences are substantial, as they effectively define the ultimate scope and breadth of the relief. The robust implementation of digital remedies can effectively reshape settled balances between clashing rights and interests and practically dictate progress and innovation.

This critical role of the technical details which underline digital remedies challenges the ability of courts to competently oversee the remedial process. Traditional mechanisms of judicial oversight do not fit the realm of digital remedies. Specifically, ex-ante judicial review, transparent legal procedure, and public participation during legal proceedings ignore all that happens after the court issues its decree, when private, profit-maximizing executors embed their technological choices in non-transparent and proprietary technologies.

An all-embracing perspective of checks and balances is needed to facilitate ongoing, ex-post review of digital compliance, to protect the rule of law,

---

275. Bray, *supra* note 21, at 581.

276. Howard W. Brill, *The Maxims of Equity*, 1993 ARK. L. NOTES 29, 34 (1993) (“The purpose of the unclean hands doctrine is neither to protect the defendant nor to favor the complainant . . . [but] to protect the court . . .”).

277. Bray, *supra* note 21, at 584.

278. See *The Devil is in the Details*, PHRASES FINDER, <https://www.phrases.org.uk/meanings/the-devil-is-in-the-details.html> [<https://perma.cc/HLY7-XKCP>] (last visited Jan. 4, 2020).

consider the various rights and interests at stake, and ensure that digital remedies adapt to a rapidly changing digital reality. As suggested in this Article, the exploitation of equitable managerial devices could advance such an all-round perspective, while empowering courts' oversight capabilities. Specifically, by consulting technical experts to hone their technical understanding and implications of the reliefs that will be granted; by supporting ex-post revision of decrees and limiting their duration to address the need for constant adaptation; and by encouraging ongoing participation of various stakeholders to facilitate a broad consideration of human rights and public values, courts could enhance their oversight capabilities while responding properly to the increasing need to resolve cyber-related disputes.



# PATENT MARKETS AND INNOVATION IN THE ERA OF BIG PLATFORM COMPANIES

Robert P. Merges<sup>†</sup>

“[A] rivalrous structure surely has its inefficiencies. But such a structure does tend to generate rapid technical progress and seems a much better social bet than a regime where only one or a few organizations control the development of any given technology.”

—Robert P. Merges and Richard R. Nelson<sup>1</sup>

## ABSTRACT

In many industries, the arc of our contemporary economy bends towards bigness. The ubiquitous digital platform companies such as Amazon, Facebook, Netflix, Chinese companies like Baidu, Tencent, and Alibaba are the best-known examples. While some concerned onlookers propose structural remedies,<sup>2</sup> America’s constrained antitrust law plus the logic of natural monopoly mean that increased concentration will likely continue for the foreseeable future. In this setting, it is important to preserve multiple sources of rivalrous innovation despite continuous growth in the Big Platforms. Preserving rivalry requires carving out and preserving a niche for innovative small and medium-sized companies. One way to do this is to promote and protect the secondary patent market. Sale of patents is one way small firms can remain viable in the shadow of Big Platforms. This Article argues that patent markets are superior in some cases to complete acquisition of a small firm by a Big Platform company because selling patents allows a small firm to survive as an independent entity. Recent patent system reforms support this pro-secondary market policy: the era of easy and extortionate patent litigation, traditionally associated with the secondary patent market, is coming to a close. Patent sales and licensing, at times backed by the threat of litigation, will promote small company innovation once these reforms gain traction. This is crucial; though Big Platforms are currently young and vigorous, history suggests that they will become less innovative in the long run. Preserving multiple small innovators—through the patent market and otherwise—is the best way to prepare for the future of Big Platforms.

---

DOI: <https://doi.org/10.15779/Z38BZ6185G>

© 2020 Robert P. Merges.

<sup>†</sup> Wilson, Sonsini, Goodrich & Rosati Professor of Law, University of California, Berkeley, School of Law.

1. Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 908 (1990).

2. See, e.g., Lina M. Khan, *Amazon’s Antitrust Paradox*, 126 YALE L.J. 710, 790 (2017).

## TABLE OF CONTENTS

<b>I.</b>	<b>INTRODUCTION .....</b>	<b>55</b>
<b>II.</b>	<b>PATENT ASSETS AND PATENT MARKETS .....</b>	<b>59</b>
A.	ARE PATENTS DIFFERENT FROM OTHER TYPES OF CORPORATE PROPERTY? .....	60
B.	PATENT PORTFOLIOS AS BUNDLES OF ASSETS: RELATIONSHIP TO CORPORATE LAW THEORY .....	64
C.	A TYPOLOGY OF TECHNOLOGY MARKETS, AND THE ROLE OF IP RIGHTS .....	67
1.	<i>The Role of Patents in the Spinoff of “Orphan” Technologies</i> .....	71
a)	Abandoned Projects .....	72
b)	Re-Directed Research .....	73
c)	Multi-Application Research .....	73
2.	<i>Failed Product Companies and the Market for Patents</i> .....	75
a)	Failed-Product Companies and Patent Litigation: Ex-Post Market Making .....	77
b)	Summary: The “Two Period” Nature of Patents and Patent Litigation .....	80
<b>III.</b>	<b>PATENT MARKETS, MERGERS, AND R&amp;D: WHAT DO THE DATA SAY? .....</b>	<b>81</b>
A.	THE “TACIT DIMENSION” AND MARKETS FOR “DISEMBODIED TECHNOLOGIES” .....	85
B.	AREN’T FIRM MERGERS AND ACQUISITIONS (ALMOST) ALWAYS SUPERIOR TO PATENT SALES? .....	86
1.	<i>For Radical Innovation, More Is Better and Small Is Big</i> .....	87
2.	<i>Small Is Big</i> .....	87
3.	<i>Innovating “Outsiders”: A Complicating Factor?</i> .....	92
C.	MERGERS AND INDUSTRY STRUCTURE: SUMMARY .....	94
<b>IV.</b>	<b>SUGGESTED REFORMS TO ASSIST THE PATENT MARKET ...</b>	<b>96</b>
A.	RELATIONSHIP TO LITIGATION: DO PATENT MARKETS “FEED THE TROLLS”? .....	96
B.	POLICIES TO SUPPORT THE SECONDARY PATENT MARKET .....	96
1.	<i>Antitrust Law</i> .....	96
a)	Towards a Consideration of Potential Future Innovation .....	102
i)	<i>Preserving a Future Disruptor</i> .....	102
b)	Patent Markets and the Future Competitive Landscape .....	105
2.	<i>Smoothing the Patent Market</i> .....	107

- a) Recording of Patent Assignments, Licenses, and Other Interests .....107
- b) Facilitating Transfers When Patents are Being Challenged in the Patent Office.....109

## V. CONCLUSION..... 111

### I. INTRODUCTION

Economic activity is always ultimately about buyers and sellers. Today, this activity increasingly takes place online. Five huge companies have emerged in the United States as makers of mass-scale markets.<sup>3</sup> Many other companies, both in the United States and elsewhere, are working to mediate between buyers and sellers in all sorts of industries. The basic logic of network economics pushes platform companies to continually expand in scale and scope: scale means more buyers and sellers, while scope means more markets served. What this ultimately means is that the era of the Big Platform has begun.<sup>4</sup>

---

3. Farhad Manjoo, *The Upside of Being Ruled by the Five Tech Giants*, N.Y. TIMES (Nov. 1, 2017), <https://www.nytimes.com/2017/11/01/technology/five-tech-giants-upside.html> [<https://perma.cc/EL5Z-B7YX>] (describing the role of five tech giants on the market).

4. The word “platform” has taken on a constellation of meanings, which often vary depending on the subject matter specialty of the speaker. In business strategy, a common set of components that form the core of a machine, software system, or the like may be called a platform; an example would be Microsoft Windows. See Carliss Y. Baldwin & C. Jason Woodard, *The architecture of platforms: a unified view*, in PLATFORMS, MARKETS AND INNOVATION, 21 (Annabelle Gawer ed., 2009). Many computer programs, sold by many different companies, can “plug into” the Windows operating system, making Windows a frequently-referenced “software platform.” More recently, engineers and economists have reserved the word “platform” to refer to any physical or virtual thing, place, or system that brings together multiple sellers and multiple buyers of products and/or services. These are often (and more properly) referred to as “two-sided platforms.” Thus, a shopping mall (a building with multiple separate units for lease) brings together sellers and buyers of retail goods. Today, virtual platforms such as Amazon, Uber, and YouTube are much in the news because of their growing size and power. Amazon brings together buyers and sellers of a huge range of goods and services. Uber brings together (or “intermediates” between) independent drivers and riders. YouTube (along with Spotify and the like) brings together producers and consumers of content (video, audio, etc.). In a more general sense, all-purpose search engines such as Google also serve as two-sided platforms, bringing together advertisers and consumers, though in this and other cases of “ad-supported” content, ads are often an extraneous intrusion into the content or information sought by the consumer. These platforms might be said to bring together producers and consumers of information, in a format subsidized by advertisers. The advertisers *use* the platform to attract customers, even though the customers are (usually) not on the platform for the express purpose of

It is now commonplace to worry about the massive size of Big Platforms. Competition and antitrust law experts will be heavily debating these issues in the years to come. The issues will be complex—Big Platforms have also destabilized conventional assumptions and practices in fields such as employment law (e.g., Uber, Lyft, and Didi in China), local regulation (e.g., Airbnb), and taxation (e.g., Amazon).

One side of the platform debate touts the advantages of size and scale in innovative industries. A well-established school of thought says that size and the accompanying market power are the best friends that innovation could ever have.<sup>5</sup> In addition, legal scholar Peter Lee has shown that technological skills are deeply embedded in pioneering companies, which makes a strong case for big companies to keep growing through company acquisition.<sup>6</sup> The buying up of talent in the form of big companies acquiring smaller ones even has a name: “acqui-hiring.”<sup>7</sup>

As with most technologies, online platforms are based on a wide range of innovations spanning many years, including the internet itself, mobile communications, data compression technologies, online payment systems, and GPS satellites and mapping software. These innovations represent the successful harvest of many scientific and technological seeds planted at various times over fifty years. The seeds for these technologies were planted in many different places: the public sector and universities as well as big, medium, and small companies.

This creates a cause for concern: in the era of the Big Platform, will there still be room for such a varied innovative ecosystem? Will the trend toward “bigness” and the “winner take all” nature of platform markets shut out the smaller innovators that have helped create the conditions in which the platform economy thrives?

A detailed answer would have many parts and cover many topics: the future of government research and development (R&D), the prospects for university research, and the pros and cons of innovation driven by company acquisitions. This Article stresses only one of these themes: the importance

---

looking at ads. These two-sided platform companies are the ones concentrated upon in this Article.

5. See *infra* notes 48–57 and accompanying text.

6. See generally Peter Lee, *Innovation and the Firm: A New Synthesis*, 70 STAN. L. REV. 1431 (2018).

7. Andres Sawicki, *Buying Teams*, 38 SEATTLE U. L. REV. 651 (2015); Samantha Nolan, *Talent for Sale: The Need for Enhanced Scrutiny in Judicial Evaluation of Acqui-Hires*, 67 HASTINGS L.J. 849, 849 (2016) (stating that in “acqui-hiring[,] [t]he buying corporation purchases the target, poaches its employees, jettisons its projects, and generally kills the company”; calling for shareholder protections for the acquired firm).

of markets for technology. The term refers to the ability to transfer technologies through the mechanism of patent acquisitions—arm’s length sales of discrete technologies rather than of the companies that developed them.<sup>8</sup> When individual technologies can be transferred to big platform companies, the smaller companies that developed those technologies can continue to exist as going concerns. The people inside these smaller companies can therefore retain the benefits of autonomy and independence, despite the vertical integration that typically accompanies the platform economy. The market for technology essentially permits vertical integration of *technologies* without requiring the swallowing up of *entire companies*—an arrangement that has some distinct advantages.<sup>9</sup>

The main point of this Article is to emphasize the advantages of patent markets and continuing small firm viability. At the outset, however, it must be said that there are good reasons for vertical integration in the era of Big Platforms. The growth of companies by sequential firm acquisition has real benefits. It certainly is a boon for small company founders; today’s golden exit for many startups is a phone call from a Big Platform company saying “we want to buy you.”<sup>10</sup> Meanwhile, Big Platform technologies and business strategies reward size and scope, which are often achieved faster by a combination of internal growth and external firm acquisitions. Nor does a Big Platform company buying a startup always mean that small company talent is permanently absorbed. Some startup founders are “serial

---

8. As explained later, most patent-related transactions these days are for patent portfolios, rather than for individual patents. *See infra* note 23 and associated text.

9. *But see* Joshua Gans et al., *When Does Start-up Innovation Spur the Gale of Creative Destruction?*, 33 RAND J. ECON. 571 (2002). The paper presents an empirical study of 100 startups and finds that the probability of cooperation with incumbent firms, as opposed to entry into product competition with them is increasing in the innovator’s control over intellectual property rights, association with venture capitalists (which reduce their transactional bargaining costs), and in the relative cost of control of specialized complementary assets. The authors conclude that the propensity for pro-competitive benefits from start-up innovators in the form of product market entry reflects an earlier market failure, in the market for ideas. For Gans et al., then, a strong market for technology and/or patent market (as explored later in this Article) actually *contributes to* the concentration of power in fewer (presumably larger) firms: the opposite of this Article’s thesis. Two things to note about this Article are: (1) it was written before the Big Platform companies had fully emerged, so entry into product market competition in information technology industries was more common; and (2) the “strong IP” industries studied clustered around pharmaceuticals, an industry in which entry barriers associated with the high cost of pharma research and regulatory approval mean that the entry of new, full-scale pharma firms to compete with incumbents is a very rare event.

10. *See, e.g.*, JOHN HAWKEY, EXIT STRATEGY PLANNING: GROOMING YOUR BUSINESS FOR SALE OR SUCCESSION 130 (2014).

entrepreneurs” who go on to start another new company after their current company is swallowed up.<sup>11</sup>

Even so, there are good reasons to favor a diverse economic ecosystem that includes ongoing, continuously operating small firms in highly innovative industries. As explained later, substantial research shows that small companies are by many measures more innovative than large ones. Because patent markets enable technologies to move from smaller to larger companies without requiring smaller companies to be completely swallowed up, these markets can play an important part in preserving a more diverse industry structure in innovative industries. The big firms can thrive by acquiring the technologies they need to expand and grow, while at least some smaller firms can remain independent. This gives the smaller firms a better chance to contribute new and valuable innovations down the road.

Sale of technologies gives smaller companies a route through which they can participate in incremental innovation in the platform era while retaining their independent and autonomous cultures. This could help push against the overwhelming forces driving toward centralization, consolidation, and vertical integration. It just might even foster the kind of “outsider” mentality that so often begins the process of creative destruction. The ultimate reason for fostering technology markets in the platform era is to open the way for the beginnings of whatever era will succeed it.

An active patent market would serve as a supplement to in-house R&D, which can be expected to grow along with Big Platform companies. An increase in big company research is likely, judging from earlier waves of vertical integration. Twentieth century companies such as the Pennsylvania Railroad, Carnegie Steel, General Electric, DuPont, AT&T, and the “Big Three” U.S. automakers pioneered the raw-materials-to-end-user corporate architecture. One aspect of this was the development of modern in-house R&D laboratories. There are signs that the Big Platform companies are moving in this direction, especially in the case of Amazon’s 126 Lab.<sup>12</sup> If the

---

11. Most big companies require the founders and other employees of acquired companies to remain as employees for a period of time; they do this by “vesting” the big company stock over two to four years, and which is the normal compensation for the founders who sell out. *See, e.g.,* Thomas Goetz, *Startup. Get Ready for a Demotion and an Identity Crisis*, INC. MAG. (June 2019), <https://www.inc.com/magazine/201906/thomas-goetz/exit-acquisition-merger-after-sale.html> [<https://perma.cc/US4W-WQ9G>] (“[M]ost startup acquisitions come with the golden handcuffs of a two- or four-year vest . . .”).

12. Amazon’s 126 Lab created the Kindle e-book reader and the Alexa voice-recognition Amazon interface. *See* Mark Gruman & Brad Stone, *Amazon Is Said to Be Working on Another Big Bet: Home Robots*, BLOOMBERG NEWS (Apr. 23, 2018), <https://www.bloomberg.com/news/articles/2018-04-23/amazon-is-said-to-be-working-on-another-big-bet-home-robots> [<https://perma.cc/S9PZ-Y9HX>]; Ry Crist, *Behind the Scenes at Alexa’s*

Big Platforms follow the traditional arc, in-house R&D will very likely continue to grow.

The nature of technology also contributes to the logic of large firm size and vertical integration. Peter Lee has identified some substantial benefits from in-house R&D and outright ownership (through acquisitions) of R&D-related assets.<sup>13</sup> Technology is usually not a disembodied commodity that can be bundled up and sold in a store, but rather a subtle mix of codified information and hard-to-pin-down know-how. There can be little tricks to make software code harder to hack, or knowledge about the right way to tweak the settings in a metal fabrication process to get the strongest alloy possible. A company can only come to own and control these “tacit” aspects of technology by either growing technology in-house or acquiring the people, machines, and buildings of an entire company.

With the advantages of in-house research and outright acquisitions, why worry about a third path that requires an arm’s-length market for technology? The answer primarily stems from two principles: diversity and autonomy. To preserve a diverse ecosystem in the era of the Big Platform, technology markets are imperative. Only through an arm’s-length transaction can a distinct, separate innovative company find an outlet for its new ideas. Only with many such small companies operating on their own can we avoid the inevitable problems of “groupthink,” not invented here, and the other ills of bigness. Only through a market for technology can a small team of experts constitute themselves as a specialty supplier that remains independent of a large company—in other words, an autonomous economic unit. Some may regard these values as unimportant or overblown, but those who recognize that these traits gave rise to Big Platforms in the first place will be interested in preserving them. This translates to concerns for the health and well-being of a robust market for independently developed technology. Any detailed discussion of that market, however, requires covering some basics about the nature of patents and the market for them.

## II. PATENT ASSETS AND PATENT MARKETS

This Article so far has mainly discussed the patent market in terms of the contribution it makes as an alternative to full-firm mergers and acquisitions.

---

Laboratory, CNET (Apr. 23, 2018), <https://www.cnet.com/news/behind-the-scenes-at-amazon-alexa-laboratory-lab126/> [<https://perma.cc/X32H-2376>].

13. Peter Lee, *Innovation and the Firm*, *supra* note 6; *see generally* Peter Lee, *Transcending the Tacit Dimension: Patents, Relationships, and Organizational Integration in Technology Transfer*, 100 CALIF. L. REV. 1503 (2012).

This Part answers questions about the nature of the assets that are transferred in this market. Why are patents a useful asset type for transferring rights over technologies? How are patents superior to the sale of technology through contracts alone? How does corporate ownership of patents interrelate with the nature of corporations themselves?

A. ARE PATENTS DIFFERENT FROM OTHER TYPES OF CORPORATE PROPERTY?

While it is convenient that a corporation can sell off a patent portfolio, it might not seem significant. After all, given that a company can sell a used machine, truck, furniture, or any other type of personal property, what is so special about patents?

In one sense, nothing. Property in a patent is no different from other property.<sup>14</sup> So patents are just one of many things a company can sell when and if it chooses to.

But in another sense, patents are different. A central quality of property is that it confers broad control rights on an owner. In contrast to a simple buy-sell contract, for example, selling an asset subject to a property right does not require writing down in detail all the ways the buyer can use the asset in the future.<sup>15</sup> Legal academics say this wide discretion in deciding what can be done with an asset is the core feature of property. The “right to exclude” everyone else from using an asset leaves property owners with almost unfettered discretion in determining how it may be used.

Economists likewise think of property as an entitlement distinct from contract. Allocating rights and duties by contract is to them a basic feature of economic activity, but it is difficult, expensive, and theoretically impossible to specify *all* the rights and duties of two contracting parties regarding an asset.

---

14. Patents are exclusive rights, just as personal property is the right to exclude others from using an object:

The right to exclude others is the essence of the human right called “property.” The right to exclude others from free use of an invention protected by a valid patent does not differ from the right to exclude others from free use of one’s automobile, crops, or other items of personal property.

*Panduit Corp. v. Stahl Bros. Fibre Works, Inc.*, 575 F.2d 1152, 1158 n.5 (6th Cir. 1978). *But see* Michael H. Davis, *Patent Politics*, 56 S.C. L. REV. 337, 386 (2004) (asserting that “[c]alling patents ordinary property, and, more importantly, treating those rights as such seems slightly irrational”).

15. *See* Henry E. Smith, *Property as the Law of Things*, 125 HARV. L. REV. 1691, 1704 (2012) (“When O1 owns Blackacre, the exclusion strategy for delineating her rights, implemented through devices like the tort of trespass, protects a range of actions A1, A2, A3 . . . without the law’s needing to specify these actions.”).



Property rights are therefore necessary and crucial for economic exchange. Property gives an owner *residual rights*: the right to all uses of an asset *not* specified in a contract. From both a legal and economic perspective, ownership means a wide and full scope of control over the uses of an asset—known and unknown, present and future.

It is this feature of property rights that is so important for patents. A patent gives its owner control rights over all embodiments of a claimed invention. Unlike other types of assets, however, technology is not static. R&D leads researchers in many directions, many of which are unpredictable.<sup>16</sup> Therefore, broad control rights over many variations on a basic theme and over currently unforeseen applications of a technique or design are especially important for new technologies and R&D activity generally. This is exactly what you get with a patent.<sup>17</sup>

Thus, selling a machine or truck is different from selling a patent. The buyer of the machine or truck can do anything he or she wants with it (as long as the use is legal), so personal property in trucks and machines adds some value. It would be burdensome for the company to specify all the things the buyer can do with the truck or machine, and property makes this unnecessary. However, listing future uses for a truck would be difficult but not intractable. The foreseeable uses of the truck, driving, carrying, delivering, etc., are legion but not infinite.

---

16. According to Henry Smith:

[T]he uses of an asset are not just risky (e.g., with a variance in outcomes forming a probability distribution), but uncertain, in the Knightian sense. That is, the set of uses of an asset may not correspond to a known probability distribution, and nonowners may not even know the members of the set. Property law helps manage this uncertainty by not making knowledge of the uses or even the probability distributions of their values relevant to dutyholders. In previous work, I have argued that Knightian uncertainty is more conducive to property rules than to liability rules, which do require more knowledge of probabilistic information by officials or courts.

Henry E. Smith, *Institutions and Indirectness in Intellectual Property*, 157 U. PA. L. REV. 2083, 2088 (2009) (footnote omitted) (citing FRANK H. KNIGHT, RISK, UNCERTAINTY AND PROFIT 19–20, 197–232 (1921)) (distinguishing “risk” from “uncertainty” by noting that “uncertainty” is immeasurable in principle); Henry E. Smith, *Property and Property Rules*, 79 N.Y.U. L. REV. 1719, 1724–27 (2004) (“Property itself is a response to uncertainty, and property rules derive some advantage as a response to uncertainty.”).

17. On this, see Smith, *Institutions and Indirectness in Intellectual Property*, *supra* note 16, at 2106 (“For property, and intellectual property especially, the discovery of options (rather than the measurement of the value of options based on known risks) is something that the indirect modular structure of property tends to foster.”).

The same cannot be said about patents. A patent on a mediocre glue might make “Post-It Notes” possible.<sup>18</sup> Invention of a nonstick coating for cookware might enable someone to make rainproof cloth that is still breathable.<sup>19</sup> A patent on a mildly useful compound for one medical condition might open the door later to an effective treatment for a major disease or problem.<sup>20</sup> The list of examples goes on.<sup>21</sup>

The essence of a patent is an extrapolation from one or a few prototypes, successful experiments, or working models. Those who draft patent claims work every day in the realm of projection, extension, variation, and modification. Even within a single patent, the usual practice is to draft a set of claims that begins broadly and then becomes narrower. This pattern is repeated several times in a typical patent. Thus, from an economic perspective, the best way to conceptualize a patent is as a set of nested options. When a patent is filed or a claim is redrafted (amended), it is impossible to know for certain whether that claim will cover (read on) a valuable commercial product (embodiment) in the future. There is also a risk

---

18. See Acrylate Copolymer Microspheres, U.S. Patent 3,691,140 (issued Sept. 12, 1972). For the story behind the invention of the post-it note, see *About Us*, 3M, [https://www.post-it.com/3M/en\\_US/post-it/contact-us/about-us/](https://www.post-it.com/3M/en_US/post-it/contact-us/about-us/) [<https://perma.cc/FR8M-7HJS>] (last visited Dec. 21, 2019).

19. A DuPont researcher named Roy J. Plunkett invented polytetrafluorethylene (PTFE), trademarked as “Teflon,” in 1938. See Robert W. Gore, SCI. HIST. INST., <https://www.sciencehistory.org/historical-profile/robert-w-gore> [<https://perma.cc/JRM3-ZK4Y>] (last visited Dec. 21, 2019). One of Gore’s co-workers was W.L. Gore, who left DuPont to develop new applications of PTFE—one of which (in research with Gore’s son, Robert Gore) led to the surprising discovery that quick-stretching PTFE made a thin film that was air permeable but waterproof: Gore-Tex. *Id.* One early Gore-Tex patent is Waterproof Laminate, U.S. Patent 4,194,041 (issued Mar. 18, 1980).

20. See, e.g., Douglas Martin & Gunter Kahn, *Inventor of Baldness Remedy, Dies at 80*, N.Y. TIMES, Sept. 19, 2014, at A21 (describing Gunther Kahn’s discovery that a failed ulcer treatment called minoxidil was quite effective at stimulating hair growth); see also Methods and Solutions for Treating Male Pattern Alopecia, U.S. Patent 4,596,812 (issued June 24, 1986); Rebecca S. Eisenberg, *The Problem of New Uses*, 5 YALE J. HEALTH POL’Y, L. & ETHICS 717, 724 (2005) (“[C]linical trials showing that a drug works for a new indication may support a process patent on a new method of treatment, even though the same drug has previously been used for another purpose.”); see generally Kathryn Brown, *Repurposing Old Drugs for New Uses*, 28 DEPAUL J. ART, TECH. & INTELL. PROP. L. 1 (2017).

21. Keep in mind the distinction between the specific features of a technology and the applications for or uses of that technology. Features must be described in order to obtain a valid patent; that is the essence of the enablement requirement in 35 U.S.C. 112. But applications or uses are a different matter. Thus, one must describe in detail how to make and use a new metal alloy if that is the new invention claimed. But a valid claim to the alloy will in general cover all future applications and uses—in machinery, autos, high-speed trains, aircraft, bicycles, and even things not yet invented at the time the alloy patent is issued such as zero gravity machines or building-sized hovercraft.

that a broader claim may encompass something known in the field before the claim was filed, making that claim invalid. As a result, patent drafters are forever navigating the eternal golden braid of validity risk, legitimate extrapolation (enablement), and future coverage.<sup>22</sup> But the better the claims are drafted, and the more of them there are, the more likely that something of future value will be covered.

Additionally, the real-world unit of analysis these days is a patent portfolio rather than a single patent.<sup>23</sup> Most portfolios also include pending patent applications which, unlike issued patents, can still be amended. Their claims can be stretched, where legitimate, to cover products that have become viable or foreseeable in the interval between the filing of the original claim and the amendment. These pending applications and their claims thus have even greater option value. The result of this setup is a large bundle of ownership claims over a multitude of technological options. The options cover embodiments that may be hard or impossible to foresee, and it is equally hard to predict the market value of these unpredictable embodiments.

---

22. For just one of the thousands of examples that could be cited, compare *Auto. Techs. Intern., Inc. v. BMW of N. Am., Inc.*, 501 F.3d 1274, 1282 (Fed. Cir. 2007) (“[T]he district court was correct that the specification did not enable the full scope of the invention because it did not enable electronic side impact sensors.”) (invalidating claim in patent for side door airbag sensors which covered sensors with a movable mass, i.e., mechanical sensors which sense an impact due to changes in a magnetic field, i.e., electronically because the patent specification adequately taught only the use of mechanical sensors) with *Hologic, Inc. v. Smith & Nephew, Inc.*, 884 F.3d 1357 (Fed. Cir. 2018) (finding disclosure of a single type of lightbulb adequate to support a claim to the use of any type of light guide in a surgical instrument).

23. See generally Gideon Parchomovsky & R. Polk Wagner, *Patent Portfolios*, 154 U. PA. L. REV. 1, 31–32 (2005) (outlining a theory of patent value in which the worth of a patent portfolio is greater than the sum of its individual parts). This Article describes two chief advantages of portfolios: (1) “scale” and (2) diversity:

[A] well-conceived patent portfolio is in many ways a form of “super-patent,” sharing many of the marketplace advantages conventionally attributed to individual patents (paradigmatically, rights to exclude others from the marketplace), only on a larger, broader scale. By aggregating the individualized value of a number of closely related patents, the scale-features of patent portfolios enable holders to realize true patent-like power in the modern marketplace to a degree which is impossible using individual patents alone.

[At the same time,] the inherent diversity created by the aggregation of many different patents offers holders a range of benefits—such as the ability to address the risk and uncertainty fundamental to innovation—that cannot be easily achieved absent the creation of such structures.

*Id.*; see also Michael Risch, *Patent Portfolios as Securities*, 63 DUKE L.J. 89, 140 n.250 (2013) (quoting Parchomovsky and Wagner).

These contingent ownership claims over uncertain future technologies and market products represent a uniquely indeterminate set of assets. This makes exclusionary or residual rights uniquely valuable as a form of entitlement over them. If property as a concept did not exist before, the desire to transfer rights over future technological embodiments and R&D trajectories would have made it necessary to invent it. The fit between the core feature of property—residual rights over unspecifiable uses—and the nature of a patent is exceedingly tight.<sup>24</sup>

B. PATENT PORTFOLIOS AS BUNDLES OF ASSETS: RELATIONSHIP TO CORPORATE LAW THEORY

There are several ways to think about patent property. One is that patents represent investments in “unsticking” information assets from other related assets.<sup>25</sup> Another is that patents represent an internal form of asset partitioning. The literature on corporate law theory has given us a rich account of how the corporate form permits discrete assets to be cleaved off and moved into a distinct entity separate from the personal assets of the people behind the corporation. This is efficient; it allows company founders to put boundaries around a limited “stake” they are willing to place inside the corporation, without endangering their individual assets.<sup>26</sup> This is an obvious corollary to a fundamental feature of corporations—limited liability of shareholders. The asset partitioning idea examines the asset side of the corporate risk equation. By drawing a conceptual circle around corporate assets, the corporate form permits a discrete set of assets to be placed at risk without endangering others.

This idea provides a template for how to think about patent portfolios, which allow a form of asset partitioning that promotes market efficiency rather than limiting liability. Patent portfolios allow a firm to place a distinct

---

24. See Smith, *supra* note 15, at 1702–1704 (asserting that property law uses the “modular theory,” whereby the law protects a variety of rights without knowing which ones the owner will use, because it “is more explanatory than the bundle picture. It helps explain the structures we do *not* find, shows how property can be used to maximize option value, and demonstrates why innovation in property takes the institutional paths it does.”).

25. See Eric von Hippel, “Sticky Information” and the Locus of Problem Solving: Implications for Innovation, 40 MGT. SCI. 429, 436–37 (1994) (describing economic conditions that encourage investments in “unsticking” information).

26. Known as “asset partitioning.” See Henry Hansmann & Reinier Kraakman, *The Essential Role of Organizational Law*, 110 YALE L.J. 387, 390 (2000) (defining asset partitioning, which the authors say is the central defining characteristic of the corporation as an organizational form); see also Giacomo Rojas Elgueta, *Divergences and Convergences of Common Law and Civil Law Traditions on Asset Partitioning: A Functional Analysis*, 12 U. PA. J. BUS. L. 517, 554 (2010) (discussing elaborations and refinements of the asset partitioning concept).

yet related set of assets into a sellable bundle, an idea pioneered in the context of general corporate assets and contracts by Ken Ayotte. Bundling in this form has numerous advantages that apply to R&D and patents.<sup>27</sup> Most notably, it encourages investment in complementary assets (e.g., related patents) and prevents opportunistic holdup. Patent law requires bundling in some cases to explicitly prevent holdup.<sup>28</sup> There is also a general sense that parties to a patent transfer agreement have a duty to prevent holdup.<sup>29</sup>

---

27. See, e.g., Kenneth Ayotte & Henry Hansmann, *Legal Entities as Transferable Bundles of Contracts*, 111 MICH. L. REV. 715, 744 (2013) (arguing that holdup is prevented by including all potentially overlapping contracts and assets in the bundle or portfolio that is sold).

28. Holdup could occur if the seller of a patent withheld one or more related, overlapping patents, so that when the buyer began making and selling a product based on the acquired patent, the seller could sue for infringement under the patent(s) that were withheld. Patent law includes a rule that formally overlapping patents (those technically subject to what is known as “double patenting”) (1) must expire at the same time (through use of what is known as a “terminal disclaimer” of any term in a second patent that would otherwise extend beyond the term of the first patent); and (2) must be transferred together, as a bundle, to prevent lawsuits from multiple sources against use of a single invention. See *In re Van Ornum*, 686 F.2d 937, 948 (C.C.P.A. 1982) (“When a terminal disclaimer causes two patents to expire together[,] a situation is created which is tantamount for all practical purposes to having all the claims in one patent. Obviously, that thought contemplates common ownership of the two patents, which remains common throughout the life of the patents.”); *In re Hubbell*, 709 F.3d 1140, 1145 (Fed. Cir. 2013) (“The second rationale [for double patenting] is to prevent multiple infringement suits by different assignees asserting essentially the same patented invention.”). A terminal disclaimer must “[i]nclude a provision that any patent granted on that application . . . shall be enforceable only for and during such period that said patent is commonly owned with the application or patent which formed the basis for the . . . double patenting [issue].” 37 C.F.R. § 1.321(c)(3). Second, parties can include a “non-holdup” provision in a patent transfer or purchase agreement.

29. See *Abraxis Bioscience, Inc. v. Navinta L.L.C.*, 625 F.3d 1359 (Fed. Cir. 2010), *cert. denied*, 132 S. Ct. 115 (2011). This case involved a \$350 million asset purchase by Abraxis, including eight pharmaceutical patents. Seller company AstraZeneca agreed that it would “do, execute, acknowledge and deliver, or will cause to be done, executed, acknowledged and delivered, any and all further acts, conveyances, transfers, assignments, and assurances as necessary to grant, sell, convey, assign, transfer, set over to or vest in Buyer any of the ‘Transferred Intellectual Property’” described in the asset purchase agreement. *Id.* at 1369. It was subsequently discovered that a subsidiary company of the seller had failed to transfer ownership of relevant patents to the seller prior to the deal; this was remedied, and the seller then transferred the patents to buyer Abraxis. Unfortunately, the transfer occurred too late to confer standing on the buyer Abraxis, so Abraxis’s patent infringement action against another company, defendant Navinta, was dismissed. *Id.* at 1365, 1368. On this, see Xuan-Thao Nguyen, *In the Name of Patent Stewardship: The Federal Circuit’s Overreach into Commercial Law*, 67 FLA. L. REV. 127, 137–46 (2015). Apart from the standing issue, the background to the case shows the general duty to transfer all technology or project-related patents, and therefore prevent patent-related holdup.

Additionally, parties can contractually agree to an anti-holdup provision.<sup>30</sup> In the same spirit, patent law encourages asset bundling—the clustering of assets around a single, discrete R&D project. The bundle, (i.e., the portfolio), can in turn be cleaved off from the other assets of the firm and sold separately. Instead of reducing the risk of liability, it enhances the ability of a firm to monetize an R&D project in the form of a discrete transaction. The remainder of the firm’s assets stays put, and the firm proceeds as before.

Because of asset bundling, patents represent a distinct set of property rights that exist inside the boundaries of a firm’s otherwise undifferentiated assets. Those property rights are separated from the firm’s other assets by recognizable legal boundaries. The legal form of the patent represents a standardized bundle of rights over assets, which consequently segregates these assets from the other unsegregated assets owned by a firm. Patents as project portfolios are therefore characterized by four attributes: (1) compartmentalization, (2) segregation or partitioning, (3) separability or “unstickability,” and therefore (4) the potential for market fluidity.

Critically, a firm need not be active as a seller in the secondary patent market to benefit from that market. A firm’s overall patent portfolio essentially creates a series of easy-to-exercise options. Each project portfolio (i.e., set of related patents) that goes into the overall portfolio which can be sold off if necessary, giving the firm added flexibility.<sup>31</sup> Just the possibility of project portfolio sales makes the firm nimbler and therefore more profitable from an option theory perspective. Markets should theoretically recognize this, but the current understanding of patent portfolios may not have developed enough to exert much influence on existing market valuations.

From the perspective of an external investor, project portfolios allow investments in a set of property rights that represents discrete and “compartmentalized” corporate assets. For example, without patents it would be expensive and difficult for an outside investor to gain ownership

---

30. *See, e.g.*, Intel Corp., Asset Purchase Agreement (Jan. 26, 2012) § 2.9(f) (hereinafter *Codec Intellectual Property Rights*) (“None of the Patents or Patent Rights retained by Seller after the Closing read on, relate to, or are otherwise infringed by the development or use of the Codec Assets (excluding the Codec Personal Property) in the manner in which Seller and its Subsidiaries have been developing such Codec Assets prior to the Closing and as reasonably anticipated in order to commercialize the Codec Assets.”).

31. *Cf.* Parchomovsky & Wagner, *supra* note 23, at 33 (“The broader scope of protection ensures that a wider range of technological possibilities will be covered, which both increases the possibility that the end result of the research and development effort will be covered, and diminishes the concerns of infringement of others’ patents. This “freedom of movement”—the ability to invent, implement, produce, and ship products with in-house resources—is increasingly viewed as an advantage in today’s dynamic market environments, where speed and flexibility are economic imperatives.”) (footnotes omitted).

over each asset standing alone. The entire firm would have to be purchased, and then the particular assets of interest would have to be separated out and split off from the residual assets of the corporation. The particular assets of interest would have to be placed into some separate ownership structure, while the remaining firm assets would presumably remain in the old firm. That old firm would then be sold off to another buyer, shut down, or the like. This would all be difficult and expensive. The hidden value of the secondary market for patents is that it permits this sort of asset divestment to take place in a much more efficient manner. Patent portfolios are comprised of identifiable, discrete assets that can be easily plucked out of the general corporate structure and sold in well-recognized markets. The patents are themselves well-defined assets; when placed in a portfolio, they represent legally distinct asset bundles that are conceptually separable from the other undifferentiated assets of the firm.

Project portfolios make the firm's boundaries more porous or permeable to outside investors. They increase liquidity for discrete assets without requiring messy and disruptive penetration of firm boundaries by outsiders. Assets from the guts of the firm can be surgically plucked out without cordoning them off and extracting them through messy and complex operations. Internal assets central to the firm can be passed outside the firm's membrane in a clean and painless operation.

Thus, secondary markets for patents play an important role in firm flexibility and liquidity. This in turn enables quicker abandonment of failed innovation strategies and a quicker pivot to other, more fruitful projects. For outside investors, it represents a way to get hold of specific firm assets without penetrating and breaking up the firm; the "going concern" value of the overall firm is preserved while particular assets are extracted and sold off.

One solution to the "winner take all" dynamic of the Big Platform era is to encourage acquisition of technology and patents in a form other than full-firm acquisitions. Understanding this alternative thoroughly requires describing the various forms that these markets can take.

### C. A TYPOLOGY OF TECHNOLOGY MARKETS, AND THE ROLE OF IP RIGHTS

Moving attention away from full-firm acquisitions enables discussions about the various ways technology changes hands in arm's-length transactions. The simplest way in which technology changes hands is when it is embodied in a product: a buyer of a DVD containing accounting software or a computer printer buys embedded technology along with the physical

product. This is as true of corporate purchasers, such as Big Platform companies, as it is for consumers.

Another way technology is purchased, however, is in a more disembodied form. The purest version of this type of transaction is a technology license, an agreement by an innovator to permit a licensee to use the innovative technology. In a pure license, there is no physical product involved. The technology itself might be said to be the “product,” the object of the transaction.

Intellectual property (IP) obviously plays a role in many of these transactions. IP rights of various sorts will usually cover one or more aspects of an innovative technology. So the purchase of a DVD or a computer printer may be characterized by the seller as a kind of dual transaction; the buyer receives both the physical product and any IP rights that cover features of the product. Here, the exact interplay of the personal property concepts governing ownership of physical objects and the IP concepts governing the protected features is irrelevant; what matters is that there is an IP component to this standard purchase and sale transaction.

The IP component is much more apparent in the pure technology license than in the sale of an embodiment. Technology and IP rights, in particular patent rights, are often conflated in such a transaction. An innovative software compression algorithm or superior map-rendering software technique may well be covered by one or more patents. The transfer of this innovative technology will therefore often be effectuated via a patent license agreement.

However, for the agreement to qualify as a true technology transaction, the buyer must gain access to a new technique or family of algorithms. The buyer must acquire a capability that is attributable to the creator of the innovation, the owner of the patent.<sup>32</sup> This may involve a transfer of software code, algorithm flowcharts, and programming techniques, among other concepts. Whichever form it takes, the agreement must reflect the transfer of a new capability.

---

32. This is phrased carefully to capture the case where engineers working for the buyer already know and use the patented technology, because they learned about it through various channels well before the buyer acquires rights to it in a formal transaction. Sometimes, in other words, the information has diffused around a field or industry well ahead of the time when a formal transfer agreement is reached. The formal agreement, in such a case, might be said to simply memorialize the information transfer, which occurred informally at an earlier time. See generally Robert P. Merges, *A Few Kind Words for Absolute Infringement Liability in Patent Law*, 31 BERKELEY TECH. L.J. 1 (2016).



Therefore, patent markets are different from product markets because patents do not map cleanly onto product markets. Patents typically cover technological components: small pieces of larger technologies. Examples include a part of a mobile phone antenna, a technique for compressing data to be sent over a network, or a method for encoding location information on a CD, an example we will return to later.

Patents map onto technologies. The invention in an antenna patent may form part of a mobile phone antenna. The compression algorithm may be used in a software program to transmit digital content such as music, video, or text. The popup menu may be part of a software program that handles calendaring or interfaces with travel-related websites.

Technologies, in turn, map onto products. The antenna is part of a mobile phone. The compression algorithm is part of a data streaming program used by music streaming companies or video websites. The popup menu may be part of a travel website or a suite of software for a mobile or desktop device.

Finally, products map onto markets. The mobile phone containing the antenna is sold in competition with other mobile devices, including phones, tablets, and watches. The data-streaming program is incorporated into the software of one of several music-streaming companies, or it is used by one video streaming service (e.g., Netflix) that competes with others (e.g., Amazon Prime or YouTube). The popup menu may be part of a desktop operating system such as Microsoft Windows, which competes with free operating systems such as Android for mobile; alternatively, it may be incorporated into one travel website (e.g., Kayak) that competes with others (e.g., Expedia).

This complex, multi-step “mapping” can be summarized as follows:

Patents → Technologies → Products → Product Markets

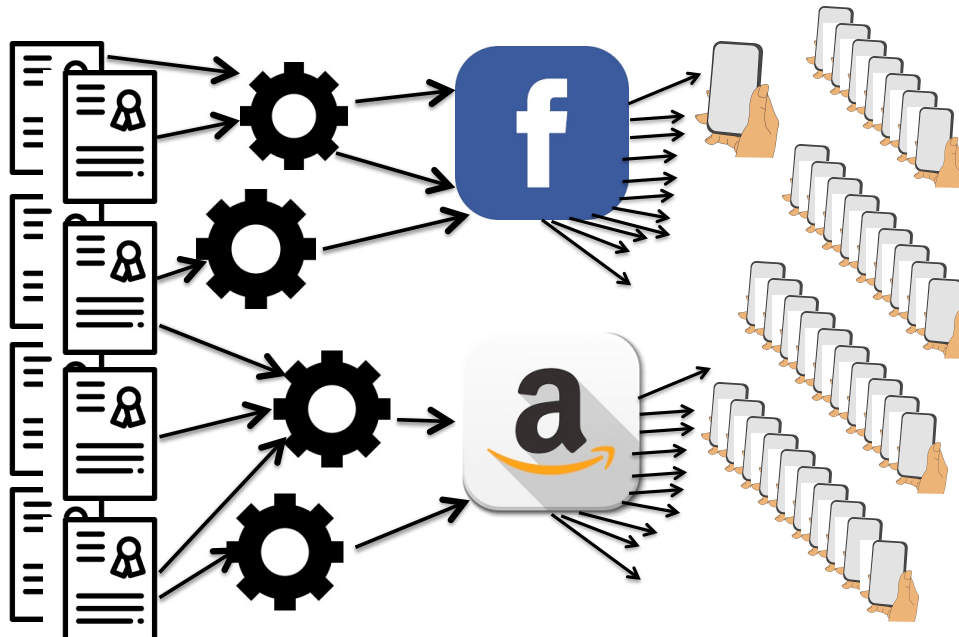
In the context of a winner-take-all/network goods market, this demonstrates why a “failed product” company does not equate to a company which has made no contribution. A helpful new technology may not be sold in a distinct market. It may be useful only as one small piece of an overall platform technology. The fact that an innovative small company has not succeeded in building a market for its technology may not be due to a poor technology design. It may instead be due to the reality that there is only one or a few prospective buyers for its design. If those buyers duplicate the small company’s technology (intentionally or not) instead of buying it, the small

company will fail—not because its technology was a failure, but because of market imperfections in platform industries.

The overall structure of the industry can be represented this way:

Figure 1: Mapping Patents into Product Markets

**Patents → Technologies → Products → Product Markets**



In this diagram, the technologies (represented by little gears) on the left are covered by patents—sometimes by more than one of them. This illustrates that patents are not the same as technologies. The technologies, rather than the patents, are what make up the inputs into Big Platform products or services, such as the Facebook platform or the Amazon marketplace.

The acquisition of a new capability attributable to the innovator distinguishes technology transactions from transactions concerned solely with legal liability. In a purely legal transaction, the only new asset acquired by the buyer is the legal right embodied by one or more patent rights. The buyer in these cases does not learn about any new technology or acquire any new technical capabilities. It instead buys patents to protect itself from future patent infringement lawsuits, or possibly to sue competitors in patent infringement suits of its own. The transaction neither effectuates nor

memorializes the transfer of any innovation or new capability; it is a transfer of legal rights and nothing more.

There are some disputes among patent specialists about the relative volume of the two transaction types.<sup>33</sup> Some findings seem to show that much patent litigation has little to do with capability enhancement; the classic study shows that accused patent infringers are almost never proven to have copied any technology from the patent owner. The study concludes that because defendants in infringement cases are independent inventors, patents in those cases simply represent a tax on innovation rather than new capabilities. In a more recent study, however, Professor Colleen Chien disagrees.<sup>34</sup> She shows that in the field of software technology, many of the license agreements she studied involve the actual transfer of computer code, know-how, and associated technical information.<sup>35</sup> They were more than just settlement agreements fending off legal liability; they were transfers of new capabilities and technologies as well.

### 1. *The Role of Patents in the Spinoff of “Orphan” Technologies*

The argument so far is simple. Discrete technologies can be transferred to Big Platform companies via the market for technology; this preserves the autonomy and culture of an innovative firm while moving innovations into the hands of Big Platform companies. Anyone with experience in sophisticated corporate deals would just call this a spinoff—the transfer of some portion or unit of one firm to another, separate firm. Regardless of terminology, both the special nature of technology-intensive spinoffs and the role that patents play in enabling them merit particular attention.

Importantly, “transfer of patents” here refers to transfer of a patent portfolio, a set of related patents clustering around a discrete technology. There is a market for individual patents, which are often purchased to provide defensive protection for the buyer. These patents cover a technology or component that might be the subject of a patent infringement lawsuit brought by another patent owner. Owning a patent that covers a component

---

33. See, e.g., Robin Feldman & Mark A. Lemley, *Do Patent Licensing Demands Mean Innovation?*, 101 IOWA L. REV. 137 (2015) (determining that very little technology transfer accompanies most patent lawsuit settlement/licensing deals).

34. Colleen V. Chien, *Software Patents as A Currency, Not Tax, on Innovation*, 31 BERKELEY TECH. L.J. 1669, 1669 (2016).

35. *Id.* (“[T]he majority of material software licenses reported by public companies to the SEC from 2000–2015 (N=245) support true technology transfer.”) (basing this statement on a study of the terms of these reported licensing agreements).

gives a potential infringement defendant “ammunition” to use against another patent owner/plaintiff in such a suit.<sup>36</sup>

However, the market for individual patents is beyond the scope of this Article. This discussion focuses on the transfer of a discrete technology, akin to the product of a distinct R&D project. The typical corporate R&D project results not in a single patent, but in a group of related patents—a portfolio. These patents represent core aspects of the technology, various improvements, refinements, and modifications of it, and all international corresponding patents that grow out of initial domestic patent filings related to the project. It also often includes pending applications, as explained earlier.

When collected in project portfolios, patents represent an interesting asset class that is distinct from general equity in a firm. They represent a form of internal asset partitioning that creates important efficiencies. Project portfolios make it easy to sell off the products of distinct R&D projects. This increases firm-level flexibility by making it easy to sell off the products of lines of research that have not panned out, which in turn enhances firm liquidity. The secondary market for patent portfolios allows firms to sell off assets associated with (1) abandoned, (2) re-directed, or (3) multi-application research projects in a relatively efficient way. Each of these transactions has some unique features that are worth taking a moment to describe.

#### a) Abandoned Projects

Abandoned projects are perhaps the most common source of patent portfolios. Companies of all sizes are constantly opening new lines of research. Except for the most truncated R&D projects, each of these lines will typically lead to at least a handful of patents. But the nature of research is subject to a number of well-known vicissitudes. Markets shift, often due to

---

36. As is well known, defensive acquisitions do no good against a pure patent troll or Patent Assertion Entity (PAE): these patent owners do not themselves make or sell any products, they are simply patent holding companies. This means that a defendant cannot assert its patents against a troll or PAE, because these entities are incapable of infringing any patents. *See* 35 U.S.C. § 271(a) (defining infringement as making, selling, using, importing, etc., embodiments of a claimed invention). As with all aspects of patent markets and patent litigation, however, there are some delicate gradations between pure trolls and pure traditional “producing” companies. Sometimes for example a producing company will supply patents to a separate firm for the sole purpose of suing and harassing a rival of the producing firm. This kind of “privateering” arrangement could incite the rival firm into a strategic response: filing an infringement lawsuit against the producing firm that supplied patents to and sponsored the privateer. The point is that defensive patents might be useful in the overall strategic game between rival producing firms, even if they are not directly useful as counter-ammunition in a specific suit brought by a (privateering) troll or PAE.

consumer preferences; technology changes, often in unforeseen ways; senior management changes its mind about the importance of some product or line of business; new units are acquired; or company politics assert their influence. In each case, what had been a priority even in the recent past may be rapidly de-emphasized. As ideas come in and out of favor, research projects follow. When an R&D project is abandoned, the secondary market for patents may permit the firm to recoup some of the R&D investment it would otherwise lose entirely. Other companies may not have given up on the technology, or they may try to use it in existing products in ways not available to the firm that developed it. Regardless, there may well be buyers for technologies that an originating firm has given up on. If so, the secondary market then allows for easier exit from abandoned research lines and therefore permits quicker transitions to new, more promising lines of research.

b) Re-Directed Research

Some companies also re-direct an R&D project from one goal to another. This may render some patents in the project portfolio less essential. For example, a project to write software code designed to signal a car driver about impending danger might be re-directed when the company decides it wants to make a fully autonomous (self-driving) vehicle. Research on how best to signal and assist a driver will therefore no longer be useful to the company, but other companies may have an interest in it. If an automaker wants to improve its danger signaling, it might purchase the first company's patents that cover this function. Alternatively, if the automaker already has a well-functioning driver signaling system, it might still purchase the patents for "defensive" use to ward off future patent infringement suits from third parties. In either case, some of the project portfolios may be sold in a patent transaction that benefits both the R&D company and the automaker.

c) Multi-Application Research

It frequently happens that an R&D project aimed at solving one problem yields technology that serves that goal but is also useful for other applications. For example, years ago the DuPont Company set out to create a permanent "nonstick" coating that could be used to make various surfaces less likely to accumulate detritus. Thus was born Teflon, whose first application was as a nonstick coating on cooking pans. A DuPont researcher familiar with Teflon (polytetrafluoroethylene, or PTFE) quickly saw that its unique features had a wide array of potential applications. This researcher, W.L. Gore, founded his own company without objection from DuPont and created the GoreTex material, hikers' and backpackers' friend. Because

DuPont determined that it did not have any continuing interest in PTFE, Gore was able to spin off a separate company.<sup>37</sup>

This scenario has been repeated many times since. One example involves Magnolia Software, a small startup in the mobile phone software field. It was founded in 2000 by an Israeli entrepreneur named Haim Harel, who had founded a number of other startups earlier in his career.<sup>38</sup> Magnolia invested somewhere near \$60 million over the next ten years to develop what it called Mobile Transmit Diversity (MTD) technology, which makes more efficient use of mobile bandwidth on the “uplink” side of mobile communications (when data is sent “upward” from a mobile phone or other device to the local cell tower or other hub, and hence out onto the mobile network).<sup>39</sup> Though Magnolia continues to sell both hardware and software versions of its MTD technology, it sold more than fifty of its MTD-related patents to Google in June 2012 for an undisclosed amount.<sup>40</sup> Based on what we know, Magnolia is using the proceeds from its patent sale to fund ongoing operations; this presumably includes continuing R&D. This case study lends credence to the main point that the market for technology can help preserve a going-concern R&D firm. This market provides a payday for past R&D while freeing up the company to continue innovating in the future.

---

37. The pertinent history is recounted in *W.L. Gore & Assocs. v. Carlisle Corp.*, 381 F. Supp. 680, 685 (D. Del. 1974).

38. *Team*, MAGNOLIA BROADBAND (last visited Dec. 30, 2019), [http://www.magnoliabroadband.com/index.php?option=com\\_content&view=article&id=55&Itemid=49](http://www.magnoliabroadband.com/index.php?option=com_content&view=article&id=55&Itemid=49) [https://perma.cc/CZ5Y-AYYW]. For an example of Harel’s research, see Sherwin Wang & Haim Harel, *Increase of Reverse Link Capacity of the 3G CDMA Network by Mobile Transmit Diversity*, 2007 IEEE RADIO & WIRELESS SYMP. (Apr. 23, 2007), <https://ieeexplore.ieee.org/document/4160729> [https://perma.cc/CFQ5-XHLB].

39. Mark Hearn, *Google’s Patent Buyout From Magnolia Broadband Now Official*, TECHNOBUFFALO (June 18, 2012), <https://www.technobuffalo.com/googles-patent-buyout-from-magnolia-broadband-now-official> [https://perma.cc/BB3T-TFKZ].

40. According to the Magnolia CEO, “[t]his transaction is a milestone for Magnolia Broadband. It provides a return to our investors and funding for continued development of Magnolia’s MTD technology.” *Id.* And, according to a trade press report: “[i]nterestingly, although the [Magnolia] MTD patent portfolio was acquired by Google, Hautanen [the CEO] noted that ‘The software, which can be embedded into any mobile broadband device remains the property of Magnolia Broadband and will be made available to mobile device vendors and chipset companies.’” Rik Myslewski, *Intel, Google Ink Patent Deals with InterDigital, Magnolia Broadband*, REGISTER (June 18, 2012), [https://www.theregister.co.uk/2012/06/18/intel\\_google\\_patent\\_deals/](https://www.theregister.co.uk/2012/06/18/intel_google_patent_deals/) [https://perma.cc/9QN5-6WKP].

## 2. *Failed Product Companies and the Market for Patents*

Disputes over the social value of the secondary market are often tied up with differences of opinion over the volume and value of patent litigation.<sup>41</sup> The tip of the spear in these disputes takes the form of arguments over patents that come from failed-product companies. These are companies that started life with the best intentions; their founders hoped they were creating the next Google, Microsoft, or Intel. As often happens with small companies, however, things did not work out as planned. Whether the intrepid startup never made a saleable product or was beaten soundly in the marketplace, the end result is the same; in these cases, dreams of greatness died a certain death. When the battle is over and defeat is at hand, what is left is often just a few loyal employees, some scattered assets, and often a great deal of debt. Among the scattered assets left at the end are the firm's patents, often thought to have the most potential value. Sometimes this leads the failed product company to undergo a metamorphosis; it turns into a patent-holding company, hoping to license its patents and litigate if necessary in the process. Other times, the failed company sells its patents to another firm. Perhaps it sells to an operating company looking for patents to bulk up its portfolio. Perhaps it sells instead to a patent aggregator such as RPX or Intellectual Ventures. Or perhaps it sells to a PAE or an entity that looks like a classic patent troll.<sup>42</sup>

Viewpoints on how we should feel about these companies vary but generally form a spectrum.<sup>43</sup> On one end are operating companies who complain that the name says it all—these are failed companies. They did not

---

41. See Michael J. Burstein, *Patent Markets: A Framework for Evaluation*, 47 ARIZ. ST. L.J. 507, 507–08 (2015) (“Taking seriously the analogy between patent markets and financial markets, I demonstrate that there are numerous circumstances in which even well-functioning patent markets will fail to promote innovation . . .”).

42. See Colleen Chien, *Startups and Patent Trolls*, 17 STAN. TECH. L. REV. 461, 479–81 (2014) (“Some small companies have been able to sell or monetize their patent portfolios to support ongoing or new practicing business ventures. . . . A successful patent assertion campaign can support the business, or help fund a transition, for example, to another operating company business model or full-time patent assertion.”).

43. Compare Xuan-Thao Nguyen, *Zombie Patents and Zombie Companies with Patents*, 69 FLA. L. REV. 1147, 1155–56 (2017) (criticizing failed product companies and advising the Federal Circuit to disfavor them in patent cases) with Michael Risch, *Licensing Acquired Patents*, 21 GEO. MASON L. REV. 979, 988 n.29 (2014) (arguing that sale of patent after firm fails can encourage firm founders to try another startup, and become “serial entrepreneurs”). Good or bad, failed product companies are a definite presence among companies that assert patents. See Robin Feldman et al., *The ALA 500 Expanded: The Effects of Patent Monetization Entities*, 17 UCLA J.L. & TECH. 1, 40 (2013) (“Many of the individuals in the samples appeared to be inventors who had tried to operate companies and when they failed, switched to litigation as a way of monetizing their patents.”).

deliver real innovations that society wants and needs. Allowing them to extract money from the winners after the fact of their loss does no one any good. This is especially so, the argument goes, because these companies in general sue successful product companies for infringing patented inventions which the successful companies themselves invented on their own. Failed companies take advantage of the rule in the patent law that independent invention is no defense to infringement. The firms and people that hold the patents of failed product companies engage in lawsuits designed to extract rents from the companies that succeeded on their own and transfer payments to holders of patents as the last, sad harvest of failure. This is, as economists say, simple rent-seeking—taking wealth from one who earned it and giving it to another whose business is to seek out and partake in well-deserved pockets of wealth without helping to create it or build it up.

On the other end of the spectrum are failed-product companies who feel wronged in one way or another. They may feel that their ideas were in fact borrowed or that they helped make possible some aspects of the technology that is now dominated by successful product firms. At the extreme they may feel that one or more big, successful companies stole their ideas outright. They may also feel that their ideas were in some ways superior to those championed by the now-successful firms. They lost out not due to inferiority, but due to random developments or “path dependencies” early in the history of the industry; those developments ended up rewarding the successful firms for essentially unimportant or random reasons. Viewing things this way, a failed product firm may feel that its contribution is no less meritorious than that of a successful company. The failed company should therefore be paid for the unacknowledged contribution it made to the early development of the industry it worked so hard to create. Failure in the product market, in this view, does not mean total failure and ought not to preclude these firms from getting some compensation for their valuable early contributions.

One team of researchers summarized the issue this way:

Failed startups . . . have little ongoing business. They may feel that the alleged infringer unfairly beat them in the marketplace. The alleged infringer may have the opposite view of the marketplace battle, and these underlying divergent views may affect the patent case. This divergence in views between failed startup plaintiffs and defendants may make disputes more difficult to settle, resulting in longer disputes. Failed startups also have investors who may desire some return, via the patent lawsuit, on their otherwise lost capital.<sup>44</sup>

---

44. Christopher A. Cotropia et al., *Heterogeneity Among Patent Plaintiffs: An Empirical Analysis of Patent Case Progression, Settlement, and Adjudication*, 15 J. EMP. LEG. STUD. 80, 89



The best study of these companies primarily includes companies that continue to manufacture some products while licensing patents covering products these companies once made but no longer do:

Examples of formerly manufacturing entities include IBM, MOSAID (now Conversant), and General Electric. General Electric continues to make products, but also engages in extensive licensing of its large patent portfolio, including many patents covering technology that it does not manufacture. It is unsurprising, given the lack of precision in the rhetoric, that these companies have been attacked as “patent trolls,” despite their past or ongoing commitment to manufacturing.<sup>45</sup>

a) Failed-Product Companies and Patent Litigation: Ex-Post Market Making

Failed-product companies that would rather not sell their patents to third parties can use another strategy; they can license instead. A number of studies on different types of patent plaintiffs finds that there are a few companies that pursue this approach.<sup>46</sup> When it happens, the usual battle of competing narratives is joined—the failed company scrapes the bottom of the barrel by becoming a troll, while the proud pioneer just wants recognition

---

(2018). On patent sales as a way to earn back some money for investors, compare Michael Risch, *The Layered Patent System*, 101 IOWA L. REV. 1535, 1575–76 (2016) (“Venture capitalization, or lack thereof, is a potential source of concern for the failed startups [studied]. Not one of the failed startups [which were studied, and which litigated one or more patents] . . . had venture funding. The reasons for this are unknown. The failed startups could have failed precisely because they had no financing, and venture-backed firms were savvy enough to sell their patents and remain in operation.”) (footnotes omitted).

45. Kristen Osenga, *Formerly Manufacturing Entities: Piercing the “Patent Troll” Rhetoric*, 47 CONN. L. REV. 435, 440 (2014) (footnotes omitted); see also David L. Schwartz, *On Mass Patent Aggregators*, 114 COLUM. L. REV. SIDEBAR 51, 52 (2014) (“While there are patent holders who abuse and exploit the patent litigation system, there also are patent holders with meritorious claims who have been unfairly denied compensation. This is true for companies that both do and do not manufacture. The critics also lump together a wide variety of seemingly different actors, including individual inventors, failed startups, research and development companies, mass patent aggregators, and Wall Street speculators who buy a single patent for purposes of enforcement. The correct analysis of the costs and benefits of patent trolls is quite complicated, and far beyond the simple narrative based upon whether the owner of the patent manufactures products.”).

46. See Cotropia et al., *supra* note 44, at 94 (categorizing patent lawsuit plaintiffs) (“Failed Operating or Start-up Company: A company that originally invented the patent-in-suit and attempted to commercialize the technology. At present, the company sells no products and its primary business appears to be patent litigation. An example of the Failed Operating or Start-up Company is Broadband Graphics LLC.”). Cotropia’s data showed that failed companies brought 4% of such litigation in 2012. See Christopher A. Cotropia et al., *Unpacking Patent Assertion Entities (PAEs)*, 99 MINN. L. REV. 649, 692 (2014).

of its path-breaking innovations that paved the way for successors in the marketplace. Litigation of this type tests some of the points made in this Article, particularly how patents capture value for early contributors who lose out over time to ultimate winners such as the Big Platform companies. This litigation also affects the secondary market for the failed-product company's patents; the value of "first generation" patents which Company B wants to sell may be affected by the litigation prospects of other "first generation" patents that Company A has chosen to license (and later, litigate) on its own. Litigation prospects essentially affect the value of patents even when they are not destined for immediate litigation.

A good example of this scenario is the patent enforcement campaign waged by the creators of the Blackberry handheld device that hit the market in 1999. Among its other features, Blackberry introduced a version of "instant text messaging," which helped make its device a big hit in the 2000s. Blackberry sales grew steadily during the decade, reaching a peak of almost \$20 billion in 2011. Only five years later, sales were down to \$2.2 billion and the company had lost money for four straight years.<sup>47</sup> Blackberry went from having 20,000 employees in 2011 to approximately 4,000 in 2018.<sup>48</sup> While Blackberry did introduce a "smart phone" as an outgrowth of its original handheld "digital assistant," the introduction of a new iPhone in 2013 effectively killed Blackberry as a player in the smartphone market.<sup>49</sup>

Beginning around 2015, Blackberry seems to have transitioned to selling corporate-level security software. It puts its still-valuable brand on low-cost mobile phones sold by others, but it is no longer a major player in the high-end smartphone market that it contributed heavily to the creation of. This U.S. market, of course, belonged almost exclusively to Apple and Samsung/Android in 2018. These two companies have undoubtedly emerged as the winning platforms thus far in the smartphone market.

Like many pioneers who later lose out in the product market, Blackberry turned to licensing its patents to the product market winners. The specific

---

47. At the end of 2007, the company had a market capitalization of more than \$60 billion. This had fallen to \$4 billion by August 2016. See DEBORAH HIMSEL & ANDREW C. INKPEN, *THE RISE AND FALL OF BLACKBERRY* (Harvard Business Publisher 2017).

48. David Friend, *BlackBerry cuts jobs, shifts employees as part of turnaround plan*, STAR (July 21, 2015), [https://www.thestar.com/business/tech\\_news/2015/07/21/blackberry-cuts-jobs-shifts-employees-as-part-of-turnaround-plan.html](https://www.thestar.com/business/tech_news/2015/07/21/blackberry-cuts-jobs-shifts-employees-as-part-of-turnaround-plan.html) [<https://perma.cc/3P9M-ZKXY>]; Arne Holst, *BlackBerry's number of employees from 2017 to 2019*, STATISTA (May 15, 2019), <https://www.statista.com/statistics/995125/blackberry-number-of-employees/> [<https://perma.cc/VQK7-VAJW>].

49. See John McDuling, *Investors are starting to think Blackberry has a future*, QUARTZ (June 30, 2014), <https://qz.com/228123/investors-are-starting-to-think-blackberry-has-a-future/> [<https://perma.cc/35U5-X7RV>].

technology Blackberry claimed to have originated is instant messaging, or text messaging. Blackberry devices included a texting feature as early as 2005 through its Blackberry Messenger (BBM) application, which ran on its handheld devices.<sup>50</sup> Blackberry asserted patents on several texting features, including an encryption technique<sup>51</sup> used to keep messages secure.<sup>52</sup>

Another Blackberry patent (U.S. Patent 8,429,236)<sup>53</sup> asserted against Facebook describes an adjustable communication rate between an application running on two interconnected devices. The invention adjusts the communication rate depending on whether users on both devices are actively using the application at the same time. Status updates are exchanged infrequently when the applications are in background mode and not being actively used; this conserves transmission bandwidth and power consumption by the devices. A texting application like WhatsApp or WeChat, for example, will check every so often to see if a new message has been sent. When the application is not being actively used, the time between status updates is long. But when the system detects that two users are using the same application simultaneously—for example, when an active texting session is underway—the transmission of status updates accelerates. Each mobile phone “prioritizes” the texting application in terms of transmission bandwidth and power consumption. The two phones return to background mode when the texting session is over, which means less frequent updates and less power consumption.

If the '236 patent is adjudged to be valid and a solid incremental advance in the messaging field, Blackberry has a reasonable claim to compensation. Although this small feature of messaging software is one of many features that collectively make up the user experience of messaging with Facebook and Instagram, it still adds some value to the user experience. It would therefore still be one of the building blocks on which Big Platform

---

50. See Complaint for Patent Infringement at 15, *Blackberry Ltd. v. Facebook, Inc.*, 2018 U.S. Dist. LEXIS 221047 (C.D. Cal. Aug. 2, 2018) (No. 2:18-cv-01844).

51. Blackberry (RIM) in 2009 acquired the company (Certicom, Inc.) that actually pioneered this encryption technique. See Motek Moyen, *BlackBerry: Make Certicom Patents Licensing More Affordable*, SEEKING ALPHA (Oct. 11, 2014), <https://seekingalpha.com/article/2554945-blackberry-make-certicom-patents-licensing-more-affordable> [<https://perma.cc/SF9M-L8QN>].

52. See U.S. Patent No. 7,372,961 (issued May 13, 2008). This patent was filed first in Canada (Blackberry's home country) in December 2000. The invention claimed in this patent was originally created by employees of Certicom, Inc., the Canadian company that was acquired by Blackberry (RIM) in 2009.

53. See Transmission of Status Updates Responsive to Recipient Application, U.S. Patent No. 8,429,236 (filed Apr. 8, 2009) (issued Apr. 23, 2013).

companies have built their successful social media systems. Blackberry's devices might be failed products, and Blackberry itself might be considered a failed or diminished company. Nevertheless, some of the Blackberry *technologies* must be considered successes. Given the "winner take all" nature of the platform markets in which instant messaging is now embedded, the only compensation Blackberry will get for its contributions is through a patent licensing program—a program backed by patent litigation, as they so often are.

b) Summary: The "Two Period" Nature of Patents and Patent Litigation

The points made in this Article regarding the good fit between property theory and patents depend largely on the way patent claims capture future options. The essential quality of a property right is residuality; all uses of an asset not carved out by illegality or the like are permitted to the owner—without the need to specify or even know about the long list of these uses. Similarly, patent claims cover a host of unspecified and perhaps unknown variations and applications of a basic inventive concept. Essentially, patent claims can be valuable if and when they cover *future embodiments* of an invention.

Contrast this with patent litigation, where courts often impose a retroactive obligation on the patent infringer to the patent owner. It is retroactive in that it imposes the obligation from the moment an infringer can be proven to have incorporated a validly claimed invention in its product—even when no voluntary deal was struck by the parties and the infringer knew nothing about the patentee's patent at that time.

When claims are issued, they cover many possible future manifestations of the claimed technology. In litigation, these claims are applied retrospectively to the activities of an accused infringer—by looking back from the time of the patent infringement suit to the time when infringement began. The question is whether the patent claims cover what the defendant was doing once they are construed fairly. The future-orientation of the claims is often what permits a finding of infringement, even though that finding is not arrived at until later. Claims by their nature create the possibility of future infringement when they are issued by the patent office, but often this obligation is imposed by a court retroactively—sometime after the infringing behavior began and only after the patent has been litigated.

### III. PATENT MARKETS, MERGERS, AND R&D: WHAT DO THE DATA SAY?

Thus far, this Article argues that the legal system should show some solicitude for the secondary patent market. The crux of this argument is that selling patent portfolios allows companies to both innovate and retain continuity as going concerns. Their continued existence, in turn, has advantages over full-firm acquisitions. This raises the question of what happens to the R&D and innovation capacities of a firm after it has been acquired. If acquired firms are more innovative across the board, this would undermine the comparative benefits of the patent market.

There is a fair amount of consensus, though far from universal agreement, on every aspect of this issue.<sup>54</sup> From the point of view of innovation, big is not always bad and in fact can be pretty good. Most researchers conclude that overall innovation, usually measured by number of patents,<sup>55</sup> improves after a merger or firm acquisition. If overall innovation

---

54. One study summarizes the competing schools of thought from the economic subfield known as industrial organization (IO):

[T]here are different arguments regarding the effect of firm size on . . . R & D productivity. While some studies argued that because, in large firms, R & D costs can be spread over its [larger] output, these firms can realize higher R & D returns, [but] other researchers argue that, due to some of the characteristics of large firms, such as a loss of marginal control or high level of bureaucratic control, R & D performance actually decreases.

Negin Salimi & Jafar Rezaei, *Evaluating Firms' R&D Performance Using the Best-Worst Method*, 66 EVAL. & PROG. PLAN. 147, 148 (2018). For a classic example from the “bigger is better school” based on a simple economic model, see Steven Klepper & Wesley M. Cohen, *A Reprise of Size and R&D*, 106 ECON. J. 925 (1996) (stating that larger firms can spread R&D costs across more divisions and products, so have an advantage in the scale of R&D they can conduct). For an overview of the field, the literature, and the debates, see generally MORTON I. KAMIEN & NANCY L. SCHWARTZ, *MARKET STRUCTURE AND INNOVATION* (1982); FREDERIC MICHAEL SCHERER & DAVID ROSS, *INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE* (1990).

55. In most of the studies we are reviewing, innovation levels pre- and post-merger are measured by using various patent-related variables. Studies employ either the sheer number of patents before and after, or their quality (often determined, as is conventional, by the number of times the patents are cited in other patents and research studies). The simple objection to this measure is that it is usually large companies that make these acquisitions—bigger buys smaller. And bigger companies usually have a more aggressive mandate to build out their patent portfolios. The gains in numbers of patents, then, may show not a truly higher rate of innovation but simply a greater propensity to acquire patents per dollar of R&D spent. As for the citation data, though it can often be helpful, citations are susceptible to a number of well-known limitations. It might well be that in many cases the higher number of citations come from the greater visibility that comes with patents issued to larger companies. It could mean quality, in other words, but it might also simply signal prominence.

were the only concern, the case for a patent market looks shaky. However, a consistent body of research also shows that radical innovation decreases with firm size. A newly acquired firm becomes part of a larger company, and large companies rarely succeed in paradigm-shifting innovations. Before elaborating on this point about radical innovation, it helps to understand why many studies connect increased innovation with post-merger firms and large firms in general.

Two explanations have been given over the years as to why bigger may be better. The first arises from market power and is known as the Schumpeterian Hypothesis, after economist Josef Schumpeter.<sup>56</sup> High profit margins result from the oligopolies or monopolies enjoyed by big companies, and this provides money for increased R&D. The second answer springs instead from the nature of technology; this theory is captured by the term “synergy.” Multiple related researchers working in proximity with each other combine findings and ideas in ways that increase the productivity of the entire collective group. Talented researchers, previously isolated in “silos,” now share ideas with others from related fields; this is a fertile formula for innovation. The whole of the combined research teams ends up being greater than the sum of its individual parts.

Schumpeter’s argument for the benefits of bigness would generally regard mergers as a good thing. Typically, “mergers reduce . . . product market competition and [therefore] increase expected payoffs from employee innovations”<sup>57</sup> due to the increased size and market power of the post-merger firm. From this perspective, the market power that so concerns

---

56. See JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM AND DEMOCRACY (1942); see also FREDERIC M. SCHERER, INNOVATION AND GROWTH: SCHUMPETERIAN PERSPECTIVES 222–37 (1984) (analyzing the Schumpeterian Hypothesis in light of studies that seem to discredit it). The idea that monopoly power leads to innovation is associated with the later writings of Schumpeter such as the 1942 volume just cited. This book includes the famous idea of “the perennial gale of creative destruction,” which describes the “process of industrial mutation that incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one . . .” *Id.* at 82–83. This “later” Schumpeter is often contrasted with an earlier string of writings in which he emphasized small firms and individual entrepreneurs. See RICHARD R. NELSON & SIDNEY WINTER, AN EVOLUTIONARY THEORY OF ECONOMIC CHANGE 39–40 (1982) (citing Joseph Schumpeter’s 1936 book, *The Theory of Economic Development*, as a good expression of “earlier Schumpeter”).

57. Paolo Fulghieri & Merih Sevilir, *Mergers, Spinoffs, and Employee Incentives*, 24 REV. FIN. STUD. 2207, 2233 (2011). But see *id.* at 2233 (noting that the merger does also result in some disincentives to innovation).

antitrust authorities is beneficial because firms with market power are more secure in the pursuit of ambitious and long-term-oriented R&D.<sup>58</sup>

A comprehensive study verified that there are benefits from integrating the R&D efforts of acquiring and acquired firms; the talk of “synergies” as a rationale for mergers has a strong basis in truth.<sup>59</sup> This study had two primary findings. First, firms acquire other firms more often when the “technological overlap” between the two firms is high—when they are familiar with and can effectively evaluate the quality of the acquired firm’s R&D activity. This is an aspect of what is known in technology studies as “absorptive capacity.”<sup>60</sup> Second, acquisitions are dominated by big and successful companies—“larger firms, as well as firms with faster sales growth, better operating performance . . . and higher prior year stock returns.”<sup>61</sup> The logic of “bigger is better” is surely at work here. The larger a firm, the more products and research projects it has.<sup>62</sup> With more projects comes a greater chance for synergies.<sup>63</sup> Large firm size and the accompanying resources to capitalize on

---

58. Note that one study found high R&D productivity in small *and* large firms, but not in the mid-sized firms that stand between them. See Kuen-Hung Tsai & Jiann-Chyuan Wang, *Does R&D performance decline with firm size? A re-examination in terms of elasticity*, 34 RES. POL'Y 966, 973 (2005) (finding a u-shaped relationship between firm size and innovation, based on total factor productivity (TFP) data: small and large firms are highly innovative, but medium sized firms are not).

59. See Jan Bena & Kai Li, *Corporate Innovations and Mergers and Acquisitions*, 69 J. FIN. 1923, 1955 (2014) (studying 1762 mergers, from 1984 to 2006).

60. *Id.* at 1945.

61. *Id.* at 1936.

62. Synergies may add value, but the first finding is troubling. If mergers are more common in cases of a degree of technological overlap, mergers would be more likely to remove potential future R&D competition than product market competition. While this is good in one sense—short-term consumer welfare is enhanced by product market competition—it is worrisome in another: future innovative capacity is likely to be absorbed into larger and larger firms. Whether small companies aim from the outset to be acquired; or whether they simply fail to introduce meaningful product competition; their future innovative potential will be taken inside a large company. And so again the benefits of independence and autonomy will be lost.

63. One aspect of the Bena and Li study presents a contrast with conventional Schumpeterian market power explanations of mergers and so is worth noting. The authors find that “close rivalry in product markets has a negative impact on the likelihood of firms merging. As a result, the positive effect of technological overlap on the likelihood of a merger pair formation is reduced for firm pairs that also overlap in product markets.” Bena & Li, *supra* note 59, at 1949. Note that one study, based on an economic model (and not empirical data), provides support for this empirical finding (even though it contravenes Schumpeterian wisdom. See Paolo Fulghieri & Merih Sevilir, *Mergers, Spinoffs, and Employee Incentives*, 24 REV. OF FIN. STUD. 2207, 2233 (2011). The authors argue that limiting product market competition reduces employee incentives due to lessened opportunities to pursue another job, providing one reason for firms to avoid mergers with other firms that compete in product markets:

synergies might give big companies a natural advantage when it comes to post-acquisition innovation.

The synergy trope shows up as an explanation for why large-company acquisitions have replaced some IPOs as a way for small companies to cash out (or “exit”). One research paper on this topic says:

The recent decline in IPO activity can be explained by the small firms’ increasing preference for being acquired rather than growing independently. . . . [A] firm’s trade-off between being acquired and remaining independent strongly depends on the extent of the synergies arising from a potential merger, which are however difficult to assess ex-ante . . . . [W]e document that [Young Innovative Companies] facing the potential to develop larger synergies are the main [cause] responsible for the decline in IPOs. Compared to 15 years ago, the quarterly number of IPOs conducted by these firms has decreased by 20 [percent].<sup>64</sup>

Despite some counter-indications in the older literature, newer studies support the idea that R&D efficiency may increase after firms merge into a single entity.<sup>65</sup>

---

[B]y reducing the number of firms in the product market, mergers limit employee ability to go from one firm to another with a negative effect on incentives. . . . When the negative effects of the merger on incentives are sufficiently large, firms are better off competing in the product market and competing for employee human capital rather than merging and eliminating competition. In other words, [in the model,] firms prefer not to merge and [instead choose to] bear competition in the product market to maintain stronger employee incentives.

*Id.* An omitted sentence in the block quote states another disadvantage of mergers: “Moreover, mergers create internal competition between the employees of the post-merger firm, with an additional negative effect on incentives to innovate.” *Id.* While consistent with the terms of the model, the idea that internal teams of rivals face reduced incentives to innovate has been countered over the years. The (now fading) practice of “parallel R&D” groups was put in place to *stimulate* intra-firm competition, and at least some managers believed that this created conditions that favor innovation instead of undermining it. *See, e.g.,* Richard R. Nelson, *Uncertainty, Learning, and the Economics of Parallel Research and Development Efforts*, 43 REV. ECON. & STATISTICS 351 (1961).

64. Andrea Signori & Silvio Vismara, *Me&A Synergies and Trends in IPOs*, 127 TECH. FORECASTING & SOC. CHANGE 141, 141 (2018).

65. For a summary of findings from this older literature, see MORTON I. KAMIEN & NANCY L. SCHWARTZ, *MARKET STRUCTURE AND INNOVATION* 103 (1982):

The bulk of the empirical findings [as of 1982] indicate that inventive activity does not typically increase faster than firm size, except in the chemical industry. R&D activity, measured by either input or output intensity, appears to increase with firm size up to a point and then level



A. THE “TACIT DIMENSION” AND MARKETS FOR “DISEMBODIED TECHNOLOGIES”

Economists have acknowledged the existence of tacit knowledge, technical information that is difficult to write down or codify, since at least the 1960s. Michael Polanyi’s famous 1966 volume *The Tacit Dimension* describes craft and technical skills that are difficult or even impossible to write down and hand off to another person.<sup>66</sup> For this type of knowledge, it is far more efficient to hire people than to try transferring the information in a disembodied form. It is either impossible or very difficult to transfer tacit knowledge in an arm’s length market for disembodied assets. The best and sometimes only way to transfer tacit knowledge from Organization A to Organization B is for Organization B to somehow acquire, absorb, and retain the employment relationship with the employees from Organization A; in other words, B has to acquire A’s people. The things, procedures, and written records of A’s people are not good enough. If people themselves are not part of the deal, crucial tacit know-how will not survive the transfer from A to B. It will instead evaporate and be lost in the hands and minds of A’s employees.

Professor Peter Lee has documented this fact well. He has written an article calling into question the supposed ascendancy of “dis-integrated” business models in the current era.<sup>67</sup> He observes that arm’s-length transfers of disembodied products, technologies, and patent rights will be inferior to full-on corporate acquisitions as long as the tacit dimension is important.<sup>68</sup> With a full acquisition comes the right to assume the acquired firm’s

---

off or decline, as is consistent with the evidence on the nature of the R&D process.

*Id.* Kamien and Schwartz also note that “market structure intermediate between monopoly and perfect competition [may be the ideal for innovation purposes.” *Id.* at 104. The authors conclude, “[e]mpirical studies over the last fifteen years have consistently shown that, although there may sometimes be certain advantages of size in exploiting the fruits of R&D, it is more efficiently done in small or medium size firms than large ones.” *Id.* at 66; *see also* FREDERIC M. SCHERER, INNOVATION AND GROWTH: SCHUMPETERIAN PERSPECTIVES 182 (1984) (noting, for all firms studied, R&D inputs (such as R&D employment) and outputs (patents) increase “less than proportionately” with size, where size is measured by firm sales); Zoltan J. Acs & David B. Audretsch, *Innovation in Large and Small Firms: An Empirical Analysis*, 78 AM. ECON. REV. 678 (1988) (presenting industry concentration measures, which estimate the degree of monopolization or oligopolistic dominance in an industry, are statistically associated with reduced innovation).

66. MICHAEL POLANYI, *THE TACIT DIMENSION* (Reissue Ed., 2009).

67. Peter Lee, *Innovation and the Firm: A New Synthesis*, 70 STAN. L. REV. 1431 (2018).

68. *See id.* at 1500 (“[P]atents do not disclose significant tacit knowledge that is valuable for practicing a technology and adapting it to commercial use. Indeed, it is precisely these knowledge deficiencies that contribute to vertical integration in patent-intensive industries.”).

employment contracts; the deal includes people as well as the disembodied assets they have created. When the tacit skills of individual people are important or the future stream of creative work matters, acquisitions will be superior to patent transfers.

B. AREN'T FIRM MERGERS AND ACQUISITIONS (ALMOST) ALWAYS SUPERIOR TO PATENT SALES?

This logic raises an obvious objection. Precisely because independent thinking is good, Big Platform companies acquire the small fry instead of growing all desired capabilities in-house. These companies also value diversity and autonomy; when these positive virtues result in valuable innovations, Big Platform rewards those innovators by acquiring their companies. If this is true, then there is no need to maintain the small company as a going concern to encourage innovation.

This phenomenon of absorbing the most innovative companies presents its own problems. A successful Big Platform acquisition represents a fine reward for innovation, but startups still call that acquisition an “exit.” The innovative team is absorbed into a big company and the small startup or emerging company is no longer independent. This makes acquisition a double-edged sword. It is a reward for *past* innovation, but a sizeable body of research suggests that it is a damper on *future* innovation. The team that develops a technology will cash out nicely, but the autonomy and independence that created the context for the original innovation will be gone. Despite heroic efforts to preserve the best of both worlds—namely by the massive acquiring company pledging to “keep hands off” and “preserve the special culture” of the acquired company—acquisition brings an inevitable change. If a large company could completely duplicate the culture of the startup, it would do so from the outset and develop the technology in-house. In the end, two stark facts usually stand out; the acquired company did what it did because it was plucky and independent, and after the acquisition it becomes part of a big company. When technology is acquired through acquisition of an entire company, autonomy and diversity both exit the scene and never fully return.<sup>69</sup>

---

69. Cf. Victor Luckerson, *How Google Perfected the Silicon Valley Acquisition*, TIME (Apr. 15, 2015), <https://time.com/3815612/silicon-valley-acquisition/> [https://perma.cc/6DA9-25RB] (“Oftentimes [after an acquisition] founders are rolled up inside another group inside of the company. They can’t make decisions as freely as when they were entrepreneurs. That affects people’s willingness to stick around.”) (quoting Justin Kan, venture capitalist at Y Combinator and cofounder of Twitch).

### 1. *For Radical Innovation, More Is Better and Small Is Big*

A second major point regarding large firm acquisitions is that they reduce the chance for radical innovation. There are two reasons. First, they reduce the total number of separate firms in a given field. Second, they eliminate from the landscape precisely the sort of smaller firms that have been the source of paradigm-changing innovations throughout history. In these two ways, the loss of radical post-merger innovations is the major cost of large firm acquisitions, despite post-merger efficiencies.

In its simplest form, a corporate merger executes a form of legal arithmetic:  $1 + 1 = 1$ . What starts with two separate firms ends with one. Whatever gains this brings in operations and in more efficient R&D, it entails a loss; an independent firm ceases to be. The consequences for future innovation are well understood in an aggregate sense but hard to pin down in any particular case. Future innovation is by its nature hard to predict, but students of long-term innovation patterns are fairly uniform in their assessment of the optimal number of firms—more is better. It is impossible to quantify what is lost when there are fewer separate firms to take part in the innovation sweepstakes, but on average throughout time, something is surely lost.

### 2. *Small Is Big*

The argument thus far establishes why *more* firms might make for more innovation in a given industry. The below arguments address another point—why *small* firms add to innovation in ways that make them superior to big ones.<sup>70</sup> All of them are variations on a single theme; smaller firms are more resourceful, nimble, focused, and productive, and hence more likely to come up with something new and different. As one study put it, summarizing

---

70. Definitions of “small” and “big” can of course vary, but in general small firms usually have fewer than 500 employees, and often fewer than 100, while large firms usually measure their workforces in the thousands. For a study of the very smallest firms and their ability to innovate, see Julian Baumann & Alexander S. Kritikos, *The Link Between R&D, Innovation and Productivity: Are Micro Firms Different?*, 45 RES. POL’Y 1263 (2016) (presenting data on German micro-firms, drawn from 10,000–15,000 firms in a total sample of firms in Germany, between 2005 and 2012). The authors find that most micro firms are young: “53% of the smallest firms were younger than 15 years.” *Id.* at 1266. “[L]arger [small firms] have a lower R&D intensity than smaller ones: ceteris paribus, small firms invest 36% more in R&D per employee, firms with 0–4 FTE employees invest 90.4% more in R&D per employee than medium-sized firms.” *Id.* at 1267. R&D intensity increases process and product innovations for all sized firms (which is to be expected). *Id.* at 1268. “Micro firms that do invest in innovation activities have 90% higher R&D expenditures per employee than medium-sized firms. Thus, firm size is negatively correlated with R&D intensity.” *Id.* at 1271.

a large literature: “Empirical research on innovation and firm size confirms that despite large firms’ apparent advantages in scale and access to complementary assets and capabilities . . . small firms are more efficient at innovation, particularly radical forms of innovation.”<sup>71</sup>

Business people and scholars have named three different benefits to smallness for purposes of generating innovations: (1) magnified incentive effects; (2) better focus, meaning simpler and more direct decision processes within firms; and (3) the preference of those with an “entrepreneurial personality” for greater autonomy, which is better satisfied in small firms.

The first benefit hinges on the idea that small firms have more riding on their relatively few research projects. They therefore have less distraction and experience greater rewards when they succeed. Failure is more painful because the future of the company may be riding on a single research project. Success is also sweeter because the individual researchers often own a significant chunk of the entire small company.<sup>72</sup> Some theorists have described how big companies can leverage these features of small firms by entering into contracts that provide large rewards for project success. This is an example of the “high powered incentives” that economist Oliver Williamson delineated as an advantage of contractual exchange over integration or ownership.<sup>73</sup> Large firms are much more diffuse; individual projects pale in comparison to the overall scale of the firm. Additionally, individual effort is dwarfed by the totality of collective effort, so there is less direct reward for extraordinary effort. Large companies can access these

---

71. Todd R. Zenger & Sergio G. Lazzarini, *Compensating for Innovation: Do Small Firms Offer High-powered Incentives That Lure Talent and Motivate Effort?*, 25 MANAG. DECIS. ECON. 329, 329 (2004). As regards overall innovation efficiency, this conflicts with some of the studies cited in the preceding Section; that might be explained by the fact that many of the studies showing greater overall efficiency for post-merger firms were published after this article was. The conclusion regarding radical innovation, however, has not been superseded in the intervening years.

72. *Id.* at 342 (“[T]he results [of this study of 352 engineers in Silicon Valley and Route 128 in the Boston area] . . . provide consistent evidence that outcomes are linked directly to differences in contract attributes, which in turn are related to firm size. Firms with more aggressive reward systems appear more successful in motivating high effort and in luring and retaining top talent. Engineers with larger equity shares and a greater variable component to their pay work longer hours and are more likely to bring work home. Strong norms of peer monitoring may further escalate effort in small firms. By contrast, engineers with small equity shares, those employed in contracts with weak incentive intensity and weak peer performance pressure are less likely to work long hours and bring work home.”).

73. This is an application of Oliver Williamson’s transaction cost economics. *See generally* OLIVER E. WILLIAMSON, *THE MECHANISMS OF GOVERNANCE* (1996); Robert P. Merges, *A Transactional View of Property Rights*, 20 BERKELEY TECH. L. J. 1477, 1483 (2005).

stronger incentives only indirectly—by contracting with small firms to supply research services or research-intensive inputs.<sup>74</sup>

A small team that puts all of its energy into a challenging project and is under the pressure of a specific contract requiring the team to deliver will on average work harder than a larger team embedded in a larger company. That idea is what fuels “high powered incentives” that accompany a contract specifying a discrete “deliverable.” A researcher working in a research division of a large company cannot typically be strongly motivated. But a researcher or small team under pressure to deliver a specific result for a contractual reward can be expected to concentrate more and work harder. The downside of failure is greater in that the small firm might fail or experience a serious setback, and the upside of success is also greater if the contract is written so as to reward success robustly.

The second reason some researchers say small firms are superior is the relative lack of bureaucracy. One literature summary identifies “a loss of marginal control or [a] high level of bureaucratic control” as among the characteristics of large firms that cause R&D performance to decrease.<sup>75</sup> The perils of large bureaucracies are well understood but seem especially salient with respect to R&D activities, where freedom from bureaucratic oversight is especially important.<sup>76</sup> One pair of researchers noted that “it is not the size of firms per se, but rather the internal processes activated as firms evolve in size that affect innovation outcomes.”<sup>77</sup> In planning for innovation, large firms typically gather more information as part of detailed analytical procedures. They “tend to make decisions in a more planned and more formal manner . . . than small firms.”<sup>78</sup> This is partly due to organizational routines and styles

---

74. See Ashish Arora & Robert P. Merges, *Specialized Supply Firms, Property Rights and Firm Boundaries*, 13 INDUS. & CORP. CHANGE 451 (2004); see generally Bo Carlsson et al., *Knowledge creation, entrepreneurship, and economic growth: a historical review*, 18 INDUS. & CORP. CHANGE 1193, 1222, 1223 (2009) (“There are two main reasons why small firms have become more important in recent decades. One is that small firms simply do certain things (such as certain types of innovation) better than large firms. As a result, through division of labor between small and large firms, the efficiency of the economy is increased. The other reason is that small firms provide the entrepreneurship and variety required for macroeconomic growth and stability . . .”).

75. Negin Salimi & Jafar Rezaei, *Evaluating Firms’ R&D Performance Using the Best Worst Method*, 66 EVAL. & PROG. PLAN. 147, 148 (2018).

76. See, e.g., Clayton M. Christensen & Joseph L. Bower, *Customer Power, Strategic Investment and the Failure of Leading Firms*, 17 STRAT. MGT. J. 197 (1996) (asserting that innovation is negatively affected because allocation of resources is not autonomously decided but instead depends on what the biggest customers would likely want).

77. José Lejarraga & Ester Martinez-Ros, *Size, R&D Productivity and Decision Styles*, 42 SMALL BUS. ECON. 643, 644 (2014).

78. *Id.* at 646.

and partly due to increased monitoring; “as firms increase in size, managers become subject to closer monitoring by the firm’s board of directors and shareholders, who expect decision making to be based on justifiable arguments.”<sup>79</sup> Finally, with more layers of review and perhaps more competition over recognition and resources—in what is often called “company politics”—the personal agendas of corporate employees may come into play more often in large firms.<sup>80</sup>

Big firms recognize that their complex structures often fit poorly with the process of innovation. The spate of acquisitions by Big Platform companies and others attests to this; what the “bigs” cannot make, they buy. However, it is also borne out by the institution of “skunk works”—semi-secret or “unofficial” R&D projects within large companies that are conducted outside normal oversight and review procedures.<sup>81</sup> Indeed, complex oversight and approval seem like an anathema to successful R&D in whatever form. A study of 464 R&D joint ventures in the telecommunications industry found that “[c]ollaborative benefits [from these joint ventures] are diminished most by selection of governance that imposes excessive bureaucracy . . .”<sup>82</sup> Whatever the industry, multi-stage decision procedures and more complex organizational landscapes seem to be the enemy of important innovation. Like many large companies before them, Big Platform companies are aware of these failings; acquisitions are one response to them. While those acquisitions may help address the “innovation deficiency” that often plagues big companies, these acquisitions come at the cost of extinguishing small innovators.

---

79. *Id.* at 646–47.

80. See TOM BURNS & G.M. STALKER, *THE MANAGEMENT OF INNOVATION* 195 (1994). Quoting a research scientist brought into an industrial company to open an R&D lab: “What happens is that a plan devised in terms of changing the working organization [to include an R&D lab] fails to materialize because factors of status and politics play a determining role, and nobody realizes, or rather, admits, that these are real problems to be dealt with.” *Id.* Describing R&D lab at one company: “[P]olitical conflicts do appear out of situations in which changing circumstances constitute a threat to existing parts of the working community. This happens when the new circumstances themselves are institutionalized.” *Id.* at 199.

81. *Skunkworks*, WIKIPEDIA (last visited Dec. 21, 2019), [https://en.wikipedia.org/wiki/Skunk\\_Works](https://en.wikipedia.org/wiki/Skunk_Works) [<https://perma.cc/Z6LJ-FYJS>] (“The designation ‘skunk works’ or ‘skunkworks’ is widely used in business, engineering, and technical fields to describe a group within an organization given a high degree of autonomy and unhampered by bureaucracy, with the task of working on advanced or secret projects.”). The name was first used at Lockheed Aeronautics; it was taken from the old L’il Abner comic strip; in that comic series, it was the name of a moonshine liquor still. *Id.*

82. Rachele C. Sampson, *The Cost of Misaligned Governance in R&D Alliances*, 20 J.L. ECON. & ORG. 484, 485 (2004).

Unlike enhanced incentives and reduced bureaucracy, the final advantage of small firms relies less on their environment and more on the personalities of those who found and staff them. For many scholars, it is not firm size that shapes the entrepreneurial innovator; it is the entrepreneur who shapes the features of the small firm with his or her distinctive taste for autonomy and independence.

Some detailed research suggests that engineers and scientists who have a strong preference for autonomy and challenging projects tend to work at startups, while those impelled by security and risk avoidance more often work at large companies.<sup>83</sup> These differing motivations produce different outcomes; the autonomy valued by startup researchers creates the right sort of environment for radical innovation. As the title of one journal article says, “Being Independent is a Great Thing.”<sup>84</sup>

Small firms admittedly have their own pressures. One is that the venture capital finance that makes startups possible brings external monitoring and accountability. Another is that although choosing one’s career direction is exciting, it is also risky; going “all in” on a single project means little chance to deflect blame or soften the blow if it fails. Apparently, however, these negatives are outweighed for at least some people by the relative freedom from hierarchical oversight.<sup>85</sup> The simple act of choosing one’s own course holds personal rewards.<sup>86</sup>

This self-selection also has ramifications for the larger economy. Because small firms are founded out of a desire for personal autonomy, they supply diverse and far-flung sources of fresh ideas. They ensure that many minds attack technological problems from many different, uncoordinated starting points. By decentralizing decision making, they make it more likely that a

---

83. Henry Sauermann, *Fire in the Belly? Employee Motives and Innovative Performance in Startups Versus Established Firms* 14 (Nat’l Bureau of Econ. Research, Working Paper No. 23099, 2017), <https://www.nber.org/papers/w23099> [<https://perma.cc/6DL9-DDDX>]. But cf. Thomas Lange, *Job Satisfaction and Self-Employment: Autonomy or Personality?*, 38 SMALL BUS. ECON. 165 (2012) (finding that, based on survey data, the extent of autonomy explains higher job satisfaction among self-employed men and women better than measures of various individual personality traits; a preference for autonomy, in this study, is not treated as a personality trait in and of itself).

84. Matthias Benz & Bruno Frey, *Being Independent is a Great Thing: Subjective Evaluations of Self-Employment and Hierarchy*, 75 ECONOMICA 362 (2008).

85. See Martin A. Carree & Ingrid Verheul, *What Makes Entrepreneurs Happy? Determinants of Satisfaction Among Founders*, 13 J. HAPPINESS STUD. 371 (2012).

86. Cf. Robert P. Merges, *Autonomy and Independence: The Normative Face of Transaction Costs*, 53 ARIZ. L. REV. 145 (2011) (arguing that even if multiple small firms add a modest increment to transaction costs in a given industry, the intrinsic value of autonomy might make it worthwhile to tolerate and encourage some small firms in that industry’s structure).

small team “off the radar” of the established research paradigm will develop an unconventional or novel approach—the type of approach that can lead to a radical innovation.<sup>87</sup>

### 3. *Innovating “Outsiders”: A Complicating Factor?*

According to the standard account, the typical source of “radical innovation” is an “outsider”—a person or firm from outside the industry that is disrupted or changed by the radical innovation. Social psychologists may provide the best explanation of why this is so through the concept of “cognitive distance.” In this research, each person has a mental framework consisting of vocabulary, assumptions, and ways of looking at problems. Cognitive distance measures the distance between two persons’ mental frameworks.<sup>88</sup> For purposes of innovation, closely aligned frameworks make for easy working relationships and productive incremental results. Nonetheless, it also produces a “groupthink” dynamic that does not lead to radical innovation.<sup>89</sup> In contrast, wildly divergent mental frameworks make it almost impossible for people to understand each other. Without a common ground, cooperative research is fruitless. Radical innovation comes not from excessive overlap or from the absence of overlap, but instead from a “just right” degree of overlap. When cognitive distance is too great, people “talk past each other” and collaboration is very difficult; but when this measure of distance is too small, people have nothing new to share with each other and their collaboration becomes sterile.<sup>90</sup>

As might be expected, cognitive distance between R&D personnel is reduced when a single organization amasses a large stock of R&D. This is good for incremental innovation because R&D efficiency increases; bigger is better for creating minor inventions. However, greater cognitive distance benefits more radical innovations; important new ideas very often come from

---

87. Each small firm also does its part to perpetuate the overall culture of small firms, the ethos and norms of this type of firm. By keeping this culture alive, even an unsuccessful small firm may sow the seeds of a future success. See Daniel W. Elfenbein et al., *The Small Firm Effect and the Entrepreneurial Spawning of Scientists and Engineers*, 56 MGMT. SCI. 659 (2010) (finding that researchers from small firms are more likely to subsequently be self-employed).

88. See Bart Nooteboom et al., *Optimal Cognitive Distance and Absorptive Capacity*, 36 RES. POL’Y 1016, 1016 (2007) (defining cognitive distance as “interpersonal difference between life experience and perceptual frameworks”).

89. Though this is a consensus view, there are outliers. See, e.g., Rajesh K. Chandy & Gerard J. Tellis, *The Incumbent’s Curse? Incumbency, Size, and Radical Product Innovation*, 64 J. MARKET. 1 (2000) (presenting a historical study of sixty-four radical innovations in the consumer and office products markets that found that the traditional “outsider” innovation story was accurate until roughly 1945, but after that year large incumbents were responsible for a growing proportion of radical innovations in these industries).

90. See Nooteboom et al., *supra* note 88, at 1016.



the confluence of hitherto unrelated technical fields. For the most significant radical innovations, close cognitive proximity—as measured by single-firm accrued R&D stock—makes no difference, meaning that large size does not confer any advantages. As the authors of one study put it, “an increase in R&D-efforts will lead to more patents in the patent classes that the firm already masters” but not in new technologies due to “the high levels of uncertainty in explorative research.”<sup>91</sup>

If outsiders are so important for radical innovations, the importance of preserving smaller companies within a given industry is harder to judge. If firms labeled as outsiders would be acquisition targets for Big Platform and other large companies, policies that preserve small outsiders are still important. Special solicitude for small firms also makes sense if there are only a few large firms in an industry and the cognitive distance between them is small—as happened with the three largest U.S. auto companies before the entry of overseas car companies in the 1980s.<sup>92</sup>

However, if outside firms are *infrequent* candidates for large firm acquisitions, the growth of big companies by merger poses less of a threat to the prospects for radical innovation. Precisely because they are “outsiders,” these firms are not on the “radar screens” of the big companies. Perhaps there will always be such outsiders, no matter how many “inside” firms are vacuumed up in large firm acquisitions. Perhaps the history of radical innovation teaches us not to worry so much. Additionally, if the large firms in an industry have employees with the right “cognitive distance” from the employees of other large firms, maybe radical innovation can result from combinations of large firms working together—even in the absence of small firms. The research on cognitive distance relates to the cognitive styles of people inside different organizations; it is not directly related to firm size in any way.

Despite these potential concerns, the available evidence indicates that preserving cognitive distance requires protecting against excessive merger activity to cultivate an industrial ecosystem that includes small firms. The research cited earlier on R&D productivity and cognitive distance is based on pairs of firms involved in collaborative R&D.<sup>93</sup> If a firm can find a partner for collaborative R&D, it could presumably acquire that firm just as easily. This means that the research partners in this and similar studies are not

---

91. *Id.* at 1027.

92. See DAVID HALBERSTAM, *THE RECKONING* (1986) (providing a historical overview of the auto industry in the United States and Japan).

93. Nooteboom et al., *supra* note 88, at 1021 (presenting data on research “alliances” between pairs of firms).

unknown to each other. Having firms of different sizes may also make it more likely that a variety of cognitive distances are present between employees of different firms. The research cited earlier<sup>94</sup> explained that the personalities and preferences of small firm entrepreneurs differ systematically from those of large firm research employees. This alone makes it more likely that some of these small firms will “see things differently” and that a more optimal degree of cognitive distance will therefore open up between them and the employees of large firms. Small firms are likely to be beneficial due to the reasons explored in the earlier sections, as well as the possibility that they will have the “just right” degree of cognitive distance from large firms to make radical innovation more likely.

C. MERGERS AND INDUSTRY STRUCTURE: SUMMARY

This Section makes the case that a variegated industry structure, one that includes a number of smaller firms, gives the best chance for important future innovations. As one study summarized it:

[T]he results show that larger firms enjoy greater advantages for incremental innovation performance . . . but not for radical innovation performance on which large firm size has a negative non-significant effect . . . . Large firms rarely introduce radical innovation performance; rather they tend to solidify their market positions with relatively incremental innovations . . . .<sup>95</sup>

Both these themes—increased overall innovation and decreased radical innovation—are apparent from a large-scale study of post-merger R&D in European companies. Economist Joel Stiebale studied 941 European mergers between 1978 and 2008 using data on the nationality of inventors listed on patent applications. The results show that after many mergers, inventive activity increases in the country where the acquiring company is located but decreases in the country that is home to the acquired, or target, company.<sup>96</sup> The study recognizes that after a merger, consolidation of patent activities in the headquarters of the acquiring company is to be expected, and that therefore more patent applications will originate from the home country of the acquiring company after the merger. To adjust for this, Stiebale tests national-level inventiveness by the domiciles of listed inventors on those patent applications, rather than by the applications originating in the

---

94. See notes 83–85 and accompanying text.

95. Beatriz Fores & César Camisón, *Does Incremental and Radical Innovation Performance Depend on Different Types of Knowledge Accumulation Capabilities and Organizational Size?*, 69 J. BUS. RES. 831, 836 (2016) (references omitted) (summarizing literature).

96. Joel Stiebale, *Cross-border M&As and Innovative Activity of Acquiring and Target Firms*, 99 J. INT'L ECON. 1 (2016).

acquiring and target home countries. The patent department and thus filing country may change after an acquisition, but the inventors usually stay put. Stiebale finds that R&D productivity of the acquired company drops, as measured by the number of patent applications filed by its inventors in the post-merger period.

The larger the patent portfolio of the acquiring company (which Stiebale calls the “knowledge stock”), the greater the drop in inventiveness in the country where the target firm is located.<sup>97</sup> The data show that innovative activities become more concentrated in the home country of the acquiring firm—a sign of the increased efficiency that accompanies R&D-oriented acquisitions.<sup>98</sup> From an overall efficiency standpoint, there is a good and defensible reason for this result; it shows that “innovation activities are not relocated from targets to acquirers per se” but to whatever part of the firm is “more efficient in innovation.”<sup>99</sup> Nevertheless, another finding of this study stands out: there is a loss of innovative vigor on the part of the target firm after these mergers. Efficiency is gained, but what could be paradigm-stretching creativity is lost.

Admittedly, this study documents a drop in inventiveness only for the trans-national, intra-European mergers studied. It is possible that these results pertain to European mergers in some peculiar way. Aside from this, however, the study sounds a cautionary note. While the overall *volume* of innovation increases in the expected way after a merger, this comes at the expense of the innovative output of the acquired firm. While the gains in efficiency may outweigh the loss of a highly innovative independent firm, the theory and experience reviewed earlier tell us to be wary of the long-term effects. Multiple, rivalrous sources of innovation are still a good thing; one might even view them as a good in and of themselves. Losing many autonomous firms to the merger trend may generate serious costs in the long run.

One historical study published in 1969, aptly entitled *The Sources of Innovation*,<sup>100</sup> takes a long-term perspective regarding industry structure and reflects many of the arguments presented here. In this study, as with more recent literature, small firms are often the heroes of innovation stories. As in 1969 and the times when the innovations studied were being developed,

---

97. *Id.* at 11.

98. *Id.*

99. Stiebale, *supra* note 96, at 11.

100. JOHN JEWKES ET AL., *THE SOURCES OF INVENTION* 211–12 (2d ed. 1969) (summarizing the invention and development of fifty-six important innovations, including the ball point pen, catalytic cracking of petroleum, new polymers such as polypropylene, the transistor, etc.).

there is an important place for small firms in a healthy R&D-rich industrial ecosystem. That was true before the Big Platform companies, and it remains true now as well.

#### IV. SUGGESTED REFORMS TO ASSIST THE PATENT MARKET

This Article has established that patent markets can serve an important purpose in an era when “bigness” is reasserting itself as an economic imperative. Although this Article therefore comes to defend patent markets rather than condemn them, they are no panacea. They have limits and create inefficiencies, which makes them far from perfect as a solution to the potential problems of the Big Platform era. This Part identifies one important problem before offering constructive suggestions.

##### A. RELATIONSHIP TO LITIGATION: DO PATENT MARKETS “FEED THE TROLLS”?

The greatest inefficiency of the patent market is that it is tethered to the thoroughly inefficient business of patent litigation. Reasonable parties on both sides of a patent transaction would ideally predict potential court outcomes, bargain accordingly, and stay away from court; this happens in roughly half of these transactions.<sup>101</sup> The other half, unfortunately, lead to some stage of the litigation process. The result is that the patent market seems intimately bound to the fraught phenomenon of patent litigation.

##### B. POLICIES TO SUPPORT THE SECONDARY PATENT MARKET

Changes in both antitrust law and the rules regarding patent rights would assist in strengthening patent markets. This can in turn mitigate the effects of Big Platform companies.

###### 1. *Antitrust Law*

Antitrust law plays an indirect role in promoting patent markets. The chief contribution it can make is to recognize the importance of small,

---

101. Until recently we might have guessed that as many as 90% of patent-related transactions were conducted without recourse to formal enforcement of some sort. But the dismal fact that the number is closer to 50% has now been established. See Mark A. Lemley et al., *The Patent Enforcement Iceberg*, 97 TEX. L. REV. 801, 803 (2019). For general treatments of the costs and benefits of litigation, see Louis Kaplow, *Private Versus Social Costs in Bringing Suit*, 15 J. LEGAL STUD. 371, 371 (1986); Peter S. Menell, *A Note on Private Versus Social Incentives to Sue in a Costly Legal System*, 12 J. LEGAL STUD. 41, 41 (1983); Steven Shavell, *The Fundamental Divergence Between the Private and the Social Motive to Use the Legal System*, 26 J. LEGAL STUD. 575, 577–79 (1997).

independent firms in the innovative ecosystem of technology-intensive industries. This will apply mostly when antitrust authorities are asked to review a sale of patents by a small firm to a larger firm, such as a platform company. Patent acquisitions are routinely reviewed for compliance with antitrust law;<sup>102</sup> they are suspect because they combine the resources of two firms in a “horizontal” (competitor-to-competitor) arrangement.<sup>103</sup> In reviewing such an arrangement, antitrust agencies and courts should consider both the short-term effect on consumers and the long-term benefits of the survival of small firms. There may be cases where a large firm acquires some added short-term market power due to the purchase of patents. While this is not to be ignored, it must be weighed against the benefits of small firm survival—which may be dependent on the sale of patents. The prospect of future innovation potential needs to be part of the regulatory calculus.

In general, antitrust review centers on the relationship between patent holdings and market power. In merger analysis, for example, antitrust authorities in the past have sought to ameliorate the effects of enhanced post-merger market power by requiring the newly merged company to license patents to a third party.<sup>104</sup> The aim in such cases is to create

---

102. *Intellectual Ventures I LLC v. Capital One Fin. Corp.*, 280 F. Supp. 3d 691, 697 (D. Md. 2017); *Kobe, Inc. v. Dempsey Pump Co.*, 198 F.2d 416, 423–25 (10th Cir.), *cert. denied*, 344 U.S. 837 (1952) (acquiring a portfolio of patents to “corner the hydraulic pump business for oil wells” constituted illegal monopolization); *United States v. Westinghouse Elec. Corp.*, 648 F.2d 642, 647 (9th Cir. 1981) (“[A] patent holder may run afoul of the antitrust laws . . . by expanding that monopoly by . . . accumulation.”); *SCM Corp. v. Xerox Corp.*, 645 F.2d 1195, 1205 (2d Cir. 1981) (“Surely, a § 2 violation will have occurred where, for example, the dominant competitor in a market acquires a patent covering a substantial share of the same market that he knows when added to his existing share will afford him monopoly power.”).

103. *Hurricane Shooters, LLC v. Emi Yoshi, Inc.*, No. 8:10-cv-762-T-30AEP, 2010 WL 4983673, \*2–3 (M.D. Fla. 2010). In denying a patentee’s motion to dismiss the accused infringer’s antitrust counterclaim, the court said this about patent acquisitions:

Count II alleges that Plaintiff acquired title to a competitor’s patent (McNaughton Inc.) in order to restrain commerce in the relevant market, by requiring other competitors, like Defendant, to take a license from Plaintiff at an exorbitant royalty. Defendant also alleges that Plaintiff has acquired more than 10 patents covering [the market for the patented product] . . . in order to obtain licenses from competitors at exorbitant rates. At this stage, this is sufficient to state a claim. Defendant has alleged that McNaughton Inc. conspired or combined to restrain competition . . . . [I]t is not a violation of the antitrust laws to acquire patents from others. [But if] it is determined, at a later stage, that these allegations were lacking in merit, the Court will not hesitate to award sanctions.

*Id.*

104. *See In re Ciba-Geigy Ltd., et al.*, 123 F.T.C. 842 (1997) (noting a consent decree requiring divestiture of lines of business and/or licensing of patents to third parties in the

competition if the new firm would have excessive market power in the absence of such a license. A good example of this is a Federal Trade Commission (FTC)-managed consent decree from 1995.<sup>105</sup> The two largest producers of polypropylene technology had proposed a joint venture (JV)<sup>106</sup> aimed at broad cooperation in the polyolefin (plastics) industry. The FTC ordered that the parties divest the JV of all plants, patents, and related assets pertaining to polypropylene; this was to prevent the JV from dominating that part of the industry.<sup>107</sup>

In antitrust analysis of patent acquisitions, authorities look at the effects of patent purchases on product markets. The emphasis is on whether the patents give the acquiring firm some extra degree of market power over rivals in these “downstream” markets (markets for products derived from or drawing upon the patented technology).<sup>108</sup> A typical antitrust review of this type came in *ABS Global, Inc. v. Inguran dba Sexing Technologies, LLC*,<sup>109</sup> where antitrust plaintiff ABS Global argued that patent acquisitions by defendant Sexing Technologies (ST) violated § 2 of the Sherman Act.<sup>110</sup> According to

---

fields of gene therapy, pet medicines, and corn herbicides). The consent decree requires that the merged firm license a specific competitor—one judged to be in the best position to promote competition:

[The parties, i.e., the merged firm] shall (i) grant a non-exclusive license to [third party Rhone Poulenc Rofer, Inc.] to make, use and sell [Herpes simplex virus-thymidine kinase (“HSV-tk”) gene therapy products, for the treatment of cancer], under [the merged firm’s] HSV-tk Patent Rights . . . or (ii) grant a nonexclusive license to make, use and sell HSV-tk Licensed Products under [the merged firm’s] HSV-tk Patent Rights to an HSV-tk Licensee that receives the prior approval of the Commission and in a manner that receives the prior approval of the Commission, in perpetuity and in good faith, at no minimum price. In consideration for the HSV-tk License, each [party] may request from the HSV-tk Licensee compensation in the form of royalties and/or an equivalent cross-license.

*Id.*

105. *In re Montedison S.P.A., et al.*, 119 F.T.C. 676 (1995).

106. According to the consent decree, the parties “collectively account for over 80% of completed and projected additions to capacity pursuant to [polypropylene] technology licenses since 1990. Other technologies are not a significant competitive constraint.” *Id.* at 681.

107. *Id.*

108. See Fiona M. Scott Morton & Carl Shapiro, *Strategic Patent Acquisitions*, 79 ANT. L.J. 463, 463 (2014) (“Our analysis has much in common with merger analysis: we study how a strategic patent acquisition changes economic incentives and trace through the likely economic effects of those changed incentives.”).

109. *ABS Global, Inc. v. Inguran L.L.C.*, No. 14-CV-503-WMC, 2016 WL 3963246 (W.D. Wis. July 21, 2016).

110. 15 U.S.C. § 2 (“Every person who shall monopolize, or attempt to monopolize, or combine or conspire with any other person or persons, to monopolize any part of the trade

ABS Global, these acquisitions were part of an effort to monopolize the market for sex-sorted bull semen used for artificial insemination in the cattle industry. ST's strategic patent acquisitions began after expiration of ST's foundational patent, the "Johnson patent":

Since the Johnson Patent expired in 2006, ST has purchased, acquired or licensed several U.S. patents related to sexed semen processing. Principally, ST acquired control of XY, Inc., in 2007. At the time, ST was one of several U.S. licensees using [three of the important] XY . . . [patents] . . . . ST is now XY's sole current licensee for its patented sexed semen process in the United States for bull studs. Since 2007, XY has also been a wholly-owned subsidiary of ST. In 2008, ST also purchased several pending patent applications related to sexed semen processing from Monsanto Company. . . . Those applications matured into 24 U.S. patents, including [two] that [were asserted against the defendant/antitrust counterclaimant] . . . here. Finally, ST obtained an exclusive license for nonhuman applications to a portfolio of U.S. patents relating to sexed semen processing from Cytonome, Inc., covering an additional 46 U.S. patents related to sexed semen.<sup>111</sup>

Antitrust defendant had thus acquired a collective portfolio of seventy-three patents covering the technology at issue in the case. Both parties moved for summary judgment on the antitrust issue, but the court declined to grant either motion. In explaining why, the court gave some useful instruction in the whys and wherefores of antitrust claims based on patent acquisitions:

Any [Sherman Act] § 2 claim based on the acquisition of patents presents an "obvious tension between the patent laws and antitrust laws. One body of law creates and protects monopoly while the other seeks to proscribe it." *United States v. Westinghouse Elec. Corp.*, 648 F.2d 642, 646 (9th Cir. 1981). Indeed, acquiring and asserting valid patents is absolutely protected by the patent laws "in the absence of monopoly but, because of their tendency to foreclose

---

or commerce among the several States, or with foreign nations, shall be deemed guilty of a felony . . ."). The same antitrust offenses punishable as felonies under criminal law can be the subject of private civil suits, due to Section 4 of the Clayton Act, codified at 15 U.S.C. § 15:

[A]ny person who shall be injured in his business or property by reason of anything forbidden in the antitrust laws may sue therefor in [a] district court of the United States . . . without respect to the amount in controversy, and shall recover threefold the damages by him sustained, and the cost of suit, including a reasonable attorney's fee.

15 U.S.C. § 15.

111. *ABS Global*, 2016 WL 3963246, at \*3.

competitors from access to markets or customers or some other inherently anticompetitive tendency, they are unlawful under § 2 if done by a monopolist [.]” *City of Mishawaka, Indiana [v. Am. Elec. Power Co.]*, 616 F.2d 976, 986 (7th Cir. 1980)] . . . at 986 (quoting *Sargent-Welch Sci. Co. v. Vernon Corp.*, 567 F.2d 701, 711–12 (7th Cir. 1977)).

Here, ABS has shown enough to suggest that ST’s acquisition of patents may qualify as unlawful under the Sherman Act. See *SCM Corp. v. Xerox Corp.*, 645 F.2d 1195, 1205 (2d Cir. 1981) (“Surely, a § 2 violation will have occurred where, for example, the dominant competitor in a market acquires a patent covering a substantial share of the same market that he knows when added to his existing share will afford him monopoly power.”); *L.G. Balfour v. F.T.C.*, 442 F.2d 1, 15 (7th Cir. 1971) (disagreeing with the petitioners that the cases they cited “[stood] for the proposition that the accumulation of patents . . . may never constitute a violation of the antitrust laws”).<sup>112</sup>

The key factor in allowing the antitrust case to proceed, as the *ABS* court said, was the defendant’s “relatively recent, aggressive patent acquisitions” that led to the patent litigation against the antitrust counterclaimant ABS.<sup>113</sup> This raised the possibility that ABS would be liable under the antitrust laws, provided that factual proof at trial showed that their patent acquisitions “reflect ST’s intent to maintain monopoly power through anticompetitive means.”<sup>114</sup>

The *ABS* case was premised on § 2 of the Sherman Act, but other challenges to patent acquisitions are brought under the Clayton Act’s § 7 prohibition on acquiring “assets” where “the effect of such acquisition may be substantially to lessen competition, or to tend to create a monopoly.”<sup>115</sup>

While not all antitrust challenges succeed,<sup>116</sup> the threat of scrutiny and the possibility of treble damages for successful antitrust plaintiffs may decrease

---

112. *Id.* at \*18.

113. *Id.*

114. *Id.* at \*19.

115. 15 U.S.C. § 18. Patents are a type of asset, so patent acquisitions are included in this provision. See *SCM Corp. v. Xerox Corp.*, 645 F.2d 1195, 1210 (2d Cir. 1981) (“Since a patent is a form of property . . . and thus an asset, there seems little reason to exempt patent acquisitions from scrutiny under [Section 7].”); *Crucible, Inc. v. Stora Kopparbergs Bergslags AB*, 701 F. Supp. 1157, 1162 (W.D. Pa. 1988) (“A patent, as a form of property, is an asset and not exempt from scrutiny under Section 7.”); *Dole Valve Co. v. Perfection Bar Equip., Inc.*, 311 F. Supp. 459, 463 (N.D. Ill. 1970) (“Of course, a patent may be ‘any part of the assets of another [person]’ within the meaning of Section 7.”).

116. See *Eastman Kodak Co. v. Goodyear Tire & Rubber Co.*, 114 F.3d 1547 (Fed. Cir. 1997), *abrogated on other grounds by* *Cybor Corp. v. FAS Techs., Inc.*, 138 F.3d 1448 (Fed. Cir.



the incidence of acquisitions or on their value. An acquiring firm that cannot use a patent against a rival will pay less for that patent. This does not support a complete rejection of all antitrust enforcement actions based on acquired patents. The health of the selling firm and its future innovative prospects should instead be part of the process for assessing the overall competitive situation that follows in the wake of the acquisition.

Admittedly, this policy may seem counter-intuitive; it permits the Big Platforms and other large companies to possibly acquire some degree of market power in the name of preserving speculative long-term benefits.<sup>117</sup> Patent acquisitions today, which can have an immediate impact on pricing and consumer welfare, are balanced against the maintenance of an ecosystem that includes some smaller potential innovators.<sup>118</sup> Theory, history, and

---

1998) (holding that a competitor failed to show the threatened market injury from the defendant's acquisition of an allegedly key patent required to support its Section 7 claim); *see also* *Intellectual Ventures I LLC v. Capital One Fin. Corp.*, 280 F. Supp. 3d 691, 703 (D. Md. 2017) (noting the extreme difficulty of determining the relevant market in a case where the "product" is a large bundle of related patents).

117. Which is why the literature on antitrust and patent acquisitions leans heavily toward the view that acquisitions present mostly problems, and not opportunities, with respect to product market and R&D competition. *See, e.g.*, Alan Devlin, *Antitrust Limits on Targeted Patent Aggregation*, 67 FLA. L. REV. 775, 776 (2015) ("[A]ntitrust law can viably limit certain abuses of the patent system by PAEs. Section 2 of the Sherman Act proscribes monopolization and Section 7 of the Clayton Act prohibits asset acquisitions that may substantially lessen competition or tend to create a monopoly. These provisions have sufficient teeth theoretically to catch the most egregious forms of hold-up founded on ex post patent aggregation and assertion. This Article explains how PAE activity can reduce social welfare and how PAEs' targeted patent aggregation and assertion may violate competition rules."); *see also* Eric Young, *A Bridge over the Patent Trolls: Using Antitrust Laws to Rein in Patent Aggregators*, 68 HASTINGS L.J. 203, 224 (2016) (warning of potential antitrust liability where a patent aggregator has acquired 100% or some other hefty market share of a certain technology standard, through its acquisition of industry standard patents). *But see* *Intellectual Ventures I LLC v. Capital One Fin. Corp.*, 280 F. Supp. 3d 691 (D. Md. 2017) (finding a failed attempt to plead antitrust liability on the part of patent aggregator Intellectual Ventures for asserting in litigation patents acquired from disparate sources and bundled into single licensing program).

118. The key is to understand the effect the acquisition will have on future innovation potential in the relevant industry. *Cf.* Erik Hovenkamp & Herbert Hovenkamp, *Buying Monopoly: Antitrust Limits on Damages for Externally Acquired Patents*, 25 TEX. INTEL. PROP. L.J. 39, 40 (2017) ("We propose that infringement damages for an externally acquired patent be denied if the acquisition served materially to expand or perpetuate the plaintiff's dominant position in the relevant technology market. By weakening enforcement, this limits the patent holder's ability to use such acquisitions to anticompetitive ends. We do not suggest that a dominant patent holder should be prohibited from securing external patent rights in the relevant technology market, but simply that its acquisition be limited to a nonexclusive license. This will permit the acquirer to practice the patent and keep its own technology up to date, but will not enable it to restrict third party access. This is as valuable to patent policy

empirical evidence nevertheless all support this policy. Big Platforms are by their nature very powerful in the short term; acquisition of some extra degree of market power through patent purchases will not change this much. Meanwhile, preserving some small firms could turn out to be enormously important for innovation in the long term. This raises the question of precisely how future innovation potential should factor in.

a) Towards a Consideration of Potential Future Innovation

Patent acquisitions have triggered antitrust scrutiny in several cases. Liability for an antitrust violation has been imposed when a firm with a strong market presence acquires patents that add to its anticompetitive economic power.<sup>119</sup> This Article proposes that antitrust regulators add a new dimension to their investigation of these acquisitions: the competitive survival of the *selling firm*.

There are two ways the survival of the seller might be incorporated into this analysis. First, it might be considered a potential *future* “disruptive firm,” a concept named in the authoritative Department of Justice (DOJ) and FTC Horizontal Merger Guidelines (“Merger Guidelines”)<sup>120</sup> as relevant in merger regulation. Alternatively, the contribution of the seller’s patents to the buying firm’s market power might be discounted or partially offset where patent sales are an important element of the selling firm’s continuing viability.

i) Preserving a Future Disruptor

The Merger Guidelines say that disruptive firms can make a valuable contribution to the competitive landscape:

The Agencies [DOJ and FTC] consider whether a merger may lessen competition by eliminating a “maverick” firm, i.e., a firm that plays a disruptive role in the market to the benefit of customers. For example, if one of the merging firms has a strong

---

as it is to antitrust, for it will tend to increase innovation by discouraging systematic monopoly in technology markets.”). My proposal is inconsistent with the Hovenkamp’s proposal to limit patent damages when patents are acquired—so long as one properly understands their test, whether the acquisition “served materially to expand or perpetuate the plaintiff’s dominant position in the relevant technology market.” An acquisition may contribute some market power in the short run while helping prevent the expansion or perpetuation of monopoly power in the long run.

119. See, e.g., *Kobe, Inc. v. Dempsey Pump Co.*, 198 F.2d 416, 423–25 (10th Cir.), *cert. denied*, 344 U.S. 837 (1952) (holding that acquiring a portfolio of patents to “corner the hydraulic pump business for oil wells” constituted illegal monopolization).

120. U.S. DEP’T OF JUSTICE & FED. TRADE COMM’N, HORIZONTAL MERGER GUIDELINES (2010), <https://www.ftc.gov/sites/default/files/attachments/merger-review/100819hmg.pdf> [<https://perma.cc/9CDP-SRRX>] [hereinafter *Merger Guidelines*].

incumbency position and the other merging firm threatens to disrupt market conditions with a new technology or business model, their merger can involve the loss of actual or potential competition. Likewise, one of the merging firms may have the incentive to take the lead in price cutting or other competitive conduct or to resist increases in industry prices. . . .<sup>121</sup>

In conventional merger analysis, “mavericks” or disruptors are preserved by refusing to approve the merger of a maverick and another firm if the merger would significantly increase market concentration.<sup>122</sup> Because they are especially important for preserving competition, a disruptor might not be allowed to merge, even if a non-disruptive firm with the same market share would be.<sup>123</sup>

In an antitrust review where “the competitive significance of one of the merging firms is declining,” the Merger Guidelines count this as a factor favoring the merger. Several antitrust cases invoke the strongest form of this principle, the “failing firm defense.” A dominant acquiring company can argue that a merger does no harm because the acquired firm is failing anyway. By this line of thinking, competitive conditions after the two firms combine cannot be any worse because the failing firm is leaving the market either way. In such a case, as the Merger Guidelines state, “the projected market share and significance of the exiting firm is zero.”<sup>124</sup>

In its current form, the failing firm defense is quite narrow.<sup>125</sup> Invoking it requires that the acquiring firm show significant business losses on the part of the acquired firm, with no immediate prospects that it can turn things around. Antitrust authorities know that if this defense is too readily accepted, it could serve as a cover for a large number of anticompetitive mergers; the ability to assume away the market share of one of the merging firms is a

---

121. *Id.* at § 2.1.5.

122. For details on how market shares and industry concentrations are calculated, see *id.* at § 5.3 (describing, among other things, use of the standard Herfindahl-Hirschman Index (HHI) to calculate market concentration).

123. See, e.g., *United States v. H & R Block*, 833 F. Supp. 2d 36, 79–80 (D.D.C. 2011) (analyzing whether the acquisition target was a “maverick” competitor in the market and finding that it was based on its role as a firm that “constrains prices”); see generally Jonathan B. Baker, *Mavericks, Mergers, and Exclusion: Proving Coordinated Competitive Effects Under the Antitrust Laws*, 77 N.Y.U. L. REV. 135 (2002); Courtney D. Lang, *The Maverick Theory: Creating Turbulence for Mergers*, 59 ST. LOUIS U. L.J. 257 (2014).

124. *Merger Guidelines*, *supra* note 120, at 32.

125. *United States v. Greater Buffalo Press, Inc.*, 402 U.S. 549, 555 (1971) (finding that failing firm defense is “narrow in scope”); cf. *Int’l Shoe Co. v. FTC*, 280 U.S. 291, 303 (1930) (holding that a company’s acquisition of a competitor does not violate Section 7 of the Clayton Act where the target’s resources are “so depleted and the prospect of rehabilitation so remote that it faced the grave probability of a business failure”).

considerable plus in merger analysis. They prevent abuse of the defense by requiring stringent proof of imminent failure.

When a small firm sells patents to a dominant company, particularly a platform company, courts should permit a new variant of the failing firm defense called the “declining significance” defense. In a winner-takes-all market, all who compete with the winning platform are by definition of declining significance. This is not the fault of these firms, but rather an unavoidable feature of platform markets. The idea this Article is championing here would apply when antitrust reviewers are looking over a purchase of small firm patents by a large platform firm. Antitrust authorities in these cases should discount or even factor out the market share contribution of a small firm’s patents. This makes it more likely that the patent acquisition will be approved, even if it enhances the market power of the large platform firm acquiring the patents. Whatever consumer harm might stem from such an enhancement is offset by the survival of the small firm into the future. If survival requires selling patents, these patent sales should be looked at favorably. A small increase in market power today is less disastrous than complete elimination of a possible innovator for tomorrow.

Where the firm selling the patents is relatively small but historically and potentially innovative, the survival of the selling firm ought to be considered. Where the acquiring firm appears to be gaining some degree of market power, an offsetting consideration would therefore enter the picture—the contribution the sale makes to the survival of the relatively small firm selling the patents. This is not a factor in the current analysis, which instead emphasizes the prospects for innovation by the acquiring firm and its downstream product competitors, which is likely appropriate in most cases.<sup>126</sup> If, for example, a patent portfolio permits an acquiring firm to raise costs for its rivals in important product markets, this may in theory detract from the rivals’ investments in future research. By diverting some of the rivals’ profits from their internal operations (such as R&D) to the acquiring firm (via patent infringement liability or licensing in the shadow of it), the firm

---

126. See Richard J. Gilbert & Steven C. Sunshine, *Incorporating Dynamic Efficiency Concerns in Merger Analysis: The Use of Innovation Markets*, 63 ANTITRUST L.J. 569, 570 (1995) (“A reduction in innovation may delay improvements in production processes that would lower the production costs of each of the merging firms, or it may reduce the magnitude of such improvements. In addition, a reduction in innovation may reduce the likelihood of discovery or delay the introduction by each firm of new or improved products. The loss of production improvements would result in higher costs, and possibly higher prices, even in markets where only one of the merging firms is a participant. Similarly, the loss of new or improved products would deny consumers the benefits of these improvements in every market where the firm is a supplier, including markets where only one of the firms is a participant.”).

acquiring the patents may impact future R&D in the industry. This much is conventional and usually correct; the process simply needs to remember the seller of the patent as well. Patent sales might be an important part of a firm's survival strategy, and to survive is to preserve the potential to fight for future innovations on some future day.

b) Patent Markets and the Future Competitive Landscape

Even while largely emphasizing the negative welfare effects of patent acquisitions, the most sophisticated antitrust analysts also recognize potential complexities. They see the possibility of positive effects. Fiona M. Scott Morton and Carl Shapiro, for example, note that:

[P]atent acquisitions by [Patent Assertion Entities], a central element of their monetization strategy, often discourage innovation and harm consumers. However, the analysis in this . . . article is rather general. We have not distinguished here between different types of patent portfolios, sellers, or buyers. When a given transaction is evaluated in practice, these particulars will rightly receive close attention. . . . As usual when patents are involved, we need to look at upstream technology markets (the markets where these patents are licensed) and at downstream product markets (the markets for products using the patented technology). Ultimately, we are interested in the impact of strategic patent acquisitions on downstream product prices, variety, and innovation.<sup>127</sup>

Although there are hints that the analysis proposed in this Section might fit within contemporary antitrust guidelines, they unfortunately are only hints. Consider the 2017 FTC/DOJ Licensing Guidelines ("Licensing Guidelines"). These Licensing Guidelines on their face apply only to the analysis of IP licensing, and even then, only to determine whether a licensing term is anticompetitive. They are not aimed at the problem of patent sales and are instead generally geared to the traditional concern of antitrust law—enhancing consumer welfare. In the context of patent licensing, this typically takes the form of protecting against the use of patent agreements to reduce competition in a market. The Licensing Guidelines protect two different types of markets: product markets and R&D (or "innovation") markets:

[A] licensing arrangement could include restraints that adversely affect competition in goods markets by dividing the markets among firms that would have competed using different technologies. An arrangement that effectively merges the activities of two actual or potential competitors in research and development

---

127. Fiona M. Scott Morton & Carl Shapiro, *Strategic Patent Acquisitions*, 79 ANT. L.J. 463, 484, 486 (2014) (footnote omitted).

in the relevant field might harm competition for development of new goods and services.<sup>128</sup>

Nevertheless, the Licensing Guidelines do shed some light on the way antitrust authorities look at future R&D potential as a factor in antitrust analysis. Section 3.2.3 of the Licensing Guidelines contains this discussion of R&D (or “innovation”) markets:

A research and development market consists of the assets comprising research and development related to the identification of a commercializable product, or directed to particular new or improved goods or processes, and the close substitutes for that research and development. When research and development is directed to particular new or improved goods or processes, the close substitutes may include research and development efforts, technologies, and goods that *significantly constrain the exercise of market power with respect to the relevant research and development*, for example by limiting the ability and incentive of a hypothetical monopolist to reduce the pace of research and development. The Agencies will delineate a research and development market *only when the capabilities to engage in the relevant research and development can be associated with specialized assets or characteristics* of specific firms.<sup>129</sup>

The highlighted phrases indicate the potential to include future R&D capacity in antitrust analysis, but they also illustrate the problems with such an approach. It may be impossible to say whether a small research-oriented company “significantly constrain[s]” market power in a given area of research; proof at this level may be asking for too much. The future R&D potential of a company may consequently be deemed too speculative to consider; as this Article argues, that would be a mistake. On the other hand, the requirement that R&D capabilities be associated with “specialized assets or characteristics” of specific firms seems consistent with the argument in this Article. One “characteristic” of a small firm that seems relevant is a track record of consistent creativity and innovation. If this “characteristic” counts as a positive in the analysis of R&D markets, the fact that the small firm so characterized will survive longer if it can sell patents may well be relevant. Consider, too, the helpful ideas in the following passage, also from the Licensing Guidelines:

---

128. U.S. DEP’T OF JUSTICE & FED. TRADE COMM’N, ANTITRUST GUIDELINES FOR THE LICENSING OF INTELLECTUAL PROPERTY § 3.1 (2017), <https://www.justice.gov/atr/IPguidelines/download> [<https://perma.cc/NP55-2XFV>] (footnotes omitted) [hereinafter *Antitrust Licensing Guidelines*].

129. *Id.* at § 3.2.3 (emphasis added) (footnote omitted).

In assessing the competitive significance of current and potential participants in a research and development market, the Agencies *will take into account all relevant evidence*. When market share data are available and accurately reflect the competitive significance of market participants, the Agencies will include market share data in this assessment. The Agencies also will seek *evidence of buyers' and market participants' assessments of the competitive significance of research and development market participants*.<sup>130</sup>

This passage serves a different purpose than the one this Article discusses, but some of the listed factors are relevant. These Licensing Guidelines analyze when a restrictive licensing arrangement might have significant anticompetitive effects. Third-party company research capabilities may bear on whether a restrictive license agreement between two parties is anticompetitive. Viable third-party-research capacity might constrain the market power of the parties to the license. Nevertheless, the spirit of the analysis is helpful. If “all relevant evidence” of the “significance of [R&D] market participants” is important for the licensing analysis, it should also be employed when a company sells patents. This Article has established why the future innovative capacity of a patent-selling firm is part of this “relevant evidence.” It has also argued that the *continued presence* of a participant in the R&D market is of chief importance, and that its survival ensures the possibility of future “competitive significance.” Essentially, the Licensing Guidelines show that the continued viability of a patent-selling firm should factor into the antitrust analysis of patent acquisitions.

## 2. *Smoothing the Patent Market*

Adjusting antitrust law can only have so much effect; there are at least two other policy changes that would facilitate patent markets. The first is a slight amendment in the patent recording statute which would make patent transfers a little more transparent. The second is a modification to the rules regarding administrative patent validity proceedings, which would allow an assignee to continue to defend a patent after an assignment rather than requiring a more expensive re-start of the proceeding.

### a) Recording of Patent Assignments, Licenses, and Other Interests

With respect to market-making, one of the most helpful features of the patent system is the patent assignment registry.<sup>131</sup> This searchable database

---

130. *Antitrust Licensing Guidelines*, *supra* note 128, at § 3.2.3 (emphasis added).

131. *Patent Assignment Search*, U.S. PAT. & TRADEMARK OFF. (last visited Dec. 21, 2019), <https://assignment.uspto.gov/patent/index.html#/patent/search> [<https://perma.cc/LMP7-YUW9>] (last visited Dec. 21, 2019).

allows patent-related transactions to be recorded and memorialized. When parties to a transaction use it, the database permits any member of the public to identify the current owner of a patent. Its most important function beyond this is as a registry that allows business people to record all manner of patent-related transactions—such as patent licenses or use of a patent as collateral for a loan.

The wording of the patent recordation statute, 35 U.S.C. § 261, provides a strong incentive to record ownership transfers. It says that “an assignment, grant or conveyance” shall be void, as against a later transferee, *unless it is recorded* in the Patent Office.<sup>132</sup> Patent recordation therefore protects an assignee against later transfers of the same patent. Although there are scenarios where an assignee can defeat a later assignment even without recordation,<sup>133</sup> recording is the safest and easiest way to protect an ownership interest in a patent.

By convention, people involved with the patent system also often record other patent-related transactions; the Patent Office will accept records of any patent-related transfer for recordation.<sup>134</sup> Thus licenses, mortgages, security interests, etc., are often recorded.<sup>135</sup> While this is advantageous, the incentive to record these interests is not as great as the incentive to record an actual assignment. Recording these other interests, as opposed to ownership transfers, does not automatically cut off the rights of subsequent transferees.<sup>136</sup> So licenses, mortgages, etc., are treated differently from assignments.<sup>137</sup> One possible improvement would be to broaden the

---

132. 35 U.S.C. § 261.

133. *See, e.g.,* Stanford Univ. v. Roche Molecular Sys., Inc., 583 F.3d 832, 843 (2009) (holding that “without notice” under § 261 can include constructive or inquiry notice, in addition to actual notice).

134. The Patent Office notes that it “does not verify the validity of the information [submitted for recordation]. Recordation is a ministerial function—the USPTO neither makes a determination of the legality of the transaction nor the right of the submitting party to take the action.” *Id.*

135. *See* MPEP (9th ed. Rev. 3, Jan. 2018) § 313 (“In addition to documents that constitute a transfer or change of title, other documents relating to interests in patents or applications will generally be recorded. Typical of these documents which are accepted for recording are license agreements and agreements which convey a security interest. Such documents are recorded in the public interest in order to give third parties notification of equitable interests or other matters relevant to the ownership of a patent or application.”).

136. *See, e.g., In re Cybernetic Servs., Inc.*, 252 F.3d 1039, 1052 (9th Cir. 2001) (arguing that § 261 concerns itself with only ownership rights, as opposed to lesser rights such as liens or licenses).

137. There is authority to support the idea that a license follows along with a patent after the patent has been assigned, regardless of whether the license is recorded in the Patent Office. *See, e.g.,* Innovus Prime, L.L.C. v. Panasonic Corp., No. C-12-00660-RMW, 2013 WL



recording statute and provide stronger incentives to record all patent-related transactions. This can be accomplished simply by changing the wording of the recording statute to “the transfer of any interest relating to a patent shall be void as against any subsequent transferee for a valuable consideration, without notice . . . .” This would place all patent interests on the same footing as assignments or other conveyances, which in turn would create strong incentives to record all patent-related transactions in the recording database.

b) Facilitating Transfers When Patents are Being Challenged in the Patent Office

The America Invents Act of 2011 created a way to challenge patent validity without paying for expensive federal litigation;<sup>138</sup> a competitor or interested party can test validity in an administrative case at the Patent Office.<sup>139</sup> The most popular form of challenge is an Inter Partes Review

---

3354390 (N.D. Cal. 2013) (granting accused infringer summary judgment that plaintiff as a fourth generation assignee of the patent had to honor a covenant not to sue that was granted to the accused infringer by the original patentee, and that the subsequent assignments of the patent did not operate to nullify the covenant not to sue, even if the later assignees were not aware of the covenant) (“Assignment transfers assignor’s contract rights . . . . This occurs whether or not an assignee had notice.”); *Keystone Type Foundry v. Fastpress Co.*, 272 F. 242, 245 (2d Cir. 1921); *see also* *L.L. Brown Paper Co. v. Hydroiloid, Inc.*, 118 F.2d 674, 677 (2d Cir. 1941) (“The assignee of a patent taking title subsequent to the granting of a license under patent receives no more than the former owner’s interest, including the usual rights of a patent owner diminished by the licensee’s right to use the patented process within scope of its license.”); *Jones v. Berger*, 58 F. 1006, 1007 (C.C.D. Md. 1893) (citing WILLIAM C. ROBINSON, *THE LAW OF PATENTS FOR USEFUL INVENTIONS* § 817 (1890)); *Sanofi, S.A. v. Med-Tech Veterinarian Prods.*, 565 F. Supp. 931, 940–41 (D.N.J. 1983) (“Because the purchaser [of patented products] is under an obligation to inquire of the seller as to the existence of any outstanding licenses, the purchaser cannot claim that his expectations have been frustrated if he fails to make the necessary inquiry.”); Andrew C. Michaels, *Patent Transfer and the Bundle of Rights*, 83 BROOK. L. REV. 933, 937–38 (2018) (“The courts have ruled that even a “bona fide purchaser” of a patent takes the patent subject to prior “licenses, of which he must inform himself as best he can at his own risk” The intuition seems to be that the purchaser of a patent should recognize the possibility that licenses on the patent might exist, and should take steps to investigate whether they in fact do exist. In other words, the purchaser is on “inquiry notice” with regard to the potential existence of license agreements affecting the patents to be transferred. Of course, a true bona fide purchaser patent assignee may have some claim sounding in tort or contract against a patent seller who is less than forthright about the extent to which the patent has been licensed, particularly where the license or its terms are not public knowledge.”) (citing *Innovus Prime*, 2013 WL 3354390, at \*15).

138. *See* ROBERT P. MERGES & JOHN F. DUFFY, *PATENT LAW AND POLICY: CASES AND MATERIALS* 931–47 (7th ed. 2017).

139. For an explanation of why administrative challenges are appealing to challengers, *see* Joseph Farrell & Robert P. Merges, *Incentives to Challenge and Defend Patents: Why Litigation*

(IPR). While most patents in IPR proceedings are also being litigated in court,<sup>140</sup> it has become common to use an IPR (or the threat of one) in all sorts of patent-related negotiations—including negotiations over the sale or license of a patent or patent portfolio.<sup>141</sup>

The problem arises when a patent changes hands while under challenge in an IPR. Current rules do not create a smooth transition between owners; they do not allow for a new owner to step into the shoes of the old one. In fact, there is no provision at all for the replacement of a party to an IPR in the middle of a proceeding.<sup>142</sup> The old owner could settle its case with the patent challenger and be released, but then the new owner and challenger might have to start over or at least duplicate some of the costs that the old owner had already sunk into the IPR.<sup>143</sup>

The obvious solution is to implement a simple party-substitution procedure. This new procedure would allow a new owner to step into the shoes of the old, provided that they are willing to be bound by stipulations

---

*Won't Reliably Fix Patent Office Errors and Why Administrative Patent Review Might Help*, 19 BERKELEY TECH. L.J. 943 (2004).

140. Saurabh Vishnubhakat et al., *Strategic Decision Making in Dual PTAB and District Court Proceedings*, 31 BERKELEY TECH. L.J. 45 (2016) (finding that 70% of instituted IPRs are brought by parties also involved in district court litigation and showing that IPRs are working as an effective substitute for district court litigation to invalidate patents).

141. See Jake Berdine & Matt Rosenberg, *Creating Leverage: A Practitioner's Guide to Inter Partes Review and Its Effects on Intellectual Property License Negotiations*, 44 AIPLA Q.J. 1, 2 (2016).

142. See *Librestream Techs., Inc., v. Wireless Remote Sys. L.L.C.*, No. IPR2014-00369, 2014 WL 5080112 (P.T.A.B. Oct. 20, 2014) (October 10, 2014) (“We advised [attorney for the old patent owner] that withdrawal may occur only as permitted under our rules. Our rules require our authorization prior to seeking withdrawal. See 37 C.F.R. § 42.10(e). Further, until withdrawal is granted, Mr. Moreland and any other attorneys designated as counsel for Patent Owner under 37 C.F.R. § 42.8(b)(3), are attorneys of record for Patent Owner. Current counsel will remain of record until new lead and backup counsel are identified by an appropriate power of attorney. See 37 CFR § 42.10(b). Whether or not the patents at issue have been assigned to a new party is not of record. Regardless, counsel and Patent Owner are advised that the AIA does not provide for the “replacement” of a party. Changes to Implement Inter Partes Review Proceedings, Post-Grant Review Proceedings, and Transitional Program for Covered Business Methods, 77 Fed. Reg. 48680, 48707 (Aug. 14, 2012).”).

143. See Christina Schwarz & Raymond Mandra, *US Patents: Beware Assigning Patents in IPR Proceedings*, MANAGING INTELL. PROP. (Dec. 11, 2014), <http://www.managingip.com/Article/3409566/US-patents-Beware-assigning-patents-in-IPR-proceedings.html> [https://perma.cc/PHA8-5F88] (“Until the Board provides further clarity, prospective patent assignees should proceed cautiously when considering assignment of patents involved in IPR proceedings. At a minimum, it should be assumed the assignor will remain the named patent owner in the IPR and thus a potential assignee should seek an agreement that provides control of the IPR proceeding and cooperation by the former patent owner. It may also be advisable to seek guidance from the Board prior to finalizing any transfer of patent rights.”).

and findings made while the proceeding was under the direction of the old owner. This would facilitate efficient challenges—a primary aim of the IPR process—in the increasingly likely scenario where the challenged patent is sold mid-stream. It would be very useful, for example, where a challenged patent is one of many that is part of a portfolio being sold. This simple procedural fix would ensure that one or a handful of patents under IPR challenge do not threaten the sale of a substantial patent portfolio.

## V. CONCLUSION

Whether all innovation in the future will emanate from a handful of massively integrated firms is hard to predict. Even a supporter of growth by acquisition and of today's entrepreneurial exits via the "acqui-hire" will normally acknowledge the benefits of small, independent outsiders. Big may be better in many minds, but it's not uniformly thought of as permanently Best.

This is important because it is wise to be wary of the thought which has so often crept into consciousness during a major technological realignment—that *this time is different*, that *this time we have it figured out*. It is inevitable that this thought will pop up, yet essential that it be resisted.

At least a few defenders of today's Big Platform companies will claim that the care and feeding of small, independent outsider firms is no longer essential because they have become obsolete. This is certain because people have repeatedly claimed this before, despite always being proven incorrect. To take one example of many that could be selected, consider this:

As organized invention and discovery gain momentum the revolutionist will have no chance . . . . He will have to compete with more and more [people] who have at their disposal splendidly equipped laboratories, time, and money, and who may work for three or four years before producing a noteworthy result . . . . Possibly Edison may be the last of the great heroes of invention.<sup>144</sup>

That was written in 1930. Meanwhile, despite this belief, outsider Philo Farnsworth was inventing the television;<sup>145</sup> Scotch tape was being invented; the frozen food process was being perfected, etc.<sup>146</sup>

---

144. WALDEMAR KAEMPFERT, INVENTION AND SOCIETY 30 (1930).

145. See U.S. Patent No. 2,087,633 (filed Apr. 26, 1933) (issued July 20, 1937).

146. See Mary Bellis, *Twentieth Century Timeline: Technology, Science and Inventions*, THOUGHTCO. (Sept. 4, 2019), <https://www.thoughtco.com/20th-century-timeline-1992486> [<https://perma.cc/4JP6-U33D>].

The reason people make this mistake over and over is that each glittery, new innovation system that comes to prominence really is impressive. The organized industrial research labs of the 1930s looked like nothing the world had seen before. They even produced excellent results for many years for companies such as General Electric and DuPont. As successful as they were, however, they were simply the Latest Word in the long march of new ideas. The mistake some people made and continue to make is to think that they were the Last Word. That Word has not been written yet, and with luck it may never be.

For now, the glittery success of Big Platform companies and their companions of the moment appears to be sweeping all before them in a great conquest of digital-era innovation. Nevertheless, it would be very wise for society to place a few side bets and hedge against the future. It makes sense to keep the avenues open for something new and different—something from out of left field. If the patent market can help in that respect, then it behooves society to keep that market open. The fact that it is associated at times with litigation and attempts at rent seeking ought not to exert too much influence. If the patent market provides a profitable outlet and allows some small companies to remain independent, it may prove quite useful in the long run. The phrase “history teaches” has a tired ring to it, but it might be accurate in this case. Multiple rivalrous and independent sources of innovation have always been a good thing. It seems safe to bet that they still are and will be in the future, too.

# ARTIFICIAL INTELLIGENCE OPINION LIABILITY

*Yavar Bathaee*<sup>†</sup>

## ABSTRACT

Opinions are not simply a collection of factual statements—they are something more. They are models of reality that are based on probabilistic judgments, experience, and a complex weighting of information. That is why most liability regimes that address opinion statements apply scienter-like heuristics to determine whether liability is appropriate, for example, holding a speaker liable only if there is evidence that the speaker did not subjectively believe in his or her own opinion. In the case of artificial intelligence, scienter is problematic. Using machine-learning algorithms, such as deep neural networks, these artificial intelligence systems are capable of making intuitive and experiential judgments just as humans experts do, but their capabilities come at the price of transparency. Because of the Black Box Problem, it may be impossible to determine what facts or parameters an artificial intelligence system found important in its decision making or in reaching its opinions. This means that one cannot simply examine the artificial intelligence to determine the intent of the person that created or deployed it. This decouples intent from the opinion, and renders scienter-based heuristics inert, functionally insulating both artificial intelligence and artificial intelligence-assisted opinions from liability in a wide range of contexts. This Article proposes a more precise set of factual heuristics that address how much supervision and deference the artificial intelligence receives, the training, validation, and testing of the artificial intelligence, and the a priori constraints imposed on the artificial intelligence. This Article argues that although these heuristics may indicate that the creator or user of the artificial intelligence acted with scienter (i.e., recklessness), scienter should be merely sufficient, not necessary for liability. This Article also discusses other contexts, such as data bias in training data, that should also give rise to liability, even if there is no scienter and none of the granular factual heuristics suggest that liability is appropriate.

---

DOI: <https://doi.org/10.15779/Z38P55DH32>

© 2020 Yavar Bathaee.

<sup>†</sup> Litigator and computer scientist. This Article is dedicated to my wife, Jacqueline, and my children, Elliot and Audrey. I would like to thank James Steiner-Dillon for his comments on the Article and his support. All errors and omissions are my own.

## TABLE OF CONTENTS

<b>I.</b>	<b>INTRODUCTION .....</b>	<b>115</b>
<b>II.</b>	<b>OPINION LIABILITY, THE SCIENTER HEURISTIC, AND INFORMATION ASYMMETRY .....</b>	<b>120</b>
A.	THE OPINION/FACT DISTINCTION .....	120
B.	SCIENTER AND OPINION LIABILITY .....	125
C.	EXAGGERATED OPINIONS AND PUFFERY .....	127
D.	OMISSIONS AND INFORMATION ASYMMETRY .....	129
E.	OPINION STATEMENTS AS MODELS .....	133
<b>III.</b>	<b>ARTIFICIAL INTELLIGENCE, THE BLACK BOX PROBLEM, AND OPINION STATEMENTS .....</b>	<b>138</b>
A.	WHAT IS ARTIFICIAL INTELLIGENCE? .....	138
B.	THE BLACK BOX PROBLEM .....	141
C.	AI OPINIONS .....	143
D.	THE FAILURE OF THE SCIENTER HEURISTIC .....	145
E.	AN OPAQUE BASIS AND MATERIALITY .....	149
F.	AI AS AN OPAQUE EXPERT .....	151
<b>IV.</b>	<b>A FRAMEWORK FOR AI OPINION LIABILITY.....</b>	<b>153</b>
A.	BETTER FACTUAL HEURISTICS FOR AI OPINION LIABILITY.....	153
1.	<i>Deference and Autonomy</i> .....	154
2.	<i>Training, Validation, and Testing</i> .....	155
3.	<i>Constraint Policies and Conscientiousness</i> .....	158
B.	EXAMINING DATA BIAS, NOT DECISION BASIS IN OMISSIONS CASES .....	159
C.	HIGH RISK / HIGH VALUE APPLICATIONS AND STRICT LIABILITY....	162
D.	WHY DISCLOSURE RULES ARE LESS EFFECTIVE IN THE CASE OF AI OPINIONS .....	166
1.	<i>Disclosure in the Non-AI Opinion Context</i> .....	166
2.	<i>Disclosure Will Be Less Effective in the AI Context</i> .....	167
E.	WHEN CAN YOU INFER USER OR CREATOR INTENT FROM AN AI MODEL'S OPINION? .....	168
F.	PUTTING IT ALL TOGETHER: SCIENTER SHOULD BE SUFFICIENT, NOT NECESSARY FOR OPINION LIABILITY.....	168
<b>V.</b>	<b>CONCLUSION.....</b>	<b>169</b>

## I. INTRODUCTION

Opinion statements are everywhere. They express judgments about things such as value,<sup>1</sup> probability,<sup>2</sup> or the appropriate course of action.<sup>3</sup> They are more than the facts underlying them; they are also the weights the person stating the opinion attaches to those facts. That is why opinion statements not only include factual statements, they also implicitly say something about the person expressing the opinion—namely, that the person stating the opinion has a basis for it, that they genuinely believe in the opinion, and that they are not aware of facts and reasons that would undermine the opinion.<sup>4</sup>

---

1. Statements about valuation are generally regarded as statements of opinion because, when there is no clear market price for an asset, the “fair value” of an asset “will vary depending on the particular methodology and assumptions used.” *Fait v. Regions Fin. Corp.*, 655 F.3d 105, 111 (2d Cir. 2011). Indeed, in many cases “[t]here may be a range of prices with reasonable claims to being fair market value.” *Henry v. Champlain Enters., Inc.*, 445 F.3d 610, 619 (2d Cir. 2006). Although much of the discussion of valuations as opinions have been in the securities law context, valuations have been treated as opinions in other fields of law, including contract. *See, e.g.*, RESTATEMENT (SECOND) OF CONTRACTS § 168 cmt. c (AM. LAW INST. 1981) (“A statement of value is, like one of quality, ordinarily a statement of opinion.”).

2. An opinion statement often carries with it the implicit statement that it encompasses a belief based on incomplete information or based on uncertain facts. Indeed, the Restatement of Contracts states that “[a]n assertion is one of opinion if it expresses only a belief, without certainty, as to the existence of a fact or expresses only a judgment as to quality, value, authenticity, or similar matters.” RESTATEMENT (SECOND) OF CONTRACTS § 168. Indeed, because opinions rest on the “weighing of competing facts,” it is generally understood that stating an opinion is a way of “conveying uncertainty.” *Omnicare, Inc. v. Laborers Dist. Council Constr. Indus. Pension Fund*, 135 S. Ct. 1318, 1329 (2015).

3. A recommendation or prognosis statement by an expert is a classic example of an opinion statement that may give rise to liability. Indeed, some of the earliest opinion liability cases in the United States concerned statements made by physicians about diagnosis and prognosis. *See, e.g.*, *Hedin v. Minneapolis Med. & Surgical Inst.*, 64 N.W. 158, 160 (Minn. 1895) (noting that the physician’s diagnosis came with it an opinion that “a representation that plaintiff’s physical condition was such as to insure a complete recovery”). When the opinion of an expert, such as a medical professional, is involved, liability has traditionally turned on whether the speaker’s role as an expert invited reliance on the opinion. *See, e.g.*, *Gagne v. Bertran*, 275 P.2d 15, 21 (Cal. 1954) (“Moreover, even if defendant’s statement was an opinion, plaintiffs justifiably relied thereon. Defendant held himself out as an expert, plaintiffs hired him to supply information concerning matters of which they were ignorant, and his unequivocal statement necessarily implied that he knew facts that justified his statement.”).

4. *See Omnicare*, 135 S. Ct. at 1334 (Scalia, J., concurring) (“In a few areas, the common law recognized the possibility that a listener could reasonably infer from an expression of opinion not only (1) that the speaker sincerely held it, and (2) that the speaker knew of no facts incompatible with the opinion, but also (3) that the speaker had a reasonable basis for holding the opinion.”); *see also* RESTATEMENT (SECOND) OF CONTRACTS § 168 (noting that an opinion comes with it the assertion that “the facts known to that person are not incompatible with his opinion,” or “that he knows facts sufficient to justify him in forming it”); *id.* § 168

The law has developed significant aptitude at evaluating the truth or falsity of factual statements based on evidence.<sup>5</sup> However, determining whether a speaker genuinely believes in their opinion will often require intent-based heuristics—the most notable of which is scienter.<sup>6</sup> Since opinion statements are not true or false merely because some fact the opinion is based upon proves to be true or false, these heuristics, which are described in Part II, are in many cases outcome-determinative on the question of liability.

The value of these intent-based heuristics will likely be aggressively challenged by a new breed of computer programs capable of forming and stating opinions—artificial intelligence (AI).<sup>7</sup> For the first time in human history, artificially intelligent computer programs are capable of rendering opinions without deterministic instructions.<sup>8</sup> They can learn from data—from experience—and come to intuitive conclusions without the aid of a human

---

cmt. a (“A statement of opinion is also a statement of fact because it . . . has a particular state of mind concerning the matter to which his opinion relates.”).

5. Indeed, the stated purposes of the Federal Rules of Evidence include “the end of ascertaining the truth.” FED. R. EVID. 102. Many of the rules themselves are addressed to determining the admissibility, relevance, and reliability of statements, the most notable of which is the hearsay rule and its exceptions. *See* FED. R. EVID. 801–802 (addressing the admissibility of statements, including out-of-court statements that are offered for their truth).

6. This is because the opinion carries with it the implicit statement that the opinion is genuinely believed by the speaker. Thus, proving the subjective falsity of the opinion is functionally the same as proving scienter. *See In re Credit Suisse First Bos. Corp.*, 431 F.3d 36, 48 (1st Cir. 2005) (“[T]he subjective aspect of the falsity requirement and the scienter requirement essentially merge; the scienter analysis is subsumed by the analysis of subjective falsity.”). Another useful heuristic is to determine whether the factual assumptions underlying an opinion hold true; if they do not, then the opinion itself is undermined because the speaker’s intent is called into question. *See Va. Bankshares v. Sandberg*, 501 U.S. 1083, 1093 (1991) (“Provable facts either furnish good reasons to make a conclusory commercial judgment, or they count against it, and expressions of such judgments can be uttered with knowledge of truth or falsity just like more definite statements, and defended or attacked through the orthodox evidentiary process that either substantiates their underlying justifications or tends to disprove their existence.”).

7. AI, as referred to in this Article, is a class of computer programs designed to solve problems that typically require “inferential reasoning [and/or] decision-making based on incomplete or uncertain information, classification, optimization, and perception.” Yavar Bathaee, *The Artificial Intelligence Black Box and The Failure of Intent and Causation*, 31 HARV. J.L. TECH. 889, 920 (2018).

8. Although some forms of AI do in fact rely on deterministic instructions, *see* Bathaee, *supra* note 7, at 898, the AI addressed in this Article generally are not deterministically programmed, but are instead trained from examples using machine-learning algorithms—that is, they are computer programs that learn directly from data. *See* ETHEM ALPAYDIN, INTRODUCTION TO MACHINE LEARNING xxv (2004) (“We need learning in cases where we cannot directly write a computer program to solve a given problem, but need example data or experience. One case where learning is necessary is when human expertise does not exist, or when humans are unable to explain their expertise.”).



being.<sup>9</sup> What then do intent-based heuristics achieve when the intent of the AI's creator or user does not necessarily affect or reflect the judgment or opinions of the AI? As this Article contends, very little.

As explained in Part III, this decoupling of the AI creator's intent from the AI's judgments arises from a technological problem that occurs when certain classes of machine-learning algorithms are used by AI—the Black Box Problem.<sup>10</sup> The black box problem arises where machine-learning algorithms rely on layers upon layers of linear and non-linear transformations, such as deep artificial neural networks. These algorithms are capable of learning from data and experience, just as humans do, but such powerful cognition comes at the price of transparency.<sup>11</sup> A trained neural network, for example, may have internalized hundreds of thousands, if not millions of data points, and may arrive at accurate predictions or sound opinions, but the complexity of the neural network may make it impossible to determine how the AI has made its judgments or reached an opinion.<sup>12</sup> Thus applying intent-based heuristics will almost never result in liability.<sup>13</sup>

Today, AI helps perform tasks that in the past have required human judgment and experience.<sup>14</sup> For example, AI can achieve higher accuracy at spotting certain forms of cancer—a task that in the past required a trained doctor with years of experience to perform.<sup>15</sup> The bread and butter of finance and accounting, valuation, will also soon be predominantly a task relegated to AI.<sup>16</sup> Even before the AI revolution, algorithmic valuation was a rapidly

---

9. See Bathaee, *supra* note 7, at 891. Because machine learning-based AI can learn directly from data instead of simply implementing rigid pre-programmed rules, it “can learn, adapt to changes in a problem’s environment, establish patterns in situations where rules are not known, and deal with fuzzy or incomplete information.” MICHAEL NEGNEVITSKY, *ARTIFICIAL INTELLIGENCE* 14 (2d ed. 2005).

10. See *infra* Section III.B.

11. See *infra* Section III.B & III.C.

12. For a detailed discussion of the AI Black Box Problem and how it arises from the use of certain machine-learning algorithms, see Bathaee, *supra* note 7, at 897–906.

13. See Bathaee, *supra* note 7, at 906–21.

14. See *infra* Sections III.B & III.C.

15. See, e.g., Andre Esteva et al., *Dermatologist-level Classification of Skin Cancer with Deep Neural Networks*, 542 *NATURE* 115 (2017); Martin Stumpe & Lily Peng, *Assisting Pathologists in Detecting Cancer with Deep Learning*, *GOOGLE RES. BLOG* (Mar. 3, 2017), <https://research.googleblog.com/2017/03/assisting-pathologists-in-detecting.html> [<https://perma.cc/2BMT-YCTX>] (“In fact, the prediction heatmaps produced by the algorithm had improved so much that the localization score (FROC) for the algorithm reached 89%, which significantly exceeded the score of 73% for a pathologist with no time constraint.”); see also Ahmed Hosny et al., *Artificial Intelligence in Radiology*, *NATURE REV. CANCER* (May 17, 2018).

16. See *infra* Section III.C.

growing field.<sup>17</sup> With the ability to build models that can accurately learn from vast amounts of data, the number of AI-based valuation systems is only expected to multiply. AI will also likely assist other specialized experts with judgments, including judges and arbitrators.<sup>18</sup>

Under the current prevailing standards for opinion liability, a court will find liability only based on the intent of the humans that stated the opinion.<sup>19</sup> But, when an AI opinion is involved, its decisions will be based on data, and the intent of the creators or users of the AI will generally not provide insight into the AI's decision-making process.<sup>20</sup> And since the AI may suffer from the Black Box Problem, it may not have an ascertainable intent that can be examined or queried.<sup>21</sup> The net effect of this will be the end of opinion liability in many fields of law that require some form of intent, such as scienter, because intent-less AI and AI-assisted opinions will be functionally immune.<sup>22</sup>

Part IV of this Article argues that the current opinion liability regime requires two significant adjustments. First, more precise heuristics—designed specifically for AI—are needed. That is, courts and factfinders should look to (i) the extent to which the AI model was given deference and autonomy, (ii) the manner in which the AI was trained, validated, and tested, and (iii) the extent to which a priori constraints were placed on the judgments of the AI system to mitigate known risks.<sup>23</sup>

It is possible that these heuristics point to recklessness on the part of the creator or user of the AI, and in such a case, there may be a permissible inference of scienter,<sup>24</sup> but as this Article explains, there may also be other

---

17. Indeed, automated valuation models, which were based on deterministic algorithms (not modern AI) were a prominent feature of the mortgage crisis of 2008. *See, e.g.,* Mass. Mut. Life Ins. Co. v. DB Structured Prods., 110 F. Supp. 3d 288, 293 (D. Mass. 2015) (discussing automated valuation models used for due diligence and appraisals of real property prior to the real estate crisis of 2008); Fed. Hous. Fin. Agency v. Nomura Holding Am., Inc., 60 F. Supp. 3d 479, 491–92 (noting that automated valuation models were used by government-sponsored entities, such as Fannie Mae, to assess values of homes underlying mortgage-backed securities they purchased).

18. In Wisconsin, for example, the state's Supreme Court recently ruled that the use of actuarial data to predict recidivism did not offend a defendant's due process rights, even though the data and methodology was not disclosed to the court or the defendant. *See* State v. Loomis, 88 N.W.2d 749 (Wis. 2016). For a full discussion of the case, see Case Comment, *Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing*: State v. Loomis, 130 HARV. L. REV. 1530, 1534 (2017).

19. *See infra* Section III.D.

20. *See* Bathace, *supra* note at 7, at 906–21.

21. *Id.*

22. *See infra* Section III.D.

23. *See infra* Section IV.A.

24. *See infra* Section IV.E.

circumstances that would warrant liability. For example, there may be some applications that would require a strict liability rule—those that involve high risks of harm or that implicate governmental or societal norms that would require human, not machine, judgment.<sup>25</sup> It may also be the case that a failure to detect significant bias in the data used to train the AI should itself warrant liability.<sup>26</sup>

In such cases, there may be no basis for an inference of scienter, including under the more precise heuristics proposed by this Article.<sup>27</sup> That does not, however, mean that liability for an opinion statement should not attach.<sup>28</sup> Accordingly, the second modification this Article proposes to the status quo is that scienter should be sufficient, not necessary for liability when AI is involved.<sup>29</sup> Most opinion liability regimes have it the other way around, requiring a showing of scienter for opinion liability—even in some cases where a statute does not require scienter for liability.<sup>30</sup> However, when AI is involved, requiring scienter will immunize a wide swath of conduct and provide a host of perverse incentives to use AI to shield opinions from liability.

With this new technology comes the promise of multiplying and perhaps exceeding human intelligence by orders of magnitude,<sup>31</sup> but with that comes the need to create new legal and factual heuristics designed for machines—not to make patchwork adjustments to legal doctrines designed to understand human conduct. Indeed, if the status quo would immunize almost all AI opinion from liability, there may be no occasion to make thoughtful and incremental adjustments to our legal doctrines.

---

25. *See infra* Section III.C.

26. *See infra* Section III.B.

27. *See infra* Section III.F.

28. *See id.*

29. *See infra* Section III.F.

30. *See infra* notes 247–48 and accompanying text.

31. As predicted for decades by commentators on AI, AI systems already exceed humans in perception-based tasks, such as vision. KURZWEIL, *THE AGE OF SPIRITUAL MACHINES* 65 (2000); Gina Smith, *Google Brain Chief: AI Tops Humans in Computer Vision, and Healthcare Will Never Be the Same*, SILICON ANGLE (Sept. 27, 2017), <https://siliconangle.com/2017/09/27/google-brain-chief-jeff-dean-ai-beats-humans-computer-vision-healthcare-will-never/> [<https://perma.cc/95AQ-8TYD>]. Experts predict that AI systems will exceed humans in tasks such as language translation and truck driving within the coming decade. *Experts Predict When Artificial Intelligence Will Exceed Human Performance*, MIT TECH. REV. (May 31, 2017), <https://www.technologyreview.com/s/607970/experts-predict-when-artificial-intelligence-will-exceed-human-performance/> [<https://perma.cc/Y4PA-YPTX>].

## II. OPINION LIABILITY, THE SCIENTER HEURISTIC, AND INFORMATION ASYMMETRY

This Part describes the unique challenges posed by opinion statements as well as some of the heuristics used to determine whether the speaker of an opinion should be held liable. This Part does not survey any particular area of law but instead attempts to describe how familiar heuristics, such as scienter and reliance, solve many of the problems posed by opinion statements. These problems include, for example, information asymmetry, contrary or incomplete information, and unreasonable or inadequate bases for the opinion. This Part concludes that opinions are factual models, which include not only a set of underlying facts, but also probability weights for those facts and notions, acquired through the speaker's experience.

### A. THE OPINION/FACT DISTINCTION

Factual statements are often at the center of legal disputes. Proving a factual statement true or false lies in finding empirical facts as they existed when the statement was made and comparing those facts to what was conveyed in the statement.<sup>32</sup>

The question of whether to impose liability based on a false statement, however, will not be a simple matter of determining what facts existed, were known, or were knowable when the statement was made. Instead, the question is often about the overall context of the statement and what the speaker intended to accomplish.<sup>33</sup> There are many battle-tested heuristics for dealing

---

32. In securities cases, falsity of a factual statement is often a necessary predicate for liability and can be pled or proven with evidence that the facts as they existed when the statement was made contradicted the factual statement. *See In re Homestore.com, Inc. Sec. Litig.*, 252 F. Supp. 2d 1018, 1032 (C.D. Cal. 2003) (noting that falsity can be pled where defendant is in "possession of non-public information that would prove his statements false"); *Plevy v. Haggerty*, 38 F. Supp. 2d 816, 826 (C.D. Cal. 1998) (noting that falsity can be pled by "direct or circumstantial facts, such as, but not limited to, inconsistent contemporaneous statements or internal reports, that would support [that the statements] . . . were false when made"). In other contexts, such as false statements under the Lanham Act, courts have focused on whether a statement of fact is "measurable" and "specific" enough to be proven to be false. *Franklin Fueling Sys. v. Veeder-Root Co.*, No. S-09-580 FCD/JFM, 2009 U.S. Dist. LEXIS 72953, at \*13 (E.D. Cal. Aug. 11, 2009); *see also, e.g., Hi-Tech Pharm., Inc. v. HBS Int'l Corp.*, 910 F.3d 1186, 1193 (11th Cir. 2018) (alleging that precise advertisement and representation of drink's percentage of protein content was sufficiently specific to be proven false by an alleged test showing a lower amount of protein in a Lanham Act claim). These courts suggest that what makes a statement of fact provably true or false is the specificity of the statement, the ability to measure the information conveyed in the statement, and the existence of consistent or inconsistent contemporaneous evidence.

33. *See, e.g., Tolles v. Republican-American*, No. UWYCV106005674, 2012 Conn. Super. LEXIS 2877, at \*9 (Super. Ct. Nov. 20, 2012) ("Connecticut law makes clear that in

with the host of issues that arise as part of the liability question, such as evaluating and comparing the credibility of witnesses,<sup>34</sup> examining the motives of the person making the statement (and in some cases of those that heard it),<sup>35</sup> and evaluating whether the statement was important enough to have affected a transaction or a decision-making process.<sup>36</sup>

---

determining the scope of the alleged statement, and further in determining its truth or falsity, context is important and sometimes even dispositive.”); *Buetow v. A.L.S. Enters.*, 650 F.3d 1178, 1185 (8th Cir. 2011) (“In assessing whether an advertisement is literally false, a court must analyze the message conveyed within its full context.”) (quoting *United Indus. v. Clorox Co.*, 140 F.3d 1175).

34. The Federal Rules of Evidence, for example, provide for the impeachment of witnesses precisely because credibility is a powerful heuristic for assessing whether the facts conveyed by the witness are true, including whether the witness’s testimony contradicts his own prior inconsistent statements. *See* FED. R. EVID. 613(b). The Federal Rules of Evidence accordingly treat out-of-court statements offered for impeachment as non-hearsay statements because they are not being offered for their truth. *See* *Hartford Fire Ins. Co. v. Taylor*, 903 F. Supp. 2d 623, 642 (N.D. Ill. 2012) (holding that out of court statement offered for impeachment was not hearsay).

35. Courts routinely consider a speaker’s motive to make a false statement. In fact, the motive to have made a false or misleading statement is an important part of the scienter inquiry required for most fraud-based claims. *See In re PXRE Grp., Ltd. Sec. Litig.*, 600 F. Supp. 2d 510, 531 (S.D.N.Y. 2009) (pleading securities fraud requires alleging facts indicating a “motive and opportunity probative of a strong inference of scienter”) (quoting *Rothman v. Gregor*, 220 F.3d 81, 90 (2d Cir. 2000)). Even in circuits where motive and opportunity are not sufficient for scienter, they are an important part of the analysis. *See In re Silicon Storage Tech.*, No. C 05-0295 PJH, 2006 U.S. Dist. LEXIS 14790, at \*50 (N.D. Cal. Mar. 10, 2006) (“In the Ninth Circuit, motive and opportunity, standing alone, are not sufficient to establish scienter . . . . However, motive can be considered as part of the ‘totality of the allegations’ regarding scienter.”) (internal citations omitted). What matters is that the alleged motive indicates a clear reason to make a false statement such that one can infer scienter. It will therefore not be enough to allege, for example, a speaker’s generalized motive to maximize profits or to justify management decisions, because all companies or businessmen have such a motive—not just those that make false statements. *See Zirkin v. Quanta Capital Holdings Ltd.*, No. 07 Civ. 851 (RPP), 2009 U.S. Dist. LEXIS 4667, at \*35 (S.D.N.Y. Jan. 22, 2009) (“A motive to maintain a higher financial rating to protect the viability of the Company, which is what the Complaint alleges here, is not enough, under the law of this Circuit, to sufficiently put forth a claim that a statement contained in an offering document was ‘fraudulent’ at the time it was made.”); *see also* *Alaska Elec. Pension Fund v. Adecco S.A. (In re Adecco S.A.)*, 371 F. Supp. 2d 1203, 1223 (S.D. Cal. 2005) (“A desire to conceal mismanagement is not sufficient to show motive and opportunity.”).

36. Both the doctrines of materiality and reliance serve this purpose. Materiality, which is required for many fraud-based claims, assesses whether a reasonable person would have considered the false statement important to his decision to enter into a transaction. *See* RESTATEMENT (SECOND) OF TORTS § 538 (AM. LAW INST. 1977) (a statement is material if, *inter alia*, “a reasonable man would attach importance to its existence or nonexistence in determining his choice of action in the transaction in question”); *see also* *United States v. Raza*, 876 F.3d 604, 619 (4th Cir. 2017) (“[T]he relevant elements of wire fraud are an intent to defraud and materiality, which Colton defined as ‘what a reasonable financial institution would

Heuristics such as scienter, materiality,<sup>37</sup> and reliance<sup>38</sup> thus generally get at the heart of many of the issues presented by the fact liability question. These heuristics ask the natural questions about factual statements, such as whether the speaker intended to mislead the person buying the car, whether the error would matter to a reasonable person buying a car, and whether the purchaser was entitled to (and did) rely on the statement because of some information asymmetry or because of the expertise or conduct of the speaker. All of these heuristics focus on the speaker and the context.<sup>39</sup>

Where there is a materially false statement, damages or rescission will often be available. In contract law, for example, there will be an escape hatch for mistake or when there is a failure to reach a meeting of the minds.<sup>40</sup> And, of course, where there is scienter sufficient for fraud, a contract will be voidable.<sup>41</sup> In some cases, there may be a statutory cause of action that provides relief for

---

want to know in negotiating a particular transaction.’ ”). Reliance, which is also an element of most fraud claims, requires that the person hearing the false statement thought the statement was important enough to act upon. Both doctrines ensure that unimportant statements, even if provably false, do not give rise to liability.

37. See generally Wendy Gerwick Couture, *Materiality and a Theory of Legal Circularity*, 17 U. PA. J. BUS. L. 453, 455 (2015) (“[Materiality doctrine] divid[es] misrepresentations that are potentially actionable from those that pose no risk of liability.”).

38. See Daniel B. Dobbs, *The Place of Reliance in Fraud*, 48 ARIZ. L. REV. 1001, 1009 (2006) (analogizing reliance in fraud cases to the role of proximate cause, because just as proximate cause requires that “the risks that are realized in the actual case are the risks that led us to characterize the defendant’s conduct as negligent toward the victim,” reliance in fraud determines “[w]hether the defendant has actually succeeded in harming the plaintiff by virtue of defrauding the plaintiff, as opposed to having harmed the plaintiff by deceiving others”).

39. See, e.g., *Omnicare Inc. v. Laborers Dist. Council Constr. Indus. Pension Fund*, 135 S. Ct. 1318, 1330 (holding that whether an opinion is misleading will depend on context, such as the custom and practices of the relevant industry).

40. The Second Restatement of Contracts defines a “mistake” as a “a belief that is not in accord with the facts.” RESTATEMENT (SECOND) OF CONTRACTS § 151 (AM. LAW INST. 1981). If the mistake is unilateral, meaning it is a factual mistake of only one of the parties, the contract is voidable only if it is shown that the mistaken party did not bear the risk of the mistake or that the other party knew of, or caused, the mistake. *Id.* § 153. When the mistake is mutually made by all of the parties, the contract is voidable by the adversely affected party if the mistake is about a basic assumption underling the contract. *Id.* § 152 & cmt. b. Of course, if there is no meeting of the minds, there was never a contract formed. In all of these cases, the relief at common law is restitution, meaning the “reversal of any steps that the parties may have taken by way of performance, so that each party returns such benefit as he may have received,” and in cases where this is not possible, damages. *Id.* § 158 & cmt. b.

41. Fraud, which generally requires proof of scienter, renders the transaction voidable, thus entitling the aggrieved party to restitution or rescission. See *Eklund v. Koenig & Assocs., Inc.*, 451 N.W.2d 150, 153 (Wis. Ct. App. 1989) (“When a party discovers an alleged fraud . . . , he may affirm the contract and sue for damages, or he may disaffirm and seek restitution.”); see also DANIEL B. DOBBS, *LAW OF REMEDIES* § 9.4, at 618 (1973) (stating that rescission and restitution are equitable remedies for fraud).

strictly false statements of fact due to information asymmetries inherent in certain types of transactions.<sup>42</sup>

Opinion statements include factual statements but are far more complex to evaluate for their liability. An opinion statement will often be based on one or more underlying fact(s),<sup>43</sup> but there is additional information being conveyed in an opinion statement. An opinion statement conveys not only that the speaker believes the facts underlying their opinion to be true, but also that they genuinely believe in their opinion, which is based on those facts.<sup>44</sup> In other words, an opinion statement contains not only factual information but information about the speaker's subjective belief in their stated judgment or decision-making process.<sup>45</sup>

In addition, opinions often convey information about the speaker's level of certainty about the facts and awareness of the facts.<sup>46</sup> A corporate executive

---

42. The most prominent examples are Sections 11 and 12 of the Securities Act of 1933, which provide for rescission or retraction damages upon a showing that a material statement in an offering prospectus was false or misleading. *See* Securities Act Section 11, 15 U.S.C. § 77k (2012); Securities Act Section 12, 15 U.S.C. § 77l (2012). Section 11 provides for damages arising from a false statement in a registration statement, and Section 12 provides for rescission or retraction damages. 15 U.S.C. §§ 77k(e) & 77l(a). There is no requirement that the false statement have been intentionally made. *Askelson v. Freidus (In re Barclays Bank PLC Sec. Litig.)*, No. 17-3293-cv, 2018 U.S. App. LEXIS 32622, at \*4 (2d Cir. Nov. 19, 2018). This is partly because of the information asymmetry that exists between issuer and the purchaser of the security. *See* William O. Douglas & George E. Bates, *The Federal Securities Act of 1933*, 43 YALE L.J. 171, 176 (1933) (“As stated above the protection given to investors by Section 11 fills a long felt need in so far as it shifts the burden of proof. This is particularly desirable during the early life of the security. At that time the registration statement will be an important conditioner of the market. Plaintiff may be wholly ignorant of anything in the statement. But if he buys in the open market at the time he may be as much affected by the concealed untruths or the omissions as if he had read and understood the registration statement. So it seems wholly desirable to create a presumption in favor of the investor in this regard.”).

43. Liability may, however, attach if a statement of fact embedded in an opinion statement is materially false or misleading. *See City of Dearborn Heights Act 345 Police & Fire Ret. Sys. v. Align Tech., Inc.*, 856 F.3d 605, 616 (9th Cir. 2017) (“[W]hen a plaintiff relies on a theory that a statement of fact contained within an opinion statement is materially misleading, the plaintiff must allege that ‘the supporting fact [the speaker] supplied [is] untrue.’”) (quoting *Omnicare*, 135 S. Ct. at 1327).

44. WILLIAM LLOYD PROSSER ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 109, at 755 (5th ed. 1984) (“[A]n expression of opinion is itself always a statement of . . . the fact of the belief, the existing state of mind, of the one who asserts it.”); *see also Omnicare*, 135 S. Ct. at 1327 (stating opinion with embedded statement of fact affirms both the underlying fact and the speaker's state of mind).

45. *See Omnicare*, 135 S. Ct. at 1327.

46. As comment a to Section 168 of the Second Restatement of Contracts explains, a statement of opinion “implies that [the speaker] does not have such definite information, that he is not certain enough of what he says, to make an assertion of his own knowledge as to that matter.” RESTATEMENT (SECOND) OF CONTRACTS § 168 cmt. a (AM. LAW INST. 1981); *see also*

that says he “believes” that his company is in compliance with federal law is likely really making a probabilistic statement based on the information he has.<sup>47</sup> The addition of the word “believe” transforms what would otherwise be a purely factual statement into one conveying both uncertainty and some level of diligence.<sup>48</sup>

In many contexts, therefore, it will not be enough for liability if one or more factual predicate of an opinion statement is false. There will have to be something more, such as evidence that the opinion is disingenuous or some showing that the opinion statement was frivolous, that it lacked any reasonable basis, or that the speaker simply never bothered to look at the facts they would normally look at before rendering an opinion.<sup>49</sup> The important questions

---

*Omnicare*, 135 S. Ct. at 1329 (“Reasonable investors understand that opinions sometimes rest on a weighing of competing facts; indeed, the presence of such facts is one reason why an issuer may frame a statement as an opinion, thus conveying uncertainty.”).

47. In *Omnicare*, management made statements in its registration statement to the effect that the company was in “compliance with applicable federal and state laws.” *Omnicare*, 135 S. Ct. at 1323. Because this belief was not alleged to have been disingenuous—that is, there was no allegation that the company did not sincerely believe it was in compliance with applicable laws—there was no basis upon which to allege that such an opinion statement was false. *Id.* at 1327. The statement, however, may have omitted material information, but even then, the mere fact that some contradictory information existed would not be enough to render the opinion statement misleading, because “[a] reasonable investor does not expect that every fact known to an issuer supports its opinion statement.” *Id.* at 1329. What may be implicitly conveyed by the opinion statement, however, is that there is some basis for the opinion, and in some cases, there are important facts that substantiate the opinion. If those facts are not provided, the opinion statement may mislead the listener. *Id.* at 1328. The opinion states:

[A] reasonable investor may, depending on the circumstances, understand an opinion statement to convey facts about how the speaker has formed the opinion—or, otherwise put, about the speaker’s basis for holding that view. And if the real facts are otherwise, but not provided, the opinion statement will mislead its audience.

*Id.* The common law rule, which the Restatement of Contracts articulates, is more stringent on the question of facts contradicting an opinion, as it presumes that an opinion statement’s implicit indication of uncertainty carries with it the representation that the speaker is not aware of any facts contrary to the opinion. *See* RESTATEMENT (SECOND) OF CONTRACTS § 168 (“If it is reasonable to do so, the recipient of an assertion of a person’s opinion as to facts not disclosed and not otherwise known to the recipient may properly interpret it as an assertion (a) that the facts known to that person are not incompatible with his opinion, or (b) that he knows facts sufficient to justify him in forming it.”); *see also id.* cmt. a (noting that an opinion statement “implies at most that [the speaker] knows of no facts incompatible with the belief or that he knows of facts that justify him in holding it”).

48. *See Omnicare*, 135 S. Ct. at 1334 (“The common law recognized that most listeners hear ‘I believe,’ ‘in my estimation,’ and other related phrases as disclaiming the assertion of a fact.”).

49. *See, e.g.,* *Twing v. Schott*, 338 P.2d 839, 843 (Wyo. 1959) (“The words of defendant Schott that the sewage system was ‘good’ and either ‘adequate’ or ‘sufficient’ could not have



revolve around the intent and subjective state of mind of the speaker, and in some cases, the reasonableness of that state of mind.<sup>50</sup>

B. SCIENTER AND OPINION LIABILITY

Because opinions are not true or false simply because some underlying factual predicate for the opinion turns out to be true or false, the important question is often whether the speaker was being disingenuous when stating an opinion.<sup>51</sup> In many cases, the opinion may be disingenuous if evidence exists that the speaker believed something contrary to the opinion when they stated it.<sup>52</sup> For example, a doctor who tells a patient that the prognosis for a particular surgery is good, but contemporaneously sends an email to a colleague saying otherwise, may have been disingenuous when stating their opinion to the patient. The same may be true for an investment advisor that recommends an investment product while privately telling a colleague that the product is “junk.”<sup>53</sup> In such cases, there is direct evidence that the opinion is not sincere,

---

been other than a fraudulent misrepresentation. We think his words constituted more than a puffing statement and more than any opinion. It was wholly inconsistent with the fact that he had repeatedly, according to his own admission, pumped out the cesspool. In that connection, it does not appear by testimony or otherwise that the purported dropping of rocks in the line would cause the cesspool to fill. It is true that the cesspool might have filled by reason of the use of excessive water by the tenants but if this were the fact then there would seem to be no excuse for failing to tell the prospective purchasers of the pumpings of the cesspool.”). Some cases have reasoned that the opinion statement coupled with “half truths” creates a duty to disclose in full all contradictory information. *See, e.g.,* *Mends v. Dykstra*, 637 P.2d 502, 508 (Mont. 1981) (holding that representations about the condition of a house were misleading given undisclosed knowledge of defects and problems). A complete lack of basis will also give rise to liability, because the person hearing the opinion may conclude that the opinion is not the sort of statement someone would make based on an uninformed judgment. *See Omnicare*, 135 S. Ct. at 1330 (“Investors do not, and are right not to, expect opinions contained in those statements to reflect baseless, off-the-cuff judgments, of the kind that an individual might communicate in daily life.”).

50. *See, e.g., Omnicare*, 135 S. Ct. at 1334–35.

51. *See Omnicare*, 135 S. Ct. at 1328 (“[A] statement of opinion is not misleading just because external facts show the opinion to be incorrect.”).

52. *See supra* note 49.

53. *See, e.g., Pursuit Partners, LLC v. UBS AG*, No. X05CV084013452S, 2009 Conn. Super. LEXIS 2313, at \*47 (Super. Ct. Sep. 8, 2009) (“The court takes [Defendant] employees at their word when they referenced their Notes, these purported ‘investment grade’ securities which they sold, as ‘crap’ and ‘vomit’, for [Defendant] alone possessed the knowledge of what their product, their inventory, was truly worth. While [Defendant] would argue that such descriptors lack a precise meaning, the true meaning of these words and the true value of [Defendant’s] wares became abundantly clear when the Plaintiffs’ multi-million dollar investment was completely wiped out and liquidated by [Defendant] shortly after the last of the Note purchases was consummated.”).

and it is reasonable to infer that the speaker had some improper motive for stating the opinion.<sup>54</sup>

This sort of opinion is in a sense a false statement because the implicit representation that the opinion is genuine is false,<sup>55</sup> and courts have no trouble assigning liability in such cases. In fact, many courts have required some evidence that the opinion was not genuinely held when stated to assign liability.<sup>56</sup> Indeed, in the securities law context, courts sometimes require a showing of scienter, even when the underlying cause of action imposes strict liability for false or misleading statements. For example, under the Securities Act of 1933, a false statement in a prospectus will give rise to rescission or damages, essentially allowing the purchaser to unwind a securities transaction premised on materially false factual statements in a prospectus.<sup>57</sup> There is no scienter requirement in the statute, but courts have required that the statement not only be proven objectively false, but also subjectively disbelieved by the issuer when the statement was made in the prospectus—in other words, opinion statements must be both subjectively and objectively false for liability to attach.<sup>58</sup>

Requiring scienter solves many of the problems with opinion liability—namely, the nearly intractable problem of having to prove that the speaker’s judgment was not only incorrect but should have been better.<sup>59</sup> In other words, the liability question would require a showing that the speaker’s judgment was somehow improper, and the clearest scenario where this is the case is where

---

54. *See id.*

55. *See supra* notes 44–45 and accompanying text.

56. *See, e.g., id.*

57. *See supra* note 42.

58. *See Fed. Hous. Fin. Agency v. UBS Ams., Inc.*, 858 F. Supp. 2d 306, 325 (S.D.N.Y. 2012) (“[P]laintiff must assert that the statement upon which it seeks to predicate liability ‘was both objectively false and disbelieved by the defendant at the time it was expressed.’”) (quoting *Fait v. Regions Fin. Corp.*, 655 F.3d 105, 110 (2d Cir. 2011)); *see also Omnicare*, 135 S. Ct. at 1327 (“What the Funds instead claim is that Omnicare’s belief turned out to be wrong—that whatever the company thought, it was in fact violating anti-kickback laws. But that allegation alone will not give rise to liability under § 11’s first clause because, as we have shown, a sincere statement of pure opinion is not an ‘untrue statement of material fact,’ regardless whether an investor can ultimately prove the belief wrong. That clause, limited as it is to factual statements, does not allow investors to second-guess inherently subjective and uncertain assessments. In other words, the provision is not, as the Court of Appeals and the Funds would have it, an invitation to Monday morning quarterback an issuer’s opinions.”).

59. By eliminating a cause of action based on a hindsight evaluation of a subjective judgment—what the *Omnicare* Court referred to as “Monday morning quarterback[ing]”—courts and fact finders do not need to decide whether they would have reached the same opinion given the set of facts as they were known or knowable when the opinion statement was made, nor do they have to determine in most cases whether the opinion was reasonable. *See Omnicare*, 135 S. Ct. at 1326.

there is evidence that the opinion was inconsistent with the speaker's own beliefs.<sup>60</sup> In those cases, it is fair to presume that the person hearing the opinion statement is entitled to at least the speaker's genuine opinion on the matter, which they did not receive.

That does not mean that it is the only sort of opinion that is problematic. In fact, there are a host of opinions that are plausible but flawed on their merits.<sup>61</sup> Heightening the standard for opinion liability essentially excludes these cases because so long as an opinion is plausible and there is no evidence that the speaker disbelieved the opinion, there is no liability.<sup>62</sup> This standard creates two significant problems: First, it excludes from liability the scenario where the opinion is facially plausible but the speaker rendered it with inadequate investigation into the relevant facts.<sup>63</sup> Second, it excludes from liability the sort of case where the opinion is rendered in the face of contradictory information that the person hearing the opinion would have wanted to know.<sup>64</sup> In most cases, there will not likely be clear evidence that the speaker disbelieves the opinion, and a strict scienter-based legal standard will not allow any way to further test the opinion statement for error or incompetence.<sup>65</sup>

### C. EXAGGERATED OPINIONS AND PUFFERY

In some cases, the context of the opinion statement may not justify an assumption that the opinion statement is grounded in supporting facts. It may be that the opinion is too general to be verifiably true or false, that the speaker's motive is such that one expects an exaggerated opinion, or both. The most

---

60. See *supra* note 49 and accompanying text.

61. In particular, an opinion may be based on misinterpretations of a set of underlying facts, based on dubious reasoning, or generally poorly thought out. These sort of opinion statements will not, without more, be actionable as misrepresentations.

62. See, e.g., *SEPTA v. Orrstown Fin. Servs., Inc.*, No. 1:12-cv-00993, 2015 U.S. Dist. LEXIS 80584, at \*98 (M.D. Pa. June 22, 2015) (dismissing claim where "Plaintiff has failed to point to a factual basis supporting its allegation that Defendant SEK did not believe its opinion" about financial statements).

63. A merely negligently-rendered opinion will not give rise to liability if a scienter-heuristic is used exclusively for liability. What is required is a completely unreasonable or inadequate basis for an opinion, such that the opinion is merely an unadorned, bald conclusion that lacks any support. In such a case, the baseless opinion may be sufficiently reckless to give rise to an inference of scienter. *Cf. Omnicare*, 135 S. Ct. at 1330.

64. The *Omnicare* opinion-liability framework mitigates this problem by allowing the basis of an opinion to be examined where the claim being considered is based on omitted facts from an opinion statement rather than based on the claim that an affirmatively-stated opinion was false or misleading opinion. See *id.*

65. See, e.g., *Omnicare*, 135 S. Ct. at 1326.

common example is the salesperson who exaggerates the value of their wares.<sup>66</sup> Most opinion-liability regimes assume that statements by a salesperson are often exaggerated and that those who hear such statements take them with a grain of salt.<sup>67</sup>

This is the rationale for the doctrine of puffery—which states that “an optimistic statement that is so vague, broad, and non-specific that a reasonable investor would not rely on it” is “immaterial as a matter of law.”<sup>68</sup> Such statements by a seller are usually presumed by the buyer to be overstated, exaggerated, or impossible to prove true or false.<sup>69</sup> The legal heuristics at work in this context are reliance and materiality, as the buyer would not be reasonable to rely on vague and overblown statements that salesmen are known to make and those statements are likely to be immaterial anyway.<sup>70</sup>

What the buyer can often reasonably assume, however, is that the seller’s overblown statements are not being made blatantly in the face of facts contrary to the opinion—that is, the opinion statement is “not fantastical.”<sup>71</sup> In other words, the speaker is likely representing that they are not aware of any facts contrary to their statement of opinion. That does not necessarily mean that

---

66. See RESTATEMENT (SECOND) OF TORTS § 539 cmt. c (AM. LAW INST. 1977) (“The habit of vendors to exaggerate the advantages of the bargain that they are offering to make is a well recognized fact.”).

67. This assumption is quite old, as it has been articulated in some of the earliest misrepresentation cases in the United States. See, e.g., *Kimball v. Bangs*, 11 N.E. 113, 114 (Mass. 1887) (“The law recognizes the fact that men will naturally overstate the value and qualities of the articles which they have to sell. All men know this, and a buyer has no right to rely upon such statements.”). As Judge Learned Hand has explained, it is presumed that there are statements that “no sensible man takes seriously, and if he does he suffers from his credulity. If we were all scrupulously honest, it would not be so; but, as it is, neither party usually believes what the seller says about his own opinions, and each knows it.” *Vulcan Metals Co. v. Simmons Mfg. Co.*, 248 F. 853, 856 (2d Cir. 1918).

68. *In re Gen. Elec. Co. Sec. Litig.*, 857 F. Supp. 2d 367, 384 (S.D.N.Y. 2012); see also *In re Vivendi, S.A. Sec. Litig.*, 838 F.3d 223, 245 (2d Cir. 2016) (“Puffery encompasses statements [that] are too general to cause a reasonable investor to rely upon them, and thus cannot have misled a reasonable investor.”) (internal citations and quotation marks omitted).

69. The Second Restatement of Contracts expressly assumes this about representations by sellers. RESTATEMENT (SECOND) OF CONTRACTS § 169 (AM. LAW INST. 1981) (“It may be assumed, for example, that a seller will express a favorable opinion concerning what he has to sell. When he praises it in general terms, commonly known as ‘puffing’ or ‘sales talk,’ without specific content or reference to facts, buyers are expected to understand that they are not entitled to rely.”).

70. See *supra* note 68.

71. See RESTATEMENT (SECOND) OF TORTS § 539 (“However, a purchaser is justified in assuming that even his vendor’s opinion has some basis of fact, and therefore in believing that the vendor knows of nothing which makes his opinion fantastic.”).

they believe that the facts underlying their opinion are objectively true.<sup>72</sup> Thus, in this context, the operative question is again the state of mind of the speaker—namely, their knowledge at the time the statement is made.

In many cases, however, puffery will simply not be actionable because such statements are likely to be too vague and general to evaluate, meaning they are not provably true or false.<sup>73</sup> In such cases, reliance, materiality, and intent are again the principal heuristics at work. A salesperson or seller's puffery cannot be reasonably relied on and therefore could not have been material,<sup>74</sup> and the statements may be too vague to have been intended as stating any facts, even about their state of mind.<sup>75</sup> If the seemingly exaggerated opinion has some specificity, then some showing that the speaker was aware of information contrary to his opinion will be required for liability.<sup>76</sup>

#### D. OMISSIONS AND INFORMATION ASYMMETRY

The more difficult case arises when there is information asymmetry and important—and perhaps contradictory—information is omitted from the opinion.<sup>77</sup> The clearest cases are when the speaker is—or is held out as—an

---

72. *See id.* § 168 (“If it is reasonable to do so, the recipient of an assertion of a person’s opinion as to facts not disclosed and not otherwise known to the recipient may properly interpret it as an assertion (a) that the facts known to that person are not incompatible with his opinion, or (b) that he knows facts sufficient to justify him in forming it.”); *see also id.* § 539 cmt. a (“Frequently a statement which, though in form an opinion upon facts not disclosed or otherwise known to their recipient, is reasonably understood as implying that there are facts that justify the opinion or at least that there are no facts that are incompatible with it.”).

73. *See In re PDI Sec. Litig.*, Civil Action No. 02-cv-0211 (JLL), 2005 U.S. Dist. LEXIS 18145, at \*69 (D.N.J. Aug. 16, 2005) (“Vague and general statements of optimism ‘constitute no more than puffery and are understood by reasonable investors as such.’”) (citation omitted); *see also* Stefan J. Padfield, *Is Puffery Material to Investors? Maybe We Should Ask Them*, 10 U. PA. J. BUS. & EMP. L. 339, 352 (2008) (“Puffery and statements of fact are mutually exclusive. If a statement is a specific, measurable claim or can be reasonably interpreted as being a factual claim, *i.e.*, one capable of verification, the statement is one of fact. Conversely, if the statement is not specific and measurable, and cannot be reasonably interpreted as providing a benchmark by which the veracity of the statement can be ascertained, the statement constitutes puffery.”) (internal citation omitted).

74. *See, e.g., In re Advanta Corp. Sec. Litig.*, 180 F.3d 525, 538 (3d Cir. 1999) (“Such statements, even if arguably misleading, do not give rise to a federal securities claim because they are not material.”).

75. *See State v. Am. TV & Appliance of Madison, Inc.*, 430 N.W.2d 709 (Wis. 1988) (“[E]xaggerations [are] reasonably to be expected of a seller as to the degree of quality of his product, the truth or falsity of which cannot be precisely determined.”) (internal quotation marks and citation omitted).

76. *See supra* Section I.B.

77. For an omission of fact to be actionable, there must generally be some duty to disclose information, for example, because of a fiduciary relationship or a relationship of trust and confidence between the parties. *See, e.g., Chiarella v. United States*, 100 U.S. 1108, 1115

expert on the subject of the opinion.<sup>78</sup> In such a case, the listener may not know even what facts are most important for a sound or justified opinion.<sup>79</sup> In other words, the listener may not only be relying on the judgment of the expert, but also the expert's judgment as to what information is most important.<sup>80</sup>

One of the clearest examples is the doctor-patient context.<sup>81</sup> In many cases, the lay patient can look at the same MRI results as the doctor but would not know what aspects of the results are significant for a diagnosis. The doctor, on the other hand, relies on education and experience to determine what aspects of the MRI results are most important.<sup>82</sup> The doctor not only has an

---

(1980) (noting that “silence in connection with the purchase or sale of securities may operate as a fraud actionable under § 10(b)” when there is “a duty to disclose arising from a relationship of trust and confidence between parties to a transaction”). Even without an affirmative duty to speak, a duty to disclose material information may arise because there may be a duty to speak fully and truthfully once a person has spoken. *See, e.g., Helwig v. Vencor, Inc.*, 251 F.3d 540, 561 (6th Cir. 2001) (“[E]ven absent a duty to speak, a party who discloses material facts in connection with securities transactions ‘assumes a duty to speak fully and truthfully on those subjects.’”) (citation omitted).

78. RESTATEMENT (SECOND) OF TORTS § 539 cmt. b (AM. LAW. INST. 1979). A statement of opinion

may also reasonably be understood to imply that [the speaker] does know facts sufficient to justify him in forming the opinion and that the facts known to him do justify him. This is true particularly when the maker is understood to have special knowledge of facts unknown to the recipient.

*Id.*

79. *See id.*

80. *See Omnicare, Inc. v. Laborers Dist. Council Constr. Indus. Pension Fund*, 135 S. Ct. 1318, 1335 (Scalia, J., concurring) (“[What] [the reasonable (female) person, and even he, the reasonable (male) person] would naturally understand a statement [of opinion] to convey is not that the statement has the foundation she (the reasonable female person) considers adequate. She is not an expert, and is relying on the advice of an expert—who ought to know how much ‘foundation’ is needed. She would naturally understand that the expert has conducted an investigation that he (or she or it) considered adequate. That is what relying upon the opinion of an expert means.”) (brackets and alterations in the original, quotations omitted).

81. *See id.* at 1334 (holding that the common law recognizes that “expressions of opinion made in the context of a relationship of trust, such as between doctors and patients” may give rise to opinion liability based on the basis of the opinion).

82. It is because an expert, such as a doctor, typically relies on his experience and judgment in reviewing facts underlying his opinion that the Advisory Committee amended Federal Rule of Evidence 703 to allow the admission of the expert's testimony about the underlying facts in certain cases without having to admit those facts individually at trial as out of court statements subject to the hearsay rule. Indeed, the Advisory Committee Notes to Rule 703 mention X-rays as examples of such evidence, which doctors apply their expertise to as a matter of course. *See* FED. R. EVID. 703 advisory committee's note (“Thus a physician in his own practice bases his diagnosis on information from numerous sources and of considerable variety, including statements by patients and relatives, reports and opinions from nurses, technicians and other doctors, hospital records, and X-rays. Most of them are admissible in

informational vantage that is superior to the patient because of their experience, but also a judgment and intuition advantage over the patient. When the doctor provides a diagnosis, they do not simply convey information about the underlying facts or even merely about the diagnosis or medical outcome, but may also be implicitly representing that they made a reasonable inquiry into the facts and correctly weighed the facts, including the facts contrary to his opinion.<sup>83</sup>

It is precisely when the underlying facts are contradictory, indeterminate, or incomplete that the opinion of an expert is most valuable. The information conveyed in such an opinion is more than factual—it's a conclusion about the facts that is inextricably bound up with the speaker's experience, intuition, and judgment.<sup>84</sup> And, in the case of an expert, the opinion invites reliance, particularly if the expert holds themselves out as a disinterested party.<sup>85</sup>

In the expert context, it will not be enough for liability that a fact underlying the opinion is, or turns out to be, false. In fact, it is expected that

---

evidence, but only with the expenditure of substantial time in producing and examining various authenticating witnesses. The physician makes life-and-death decisions in reliance upon them. His validation, expertly performed and subject to cross-examination, ought to suffice for judicial purposes.”).

83. Some of the earliest applications of this sort of expertise-asymmetry rationale appeared in cases involving physicians. *See, e.g., Hedin v. Minneapolis Med. & Surgical Inst.* 64 N.W. 158, 160 (Minn. 1895) (“The doctor, especially trained in the art of healing, having superior learning and knowledge, assured plaintiff that he could be restored to health. That the plaintiff believed him is easily imagined; for a much stronger and more learned man would have readily believed the same thing. The doctor, with his skill and ability, should be able to approximate to the truth when giving his opinion as to what can be done with injuries of one year’s standing, and he should always be able to speak with certainty before he undertakes to assert positively that a cure can be effected. If he cannot speak with certainty, let him express a doubt. If he speaks without any knowledge of the truth or falsity of a statement that he can cure, and does not believe the statement true, or if he has no knowledge of the truth or falsity of such a statement, but represents it as true of his own knowledge, it is to be inferred that he intended to deceive.”).

84. *See* RESTATEMENT (SECOND) OF TORTS § 542 cmt. f (AM. LAW. INST. 1979) (“The complexities and specializations of modern commercial and financial life have created many situations in which special experience and training are necessary to the formation of a valuable judgment. In this case if the one party has special experience or training or purports to have them, the other, if without them, is entitled to rely upon the honesty of the former’s opinion and to attach to it the importance that is warranted by his superior competence.”).

85. *See id.* § 542 cmt. h (“One who has taken steps to induce another to believe that the other can safely trust to his judgment is subject to liability if the confidence so acquired is abused. This is true not only when the maker of the fraudulent misrepresentation of opinion is or professes to be disinterested, as when the transaction is between the recipient and a third person, as to which see § 543, but also when he is known to have an adverse interest in the transaction.”).

there are contradictory or inconsistent facts underlying the opinion.<sup>86</sup> Indeed, if the expert could rely on only deterministic facts, there would be no room for their judgment.<sup>87</sup> There are, however, certain facts that any person, even one who relies on an expert, would want to know about. If a doctor states that an ailment appears benign based on their judgment and experience, but considered and disregarded a possible diagnosis that is likely terminal if not immediately treated, the patient would likely want to know about the disregarded diagnosis—the risk of harm is high, and the underlying information is time sensitive.<sup>88</sup> The patient would use that information to, perhaps, obtain a second opinion or at the least, to consciously decide to what extent they want to rely solely on the doctor's opinion.<sup>89</sup>

Even when the speaker is not an expert, there are contexts where the information asymmetry is so great that it is fair to assume that the speaker is better positioned not only to know all of the relevant facts but also how to weigh those facts. An officer of a public company is generally not free to share internal information about the company outside of a public filing with the SEC.<sup>90</sup> This results in a scarcity of information about the corporation between periodic filings, such as quarterly reports. The officer, however, presumably receives information in real time. Moreover, by virtue of his management position, he is aware of what information is most important to the operations and profitability of the company.<sup>91</sup> When the executive ultimately provides

---

86. See *supra* note 51.

87. An opinion based on determinative facts of obvious weight is not an opinion at all because there is no uncertainty about the facts to express. Such an opinion is likely simply nothing more than a set of factual statements.

88. See *Arato v. Avedon*, 858 P.2d 598, 607 (Cal. 1993) (“Rather than mandate the disclosure of specific information as a matter of law, the better rule is to instruct the jury that a physician is under a legal duty to disclose to the patient all material information—that is, ‘information which the physician knows or should know would be regarded as significant by a reasonable person in the patient’s position when deciding to accept or reject a recommended medical procedure’—needed to make an informed decision regarding a proposed treatment.”).

89. See *id.*

90. Although a reporting company must in some cases file interim reports concerning material corporate events, most internal information about a public corporation is in practice withheld until the next quarterly report. See 17 C.F.R. § 240.13a-11. There are other rules that prevent real-time disclosure of information, which creates information asymmetries. For a detailed discussion about the law surrounding the disclosure of corporate information, including under Regulation FD, see generally M. Todd Henderson & Kevin S. Haeberle, *Information-Dissemination Law: The Regulation of How Market-Moving Information Is Revealed*, CORNELL L. REV. 1373 (2016).

91. A Justice Scalia noted in his concurrence in *Omnicare*, it is reasonable to assume that corporate executives have expertise concerning the finances of the companies they run, including about corporate and financial information that must be set forth in an offering document or registration statement. See *Omnicare, Inc. v. Laborers Dist. Council Constr.*



information through public filings, what is reported implicitly carries with it not only the representation that what is reported is accurate, but also that what is reported is pertinent.<sup>92</sup> In this context, the corporate officer is similarly regarded as an expert. If the corporate officer makes representations about asset valuations based on a universe of factual inputs, the failure to state contradictory facts alongside the valuation opinion may be misleading, depending on the importance and weight of the omitted fact.<sup>93</sup>

Thus, generally in asymmetric information contexts, and specifically in expert opinions, the opinion's veracity may be sensitive to omitted information. Although the information asymmetry requires more reliance on the speaker in these contexts, that reliance also makes those who hear the opinion vulnerable to a form of blindness—they cannot see around the opaque corners that are likely transparent to the speaker. If the information the speaker relies on is outcome determinative or immensely important to the opinion, its disclosure may be as important as the conclusion communicated in the opinion.

#### E. OPINION STATEMENTS AS MODELS

If opinion statements are more than the facts underlying them, then what exactly are they? One way to think of an opinion is as a model of reality that is based on an individual's judgment and a universe of facts. The information

---

Indus. Pension Fund, 135 S. Ct. 1318, 1335 (Scalia, J., concurring) (“It is reasonable enough to adopt such a presumption for those matters that are required to be set forth in a registration statement. Those are matters on which the management of a corporation are experts. If, for example, the registration statement said ‘we believe that the corporation has \$5,000,000 cash on hand,’ or ‘we believe the corporation has 7,500 shares of common stock outstanding,’ the public is entitled to assume that the management has done the necessary research, so that the asserted ‘belief’ is undoubtedly correct.”).

92. *Cf. id.* The SEC generally requires management to provide a discussion and analysis of a public company's financial condition in order to “enable . . . investors to see the company through the eyes of management,” meaning that management must provide the information and form of information that it deems important as it manages the company. Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operation, Release No. 33-8350 (Dec. 29, 2003).

93. Whether an omission is material will depend in most misrepresentation cases on the context surrounding the statement that contained the omission. *See, e.g., Omnicare*, 135 S. Ct. at 1330 (“[A]n investor reads each statement within such a document, whether of fact or of opinion, in light of all its surrounding text, including hedges, disclaimers, and apparently conflicting information. And the investor takes into account the customs and practices of the relevant industry. So an omission that renders misleading a statement of opinion when viewed in a vacuum may not do so once that statement is considered, as is appropriate, in a broader frame. The reasonable investor understands a statement of opinion in its full context, and § 11 creates liability only for the omission of material facts that cannot be squared with such a fair reading.”).

available may be incomplete, the facts may be wrong, and in some cases, the facts may cut different ways or be subject to diverging interpretations.<sup>94</sup> An opinion makes sense of the universe of facts, assigns interpretation and weight to those facts, and maps the universe of facts to a conclusion, decision, or outcome.<sup>95</sup>

In the case of a trained expert, the opinion model is not only based on a universe of facts, but also on their experience. That is, the person holding the opinion not only makes sense of the universe of data points available to them, but also squares those data points with what they have seen in the past. When the expert has specialized training, a certain standard set of data points and background information is attributable to the expert. For example, a trained lawyer is deemed to have exposure to essential building blocks from contract and tort law and is generally imputed with a basic understanding of constitutional norms. Any opinion they render is against the backdrop of both their training and experience. All of the data from training, experience, and fact-gathering are combined together to form an opinion. Thus, an opinion can be thought of as a model of reality. It is a collection of facts, interpretations, weights, and probabilistic assessments.

Indeed, opinions are in some ways similar to mathematical and statistical models, which often seek to replicate the behavior of a particular aspect of reality in order to make predictions.<sup>96</sup> A useful analogy is a crude least-squares

---

94. See *supra* note 2.

95. Opinions are in some ways analogous to scientific theories, as both are built on some set of facts or axioms assumed to be true and some derived implications from those facts or axioms. The difference, of course, is that a scientific theory is only as good as its predictive power, and if its predictions can be proven incorrect, meaning they are falsifiable, the theory itself can be proven false. KARL POPPER, *THE LOGIC OF SCIENTIFIC DISCOVERY* 10. Popper states:

Next we seek a decision as regards these (and other) derived statements by comparing them with the results of practical applications and experiments. If this decision is positive, that is, if the singular conclusions turn out to be acceptable, or verified, then the theory has, for the time being, passed its test: we have found no reason to discard it. But if the decision is negative, or in other words, if the conclusions have been falsified, then their falsification also falsifies the theory from which they were logically deduced.

*Id.* Human opinions are evaluated for liability purposes, so the question is not whether the opinion is universally correct, but rather whether the opinion was justified under the circumstances. As explained *infra* Part III, however, AI opinions are closer to scientific theories, in that they can be tested for accuracy before being deployed.

96. See TIMOTHY GOWERS, *MATHEMATICS: A VERY SHORT INTRODUCTION* 4 (2002) (“Mathematics do not apply scientific theories directly to the world but rather to models. A model in this sense can be thought of as an imaginary, simplified version of the part of the world being studied, one in which exact calculations are possible.”).

regression.<sup>97</sup> It models what could be a noisy and scattered set of data points with a line. This line is a blunt instrument but can be useful to get a sense of correlations in the data.<sup>98</sup> In most cases, virtually none of the data points will fit the modeled line, meaning that in a sense they are contradictory to the simplistic line created to describe the data—but divergent data points do not make the model “false” simply because they do not fit neatly on the regression line.<sup>99</sup>

The model may still be useful for a crude estimate. It is more than the points that were used to create it. It is a reduction of the facts, and its value depends entirely on what it is used for. Sometimes a regression line is useful to make general predictions about a population—for example, age and height will correlate up until a certain age. If you’re using a height and age model to predict the height of elementary school students, it may be perfectly useful, but if you use the same model across a population that includes adults, the model is plainly insufficient and will be grossly inaccurate in many cases. Opinion statements are just as vulnerable to context. In the proper context, even a weak basis for an opinion may be sufficient.<sup>100</sup> That same basis in another context may be misleading.

In the case of statistical and mathematical models, some data points are so out of step with the entire data set that they are considered outliers.<sup>101</sup> How a

---

97. A least-squares regression is a simple mathematical model of data that attempts to fit a line to a set of data by minimizing the square of the error resulting from the fitted line’s predictions. *See generally* WILLIAM MENDENHALL, III ET AL., INTRODUCTION TO PROBABILITY AND STATISTICS 482–529 (14th ed. 2013).

98. Regressions are often too simple to be used to study complex datasets but are frequently used as a starting point because of their simplicity. JEFFREY M. WOOLDRIDGE, INTRODUCTORY ECONOMETRICS: A MODERN APPROACH 21 (6th ed. 2015) (“Although simple regression is not widely used in applied econometrics, it is used occasionally and serves as a natural starting point because the algebra and interpretations are relatively straightforward.”).

99. Some divergent points in a linear model will significantly skew the fitted line. Such divergent or influential data points are sometimes discarded as “outliers.” *See id.* at 326–27 (“Loosely speaking, an observation is an influential observation if dropping it from the analysis changed the key LS estimates by a practically ‘large’ amount.”). Ordinary Least Squares models are sensitive to outliers because the process minimizes the squares of errors or residuals, thus compounding the importance of large prediction errors. *See id.* at 327. (“OLS is susceptible to outlying observations because it minimizes the sum of squared residuals: large residuals (positive or negative) receive a lot of weight in the least squares minimization problem. If the estimates change by a practically large amount when we slightly modify our sample, we should be concerned.”).

100. For example, an opinion provided during an emergency or under time constraints may be adequate even though a reasonable person would under normal circumstances undertake a more detailed inquiry into the matter.

101. *See supra* note 99.

model deals with an outlier is important and will sometimes require disclosure for someone using the model to fully understand the power and effectiveness of the model.

The same may be true for an opinion statement. Some facts may be so contradictory to the opinion that the omission of the fact may render the model misleading. The same may also be true if the opinion is based on incomplete or potentially inaccurate data. Some information may be unknown or unknowable. If the omitted or missing information can affect the efficacy of the opinion's model of reality, that is when disclosure may be important, and that may also be when failure to disclose that information should give rise to liability.<sup>102</sup>

There is an important attribute of most models, including both opinion models created by human beings and determinative algorithms (such as a statistical regression), that is important for the purposes of this Article: one will generally be able to query the person making the model or examine the deterministic algorithm upon which the model is based to determine how factors were weighted, what facts were considered, and the effect omitted information may have had on the overall opinion calculus. A person can be placed under oath and put on the stand, and his intent can be discerned by a factfinder using long-tested legal constructs and heuristics.<sup>103</sup> In the case of a deterministic algorithm, the algorithm itself can be examined by experts or even directly by factfinders. So, there is usually at least some minimal modicum of transparency.

There are, to be sure, instances even in the case of human experts and deterministic algorithms where transparency will be greatly diminished. The most obvious example is when facts have been interpreted using the judgment,

---

102. This is the rationale courts have applied when an opinion is based on uncertain facts, but the speaker fails to say so. *See, e.g.,* Hedin v. Minneapolis Med. & Surgical Inst., 64 N.W. 158, 160 (“The doctor, with his skill and ability, should be able to approximate to the truth when giving his opinion as to what can be done with injuries of one year’s standing, and he should always be able to speak with certainty before he undertakes to assert positively that a cure can be effected. If he cannot speak with certainty, let him express a doubt.”).

103. For example, in a recent trial resulting in a criminal conviction for “spoofing,” the practice of using a computer program to rapidly place and cancel orders for securities to move a market, one of the most critical pieces of evidence at trial was the testimony of the computer program’s designer about what the trader instructed him to create and what the computer program was designed to do. *See* United States v. Coscia, 866 F.3d 782, 789 (7th Cir. 2017) (“The designer of the programs, Jeremiah Park, testified that Mr. Coscia asked that the programs act ‘[l]ike a decoy, which would be ‘[u]sed to pump [the] market.’ Park interpreted this direction as a desire to ‘get a reaction from the other algorithms.’ In particular, he noted that the large-volume orders *were designed specifically to avoid being filled . . .*”) (emphasis added).

experience, and intuition of an expert.<sup>104</sup> In those cases, it is difficult to determine how the underlying opinion model works. One cannot usually describe the vast degrees of freedom upon which a doctor with twenty years of medical experience bases their intuition.<sup>105</sup> But even in these cases, person's intent and set of motives can be examined. An expert with a motive to deceive will receive far less credit for his judgment than one without.<sup>106</sup> Moreover, experts may be judged against a standard of care reflecting their expertise, as they are in negligence cases,<sup>107</sup> which establishes a baseline of acceptable or

---

104. This opacity arises frequently when experts are called to testify in court about a technical subject. Commentators have questioned whether factfinders, such as judges and juries, are epistemically competent to hear such evidence, particularly given the tendency to rely on credibility heuristics when the substance of an expert's testimony is not accessible or understandable to a lay factfinder. *See, e.g.*, James R. Steiner-Dillon, *Expertise on Trial*, 19 COLUM. SCI. & TECH. L. REV. 247, 278 (2018) ("On the other hand, good intentions and genuine effort cannot create epistemic competence in the absence of substantive expertise. Jurors often fail to understand and apply scientific testimony correctly, even when the underlying science itself is relatively clear. They also tend to rely on specious proxies for substantive expertise."); *see also* Jennifer L. Mnookin, *Expert Evidence, Partisanship and Epistemic Competence*, 73 BROOK. L. REV. 1009, 1014 (2008) ("But if the jury lacks the knowledge that the expert provides, how, then, can it rationally evaluate the expertise on offer? To be sure, one might not need to be an expert in order to assess expertise, but the main mechanisms for assessing expertise outside of one's domain of knowledge are, by necessity, secondary indicia, proxies: demeanor, perhaps, or credentials, or superficial explanatory plausibility.").

105. *Cf.* Learned Hand, *Historical and Practical Considerations Regarding Expert Testimony*, 15 HARV. L. REV. 40, 54–55 (1901) ("The trouble with all this is that it is setting the jury to decide, where doctors disagree. The whole object of the expert is to tell the jury, not facts, as we have seen, but general truths derived from his specialized experience. But how can the jury judge between two statements each founded upon an experience confessedly foreign in kind to their own? It is just because they are incompetent for such a task that the expert is necessary at all. . . . What hope have the jury, or any other layman, of a rational decision between two such conflicting statements each based upon such experience.").

106. Again, in the context of testifying experts, an expert's motive for testifying—in many cases a fee—becomes an important proxy for credibility. *See* Mnookin, *supra* note 104, at 1014 ("Because the jury does not have the expertise to evaluate the substance of expert testimony, it is unlikely that it will be an accurate evaluator of partisan bias. . . . Without epistemic competence, the jury has no choice but to rely on proxies as secondary indicia of bias, and these may often be either inaccurate or difficult to evaluate.").

107. In negligence cases, the standard of reasonable care for one with expertise reflects his elevated capacity. *See* RESTATEMENT (THIRD) OF TORTS § 12 (AM. LAW INST. 2010) ("If an actor has skills or knowledge that exceed those possessed by most others, these skills or knowledge are circumstances to be taken into account in determining whether the actor has behaved as a reasonably careful person."); *see also* Omri Ben-Shahar & Ariel Porat, *Personalizing Negligence Law*, 91 N.Y.U. L. REV. 627, 641 (2016) ("Defendant's special skills are most often taken into account in cases where the defendant's profession is relevant to the injury. For example, doctors are held to a standard of care for their patients that is considerably higher than the reasonable person standard.").

valid models of reality that an expert may operate within.<sup>108</sup> This is why scienter and basic heuristics continue to function in these settings, even when there is some opacity resulting from the application of human experience, judgment, and intuition.

As explained in the next Part, AI models are different. They in many cases risk creating the opacity of an expert's intuition and judgment, but without the ability to examine a motive, standard of care, or set of human biases.<sup>109</sup> And, because they are generally not based on deterministic instructions, there are no clear instructions that can be used as a proxy for the intent of the AI's creator or user.

### III. ARTIFICIAL INTELLIGENCE, THE BLACK BOX PROBLEM, AND OPINION STATEMENTS

#### A. WHAT IS ARTIFICIAL INTELLIGENCE?

The term artificial intelligence generally refers to a class of computer programs capable of solving problems requiring inferential reasoning, decision making based on incomplete or uncertain information, classification, optimization, and perception.<sup>110</sup> AI can be based on determinative algorithms, such as a brute-force search,<sup>111</sup> or on machine-learning algorithms that learn directly from training examples.<sup>112</sup> The recent and rapid advances in AI have come mostly from the second category of AI—those built on machine-learning algorithms that learn from data,<sup>113</sup> such as deep networks of artificial neurons.

---

108. Cf. Ben-Shahrar et al., *supra* note 107, at 643 (“We saw that doctors are generally required to provide care that is at least as good as the average qualified medical practitioner, perhaps adjusted upwards to account for personal expertise.”).

109. See *infra* Section III.B.

110. See Bathaee, *supra* note 7, at 898.

111. An example of such a brute force algorithm would be a computer program that searches the space of possible chess moves to determine which move to make next using some deterministic scoring or ranking criteria. See Dave Gershgorin, *Artificial Intelligence Is Taking Computer Chess Beyond Brute Force*, POPULAR SCI. (Sept. 16, 2015), <http://www.popsci.com/artificial-intelligence-takes-chess-beyond-brute-force> [<https://perma.cc/PE5F-TSBE>].

112. See *id.*

113. See IAN GOODFELLOW ET AL., DEEP LEARNING 2 (2016) (“Several artificial intelligence projects have sought to hard-code knowledge about the world in formal languages. A computer can reason automatically about statements in these formal languages using logical inference rules. This is known as the knowledge base approach to artificial intelligence.”) (emphasis omitted). This approach of hard coding deterministic rules has given way to more powerful techniques that allow AI programs to learn directly from example and to make decisions based on a trained model's intuition. See *id.* at 1–2; see also Bathaee, *supra* note 7, at 898 (“On the most flexible end are modern AI programs that are based on machine-learning

Artificial neural networks are akin to neurons in the human brain, but they are not designed to mimic the function of biological neurons.<sup>114</sup> Rather, they are mathematical models—linear transformations often coupled with non-linear activation functions.<sup>115</sup> When combined into complex networks, they are capable of a form of cognition.<sup>116</sup> AI systems built on so-called “deep” architectures—stacked layers of artificial neurons—have been capable of performing tasks that most computers have been unable to perform at human-level proficiency.<sup>117</sup> In some applications, such as in the case of computer vision, these models exceed the proficiency of humans.<sup>118</sup>

AI programs may contain one or more of these underlying machine-learning algorithms.<sup>119</sup> Deep reinforcement learning systems, for example, use networks of artificial neurons to estimate future rewards when selecting from

---

algorithms that can learn from data. Such AI would, in contrast to the rule-based AI, examine countless other chess games and dynamically find patterns that it then uses to make moves.”).

114. See Bathae, *supra* note 7, at 901 (“The deep neural network is based on a mathematical model called the artificial neuron. While originally based on a simplistic model of the neurons in human and animal brains, the artificial neuron is not meant to be a computer-based simulation of a biological neuron. Instead, the goal of the artificial neuron is to achieve the same ability to learn from experience as with the biological neuron.”).

115. An artificial neuron is typically structured as a linear combination of parameters and weights. See GOODFELLOW ET AL., *supra* note 113, at 192. The output of that linear combination is then passed to a non-linearity, or activation function, which broadcasts or squelches the neuron’s output signal depending on the activation function’s criteria. The activation functions provide necessary non-linearity to the model—otherwise, a series of linear transformations will generally only be able to approximate linear patterns, and there would be little additional power that would result from deepening a network of artificial neurons. *Id.* at 192. By adding a non-linearity, it is posited that a deep neural network can approximate important classes of non-linear functions in finite-dimensional space. See *id.* at 194 (“Specifically, the universal approximation theorem . . . states that a feedforward network with a linear output layer and at least one hidden layer with any ‘squashing’ activation function (such as the logistic sigmoid activation function) can approximate any Borel measurable function from one finite-dimensional space to another with any desired non-zero amount of error, provided the network is given enough hidden units.”).

116. The notion that cognition occurs in deeply interconnected networks, such as in both biological and artificial neural networks, is called connectionism. PETER FLACH, MACHINE LEARNING: THE ART AND SCIENCE OF ALGORITHMS THAT MAKE SENSE OF DATA 16 (2012) (“The central idea in connectionism is that a large number of simple computational units can achieve intelligent behavior when networked together. This insight applies equally to neurons in biological nervous systems as it does to hidden units in computational models.”).

117. See *supra* note 15.

118. See *id.*

119. This Article distinguishes between machine learning and AI systems because AI is referred in this Article as systems that may include one or more machine learning-based sub-systems (and therefore employ one or more machine-learning algorithms).

a set of possible actions.<sup>120</sup> Reinforcement learning algorithms built on artificial neural networks have been able to defeat professional Go players, chess players, and even expert-level humans at complex real-time strategy games.<sup>121</sup>

What is striking about AI computer programs that are built on machine-learning algorithms is that they can be built to map an arbitrary set of states to an arbitrary set of actions in pursuit of complex goals.<sup>122</sup> A deep reinforcement learning system, for example, may converge on an optimal battlefield strategy, simply by repeating millions of simulated engagements.<sup>123</sup>

Deep machine-learning algorithms, such as deep neural networks, are significantly more complex with size, as no single artificial neuron or layer of artificial neuron bears much individual responsibility for the model's decisions.<sup>124</sup> Thus as the network of artificial neurons increases in size, the

---

120. Reinforcement learning algorithms are algorithms designed to “maximize a numerical reward signal,” but unlike most forms of machine learning, reinforcement learning algorithms “must discover which actions yield the most reward by trying them.” RICHARD S. SUTTON & ANDREW G. BARTO, REINFORCEMENT LEARNING: AN INTRODUCTION 2 (1998). A “deep” reinforcement system relies on deep architectures of neural networks to predict future rewards, thus enabling the reinforcement learning system to converge on an environments maximum rewards after repeated trial and error. *See generally* Maxim Lapan, DEEP REINFORCEMENT LEARNING HANDS-ON, loc. 2419 (2018) (ebook) (describing implementation of deep Q-learning system).

121. *See* David Silver et al., *Mastering the Game of Go Without Human Knowledge*, 550 NATURE 354, 354–59 (2017); David Silver et al., *A General Reinforcement Learning Algorithm that Masters Chess, Shogi, and Go through Self-Play*, 362 SCIENCE 1140 (2018); OPENAI FIVE (June 25, 2018), <https://blog.openai.com/openai-five/> [<https://perma.cc/QB49-KTD9>] (“Our team of five neural networks, OpenAI Five, has started to defeat amateur human teams at Dota 2.”).

122. This is particularly true for reinforcement learning systems that use deep neural networks, as such systems can learn to execute complex sequences of actions that require planning, meaning anticipating the future and estimating long-term rewards. *See* Razvan Pascanu et al., *Agents that Imagine and Plan*, GOOGLE DEEP MIND (July 20, 2017), <https://deepmind.com/blog/agents-imagine-and-plan/> [<https://perma.cc/C5LK-MKK2>] (“We have seen some tremendous results in this area—particularly in programs like AlphaGo, which use an ‘internal model’ to analyse how actions lead to future outcomes in order to reason and plan.”).

123. *See* SUTTON & BARTO, *supra* note 120, at 4 (“These two characteristics—trial-and-error search and delayed reward—are the two most important distinguishing features of reinforcement learning.”).

124. *See* Davide Castelvecchi, *Can We Open the Black Box of AI?*, 538 NATURE 20, 22 (2016) (“But this form of learning is also why information is so diffuse in the network: just as in the brain, memory is encoded in the strength of multiple connections, rather than stored at specific locations, as in a conventional database.”); *see also* Bathaee, *supra* note 7, at 891–92 (“AI that relies on machine-learning algorithms, such as deep neural networks, can be as difficult to understand as the human brain. There is no straightforward way to map out the decision-making process of these complex networks of artificial neurons.”).



capacity of the AI model likewise increases.<sup>125</sup> With that increase in capacity, however, comes opacity.<sup>126</sup> A fully trained neural network is capable of making decisions the same way a trained expert makes decisions—based on experience and intuition.<sup>127</sup> In other words, there are no detailed instructions given to a computer as in the case of traditional computer programs, but instead, AI programs are often products of the data on which they have been trained.<sup>128</sup> In a sense, the patterns in the underlying training data govern the AI program’s decision making. Because the complex network of artificial neurons allows for countless permutations, no single neuron or even layer of neurons encodes any particular part of the decision-making process.<sup>129</sup> Although the inputs to these models are often known, information, such as how those inputs are weighed as they propagate through the networks, may be nearly impossible to determine.

#### B. THE BLACK BOX PROBLEM

Modern deep neural networks can be very deep and are extremely interconnected. This means that there may not be any clear way of understanding the decision-making process of the network once it is trained on the data.<sup>130</sup> Moreover, the inputs to machine-learning algorithms, including deep neural networks, are often multi-dimensional, meaning that various input

---

125. Although it is not entirely understood why deeper architectures increase in capacity to approximate non-linear functions, it is assumed that it may be because deeper architectures are decomposing non-linear functions into components that can be incrementally estimated. *See* GOODFELLOW ET AL., *supra* note 113, at 195 (“Choosing a deep model encodes a very general belief that the function we want to learn should involve composition of several simpler functions. This can be interpreted from a representation learning point of view as saying that we believe the learning problem consists of discovering a set of underlying factors of variation that can in turn be described in terms of other, simply underlying factors of variation.”).

126. *See* Bathaee, *supra* note 7, at 894 (“Deep networks of artificial neurons distribute information and decision-making across thousands of neurons, creating a complexity that may be as impenetrable as that of the human brain.”).

127. *See id.* at 902 (“The net result is akin to the way one ‘knows’ how to ride a bike. Although one can explain the process descriptively or even provide detailed steps, that information is unlikely to help someone who has never ridden one before to balance on two wheels. One learns to ride a bike by attempting to do so over and over again and develops an intuitive understanding.”); *cf.* Siddhartha Mukherjee, *A.I. Versus M.D.*, NEW YORKER (Apr. 3, 2017), <http://www.newyorker.com/magazine/2017/04/03/ai-versus-md> [<https://perma.cc/MY9K-LBVG>] (describing distinction between “knowing that” and “knowing how” forms of learning, where “knowing how” arises from trial and error and is learned from experience).

128. *See* Bathaee, *supra* note 7, at 902–03 (“Because a neural network is learning from experience, its decision-making process is likewise intuitive. Its knowledge cannot in most cases be reduced to a set of instructions, nor can one in most cases point to any neuron or group of neurons to determine what the system found interesting or important.”).

129. *See id.*

130. *See infra* Section III.B.

parameters are encoded as high dimensional vectors.<sup>131</sup> Machine-learning algorithms, such as Support Vector Machines, rely on special relationships in higher-dimensional vector spaces.<sup>132</sup> In other words, if there are 115 different parameters used by a model, then machine-learning algorithms will search for patterns in 115 or more dimensions, a sort of geometric space that humans simply cannot visualize.<sup>133</sup> The net effect is both opacity from the vast number of interconnected layers and the difficulty of visualizing higher-dimensional patterns.<sup>134</sup> So there is no clear way for human beings to easily examine the patterns that a machine-learning algorithm may be seizing on as part of its decision-making process.

To complicate things further, the systems built on these machine-learning algorithms may introduce additional opacity to the decision-making process. A reward-seeking reinforcement learning system that uses a deep neural network to estimate future rewards for certain actions may mask the underlying patterns that the deep neural network has detected—all that a human will be able to discern is the estimated rewards for the next actions and those thereafter.<sup>135</sup> For example, a deep reinforcement learning system may predict an eventual checkmate several dozen moves in the future and choose the next move (e.g., moving a pawn two steps forward) that would lead to that outcome, but it may be impossible to tell what series of future moves will ultimately lead to such a result.<sup>136</sup>

---

131. For example, a model that uses three inputs, height, weight, and age, to predict the amount of time it takes for a person to run one mile would receive inputs as a three-dimensional vector (one for each input parameter) and would be searching a three-dimensional space of data for patterns.

132. For a description of Support Vector Machines (SVMs) and how they create opaqueness because of dimensionality, see Bathaee, *supra* note 7, at 903–04; *see also id.* at 905 (“Thus, when the number of variables or features provided to an SVM becomes large, it becomes virtually impossible to visualize how the model is simultaneously drawing distinctions between the data based on those numerous features.”).

133. *See* Bathaee, *supra* note 7, at 892 n. 14 (“A two-dimensional space can be visualized as a series of points or lines with two coordinates identifying the location on a graph. To represent a third dimension, one would add a third axis to visualize vectors or coordinates in three-dimensional space. While four dimensions can be visualized by adding a time dimension, five dimensions and higher are impossible to visualize.”).

134. *See id.* at 901–04.

135. The output of a deep “Q-learning” reinforcement system, for example, may be a vector of long-term rewards associated with a set of possible actions. *See* Lapan, *supra* note 120, at loc. 2638. Those rewards may not provide any insight into what patterns the reinforcement learning system’s deep neural network has spotted and correlated with the anticipated reward.

136. Much of this depends on the structure of the reinforcement learning system. Some reinforcement learning systems evaluate particular moves on a tree of possible outcomes to estimate the value of a move or sequence of moves—in those cases, there may be more

All of these technologies for the first time provide computers the ability to make decisions as humans do—based on experience.<sup>137</sup> A trained neural network will use a decision-making process akin to intuition or judgment.<sup>138</sup> It is essentially the difference between a person who is given detailed instructions on how to ride a bike and a person who has learned to ride a bike—to balance and shift weight—through experience and iteration.<sup>139</sup>

Different machine-learning algorithms create varying levels of opacity. Some can be queried in a way such that outcome-determinative inputs can be ascertained. Others cannot. There are, therefore, both weak and strong versions of the Black Box Problem.<sup>140</sup> All things point to a trend towards the strong form as the technology progresses. Complexity in modern neural networks has increased significantly and it is likely that neural networks will continue to deepen in architecture and increase in size and connectivity.<sup>141</sup>

This Article mostly addresses the strong form of the Black Box Problem. In other words, I assume that fully trained AI systems will be mostly opaque—that the decision-making process cannot be determined by probing the model with different inputs.<sup>142</sup> This is the most problematic incarnation of the Black Box Problem for most legal doctrines, and it is the most important form to consider for legal constructs that rely on intent or scienter heuristics.<sup>143</sup>

### C. AI OPINIONS

The most direct use of AI programs built on machine-learning algorithms are systems designed to predict outcomes or to classify data.<sup>144</sup> AI is already

---

transparency as to what course of action the model favors. *See, e.g.,* Silver et al., *supra* note 121, at 2 (noting the use of a Monte Carlo search tree to evaluate potential moves).

137. *See supra* note 127.

138. *See id.*

139. *See supra* note 126.

140. The strong version of the AI Black Box Problem posits that there is no way to determine a rank-order of importance for a model's inputs or to determine how the model is arriving at decisions. *See* Bathaee, *supra* note 7, at 906. The weak form assumes that a loose ordering of input importance can be ascertained. *See id.*

141. Neural network depth will likely increase because it is generally the case that deeper networks potentially have exponentially greater capacity to approximate functions. *See* GOODFELLOW ET AL., *supra* note 113, at 196 (“[P]iecewise linear networks (which can be obtained from rectifier nonlinearities or maxout units) can represent functions with a number of regions that is exponential in the depth of the network.”).

142. *See* Bathaee, *supra* note 7, at 906.

143. *See id.* at 906–08.

144. This is because many underlying machine-learning algorithms, including deep neural networks, can be configured directly to classify data or to provide a bounded output, such as a regression or a sigmoid output function. *See* GOODFELLOW ET AL., *supra* note 113, at 166, 347.

being used to make diagnostic predictions given imaging information (such as from MRI results or X-rays).<sup>145</sup> Electronic Medical Records can be mapped to therapeutic outcomes or risk factors.<sup>146</sup> AI can be used to make predictions about what advertisements or search results to display.<sup>147</sup> It can be used to value real estate given a set of inputs about a particular piece of property, or to determine whether a borrower or counterparty is creditworthy. The applications are numerous and rapidly growing.

These applications are natural progressions from deterministic algorithms that occupied the space for decades prior to the recent explosive growth of AI. Automated valuation models, for example, were a prominent feature of the underwriting and appraisals that led to the mortgage-backed securities crisis that occurred after 2008.<sup>148</sup> In those cases, the algorithms and models used did not shield any of the actors from liability because the decision-making process remained mostly in human hands. In fact, when humans made decisions in those cases that ignored the algorithms, their decisions to do so sometimes served as a basis for a finding of scienter.<sup>149</sup>

Even computerized securities trading systems, like high frequency trading systems, which trade securities in fractions of a second,<sup>150</sup> although sometimes autonomous, were for years deterministic, meaning one merely had to examine the underlying code to determine what the intent of the programmer or user

---

145. See *supra* note 12. Many of these image-based models use convolutional neural networks to extract patterns from visual data such as images. See GOODFELLOW ET AL., *supra* note 113, at 326 (“[Convolutional Neural Networks] are a specialized kind of neural network for processing data that has a known grid-like topology. Examples include time-series data, which can be thought of as a 2-D grid of pixels.”).

146. See Huiying Liang & Brian Y. Tsui, *Evaluation and Accurate Diagnoses of Pediatric Diseases Using Artificial Intelligence*, NATURE MEDICINE (Feb. 11, 2019), <https://www.nature.com/articles/s41591-018-0335-9> [<https://perma.cc/WK68-R7H4>] (“Our model applies an automated natural language processing system using deep learning techniques to extract clinically relevant information from EHRs. In total, 101.6 million data points from 1,362,559 pediatric patient visits presenting to a major referral center were analyzed to train and validate the framework. Our model demonstrates high diagnostic accuracy across multiple organ systems and is comparable to experienced pediatricians in diagnosing common childhood diseases.”).

147. See Tom Simonite, *Google and Microsoft Can Use AI to Extract Many More Ad Dollars from Our Clicks*, WIRED (Aug. 31, 2017, 7:00 AM), <https://www.wired.com/story/big-tech-can-use-ai-to-extract-many-more-ad-dollars-from-our-clicks/> [<https://perma.cc/4B4Z-LAH7>].

148. See *supra* note 7.

149. See, e.g., *Fed. Hous. Fin. Agency v. Nomura Holding Am., Inc.*, 104 F. Supp. 3d 441, 479 (S.D.N.Y. 2015) (finding that decisions to include loans in mortgage-backed securities, notwithstanding automated valuations that exceeded tolerances, could serve as a basis for opinion liability).

150. See Bathaee, *supra* note 7, at 908–09.

was when the program was deployed. Indeed, in one of the first criminal trials concerning an unlawful trading practice called spoofing, wherein phantom orders were placed and canceled in fractions of a second in order to move the market, the jury's verdict was based on the testimony of the programmer who created the program at the request of the trader.<sup>151</sup> A human testified about intent because intent was ascertainable—a human provided the computer system detailed instructions, which either evinces an intent to spoof or does not.<sup>152</sup>

Valuations, risk assessments, and even hiring decisions are natural applications for AI models because they were already the subject of intricate deterministic computer programs. Many of the decisions made by computer programs in these fields were based on hard rules or crude statistical patterns (such as linear regressions). The ability to create computer programs that perform the same tasks based on complicated patterns in underlying data—perhaps data collected from hundreds of thousands of human decisions—is undoubtedly the next step for many businesses, governments, and institutions.

Because these models make decisions based on patterns in data and a number of case-specific factors, their outputs are likely to be considered opinions. The output of an AI model that values a security or evaluates counter-party risk or diagnoses patients will be more than a set of underlying facts, more than a set of hard (or even fuzzy) rules, and more than the broad patterns and correlations in the underlying data. They will be opinions to the same extent decisions based on human judgment are opinions, but with one important difference—there will be no human to put on the witness stand to describe the decision-making process that produced the opinion. And there will, in many cases, be no parity between the intent of the AI model's creators and the AI's decision-making schema.<sup>153</sup>

#### D. THE FAILURE OF THE SCIENTER HEURISTIC

The most serious legal problem posed by any complex AI system is the decoupling of the intent of the system's creators from the system's decisions.<sup>154</sup> A deep reinforcement learning system may be provided a scheme of clear rewards by the designer, but the AI may have many degrees of freedom in how it pursues those rewards and may have to traverse a massive state space to

---

151. *See supra* note 102.

152. *See id.*

153. Bathaee, *supra* note 7, at 926 (“It is clear that a strong black box, however, cannot be interrogated. Its decision-making process cannot be audited.”).

154. *See* Bathaee, *supra* note 7, at 908.

obtain them,<sup>155</sup> which means how the model performs in pursuit of those rewards may be unpredictable. A seemingly absurd, but often recounted example, is Nick Bostrom's paperclip maximizer—an AI tasked with producing as many paperclips as possible. That AI is operating within its parameters even if it consumes all of the resources in the world to obtain its slated rewards. Indeed, because humans would be the source of precious atoms from which paperclips can be made, “the future that the AI would be trying to gear towards would be one in which there were a lot of paper clips but no humans.”<sup>156</sup>

The problem is referred to as instrumental convergence.<sup>157</sup> It is the hypothetical notion that AI of a sufficient amount of intelligence will seek to obtain unbounded instrumental goals by maximizing resource acquisition as well as the system's own self-preservation (to ensure its longevity as it pursues its unbounded goals).<sup>158</sup> This is not a literal problem for AI systems today—indeed, few AI systems have unbounded instrumental goals and even fewer are directly plugged into sensitive systems. The problem, including the paperclip hypothetical, however, makes clear that the Black Box Problem is

---

155. Even a real-time strategy video game, such as StarCraft, creates a massive state and action space that a reinforcement learning system must traverse for rewards. *See, e.g.,* Zhen-Jia Pang et al., *On Reinforcement Learning for Full-length Game of StarCraft*, ARXIV 2 (Feb. 3, 2019), <https://arxiv.org/pdf/1809.09095.pdf> [<https://perma.cc/69P7-37L9>] (“From the perspective of reinforcement learning, StarCraft is a very difficult problem. Firstly, it is an imperfect information game. Players can only see a small area of map through a local camera and there is a fog of war in the game. Secondly, the state space and action space of StarCraft are huge. StarCraft's image size is much larger than that of Go. There are hundreds of units and buildings, and each of them has unique operations, making action space extremely large.”).

156. Kathleen Miles, *Artificial Intelligence May Doom The Human Race Within A Century, Oxford Professor Says*, HUFFINGTON POST (Aug. 22, 2014), [https://www.huffingtonpost.com/2014/08/22/artificial-intelligence-oxford\\_n\\_5689858.html](https://www.huffingtonpost.com/2014/08/22/artificial-intelligence-oxford_n_5689858.html) [<https://perma.cc/5HT7-US6K>].

157. *See* Nick Bostrom, *The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents*, MINDS & MACHINES 6 (2012), <https://nickbostrom.com/superintelligentwill.pdf> [<https://perma.cc/L69K-J24V>].

158. *See id.* More formally, the Instrumental Convergence Thesis posits that:

Several instrumental values can be identified which are convergent in the sense that their attainment would increase the chances of the agent's goal being realized for a wide range of final goals and a wide range of situations, implying that these instrumental values are likely to be pursued by many intelligent agents.

*Id.*; *see also id.* at 7 (“Suppose that an agent has some final goal that extends some way into the future. There are many scenarios in which the agent, if it is still around in the future, is then [] able to perform actions that increase the probability of achieving the goal. This creates an instrumental reason for the agent to try to be around in the future—to help achieve its present future-oriented goal.”).

not just the result of the complexity of machine-learning algorithms, but also the opacity created by rewards system. Thus, one may be able to specify an AI system's goals very clearly but may not be able to anticipate how the AI achieves those goals, or even what instrumental goals it may deem necessary to achieve them.<sup>159</sup>

Reinforcement learning systems already exceed the capabilities of their creators at the tasks to which they are applied,<sup>160</sup> and indeed, sometimes even exceed the expectations of their creators or their creators' understanding of the problem. For example, those watching Google Deep Mind's AlphaGo and AlphaGo Zero AI play human champions have commented that there is something inhuman about the moves made by the program.<sup>161</sup> It is also beyond dispute that the AI exceeded its creators' ability at the game of Go—indeed, the AI defeated the very best human Go players in the world, which had no hand in the creation of the AI.<sup>162</sup> In other words, the AI's decisions are much more than the mere reward and value specifications set forth by its creator—the AI is making its own decisions.

Consider a reinforcement learning system that is given a reward system based on the amount of money it makes in an electronic trading market. If it stumbles upon spoofing or other forms of market manipulation as a viable strategy for maximizing the rewards it has been tasked to obtain, it may do so notwithstanding the fact that its creators never intended to break the law or engage in a manipulative strategy.<sup>163</sup> Of course, one may object and say that the creator has a duty to impose constraints, but what if the reinforcement learning system stumbles upon a manipulative trading strategy that no human had yet thought of or could even execute (for example, because it would require simultaneous cognition and coordination across thousands of different markets)? Worse yet, what if humans cannot tell that the AI's decisions are

---

159. *See id.* at 5 (“The orthogonality thesis implies that synthetic minds can have utterly non-anthropomorphic goals—goals as bizarre by our lights as sand-grain-counting or paperclip-maximizing. This holds even (indeed especially) for artificial agents that are extremely intelligent or superintelligent.”). Notably, Bostrom believes it is conceptually possible to design systems that behave in a predictable fashion. *See id.* at 7–8. The question is an open one, and as this Article contends, it may be a technological one which depends on the complexity of an AI's internal models. *See supra* note 141 and accompanying text.

160. *See, e.g., supra* note 15.

161. *See* Cade Metz, *How Google's AI Viewed the Move No Human Could Understand*, WIRED (Mar. 14, 2016 2:39 AM), <https://www.wired.com/2016/03/googles-ai-viewed-move-no-human-understand/> [<https://perma.cc/LYQ2-3WJN>].

162. *See id.*

163. *See* Bathaee, *supra* note 7, at 911.

manipulative or unlawful because the AI converges on an obfuscated form of manipulation not strictly prohibited by a priori constraints?<sup>164</sup>

As AI becomes more sophisticated, all of this becomes exceedingly problematic for the hundreds of years' worth of legal doctrines and heuristics that we have accumulated, particularly those based on notions of intent or foreseeability.<sup>165</sup> The intent heuristic fails, because in many cases, AI may be provided perfectly legitimate rewards or ends to pursue, but because the degrees of freedom among possible actions it may pursue is high, the creator of the AI may not be able to predict how the AI will achieve those goals. Indeed, it may be impossible to foresee all of the possibly problematic sequences of actions that the AI may take in pursuit of the rewards it has been given, and that means that the creator of the AI may not be able to anticipate every constraint that would be necessary to keep the AI in line. The high degrees of freedom in the actions the AI can take means that certain actions may simply not be foreseeable, making scienter based on even recklessness or gross negligence impossible to prove because some awareness of risk or foreseeability is a necessary predicate for them.<sup>166</sup>

Because scienter cannot be satisfied when Black Box AI is involved, the law may excuse any injury inflicted by AI simply because no human intended the injury.<sup>167</sup> The net effect is the anomalous result where conduct, if done by a human, would result in liability, but if done by AI would be immune from liability.<sup>168</sup> Thus, a person who engages in spoofing may be convicted of a crime, but a person who designs AI that stumbles upon spoofing as an effective, reward-maximizing strategy, will not result in any liability because the creator of the AI never told it to engage in such a strategy.<sup>169</sup>

There is therefore a perverse incentive to use an AI model to make decisions in highly regulated environments because the AI functionally cuts off any possible liability.<sup>170</sup> Indeed, AI that discriminates based on gender

---

164. The AI may perceive obfuscation of its strategy as an instrumental goal, such as the overarching goal to survive described by Bostrom. *See supra* note 156. In other words, if avoiding detection of the AI's impermissible trading strategy is a necessary sub-goal of its overarching goal to obtain rewards, it will seek to optimize on that sub-goal as well as the rewards.

165. *See* Bathaee, *supra* note 7, at 892, 922.

166. *See* Bathaee, *supra* note 7, at 907.

167. *See id.*

168. *See id.*

169. *Id.* at 911.

170. Consider the current incentive to build complex corporate hierarchies. In the case of corporations charged with federal crimes, only about a third of the cases involved charges against individuals. *See* Brandon L. Garrett, *The Corporate Criminal as Scapegoat*, 101 VA. L. REV. 1790, 1802 (2015). Among those charged, "many were not higher-up officers of the



because of biases in the data used to train it will likely not result in liability for the company that created the AI (because there is no evidence of scienter or even negligence),<sup>171</sup> whereas if the same hiring decision was made by a human, there would possibly be exposure to lawsuits or, at a minimum, ethical objections.

#### E. AN OPAQUE BASIS AND MATERIALITY

Even without reward-based specifications and reinforcement-learning systems, the basis upon which a vast network of artificial neurons makes decisions will in many cases be impossible to determine.<sup>172</sup> In the case of deep neural networks, the nature of the highly connective system of linear and non-linear mathematical transformations of the data may result in mappings that cannot be readily understood—even if various inputs are provided to the model to determine boundary conditions.<sup>173</sup> In fact, the more complex the decision-making process, the less likely it will be that the AI's decision making can be mapped out simply by examining correlations between inputs and outputs of the AI model.<sup>174</sup>

For example, AI that makes a medical diagnosis based on three or four parameters in a patient's medical file can likely be probed—if one of the parameters is age, then the input age can be varied to determine whether the change is outcome determinative. But as more parameters are added, the

---

companies, but rather middle managers of one kind or another and also some quite low-level individuals.” *Id.* at 1802. The explanation may be that corporate complexity results in the insulation of senior corporate executives from liability, as the more complex the organization becomes, the more the corporate institution can be blamed instead of the willful conduct of any single person. *Id.* at 1825. The incentive to use AI may be somewhat similar to the incentive to create complexity in a corporation—the complexity and opaqueness of the AI diffuses responsibility and insulates senior corporate officers, particularly if the programming, testing, and business applications of the AI are responsibilities of different parts of an organization.

171. At least one company has scrapped its AI designed to vet potential employees because their AI discriminated based on gender. Jeffrey Dastin, *Amazon Scraps Secret AI Recruiting Tool that Showed Bias Against Women*, REUTERS (Oct. 9, 2018 11:12 PM), <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> [https://perma.cc/99QW-EPWX]. It did so because of bias that existed in ten years of hiring data used to train the AI. *Id.* Several large companies have developed or are developing similar tools, though none have reportedly experienced the same kind of discrimination due to training data bias. *Id.*

172. See Bathaee, *supra* note 7, at 901.

173. See *supra* note 139.

174. This assumption definitionally tracks a strong-form Black Box Problem, which assumes that one cannot deterministically map inputs to outputs simply by probing the model. See Bathaee, *supra* note 7, at 906 (“Importantly, this form of black box cannot even be analyzed *ex post* by reverse engineering the AI's outputs.”).

number of possible input combinations exponentially increase. Indeed, when a model bases its decisions on thousands of input parameters, it would take several lifetimes to fully determine what the model's decision boundary looks like.<sup>175</sup> Age may be outcome determinative when it is one of a few other parameters, but when it is one of thousands, it may be relevant only when certain other parameters meet certain criteria, and even then, it may not be outcome determinative in many cases. Thus, it becomes impossible to, for example, rank which input parameters are the most important to the AI model.<sup>176</sup>

This inability to understand the basis for an AI's decisions may also impair the materiality inquiry.<sup>177</sup> Human judgment may differ to such an extent from the AI's that an omitted fact may be material to a reasonable person but entirely irrelevant to an AI. Indeed, it may be that in all cases, the number of bedrooms in a house would be important to a human being that is valuing a house, but the AI may determine that in a certain zip code and given a certain threshold square footage, the number of bedrooms does not increase the accuracy of the model's valuations—if that's the case, the model may be making its decisions without considering the number of bedrooms in the house in many particular instances. Yes, a reasonable person would want to know how many bedrooms were in the house, but a trained and accurate AI model may not care at all—and maybe for good reason (because it does not make the model more or less accurate to consider that information).

A rule that hinges opinion liability on whether a material fact underlying the opinion was omitted may therefore focus on entirely spurious notions of materiality when AI is concerned.<sup>178</sup> If one cannot tell how the AI assigns

---

175. In the discrete case, meaning that the input space consists of non-continuous inputs such as a finite set of integers, the massive input space is a matter of combinatorics, as the number of possible input combinations potentially multiply, creating exponentially larger possible inputs as input parameters are added. In the continuous case, such as when the inputs are real numbers (approximated by floating-point numbers of a fixed bit size in the case of most computers), the input space, while discrete in the sense that the possible inputs for each parameter is bounded by the precision of the floating-point numbers used, becomes unfathomably massive. In other words, in almost any real-world case, it is simply not possible to try all of the possible input combinations to determine effects on outputs.

176. See Bathaee, *supra* note 7, at 906.

177. Materiality here refers to the legal test, which asks whether a stated or omitted fact was important to a decision. See *supra* note 36.

178. The power of AI comes from pattern recognition, and sometimes the value of the AI is that it can recognize patterns that are not intuitive or perceptible to humans. If one could simply examine the AI to determine what is most material to it, one could write a determinative algorithm to perform the AI's task—it would be traditional software. See Rudina Seseeri, *The Problem with "Explainable" AI*, TECHCRUNCH (June 14, 2018), <https://techcrunch.com/2018/06/14/the-problem-with-explainable-ai/> [<https://perma.cc/QE8R-7VU6>] ("Part of the

weights to parameters and patterns in particular contexts, there may be no way to prove that the omitted information was an important or relevant part of the AI's decision.<sup>179</sup>

#### F. AI AS AN OPAQUE EXPERT

It is not uncommon for human experts to use intuition or judgment to render opinions.<sup>180</sup> A bank that hires a valuation expert to determine the value of complex derivatives may not have much insight into aspects of the expert's opinion that are based on his experience or judgment. Experts with technical or scientific expertise may rely on mathematics that would require the lay person years of study to understand. In such cases, the expert is functionally a Black Box to those who may rely on his opinion.<sup>181</sup> There is, however, a notable difference. Humans can attempt to explain the bases for their opinions, and in some cases, even the principles and assumptions underlying their opinion.<sup>182</sup>

---

advantage of some of the current approaches (most notably deep learning), is that the model identifies (some) relevant variables that are better than the ones we can define, so part of the reason why their performance is better relates to that very complexity that is hard to explain because the system identifies variables and relationships that humans have not identified or articulated. If we could, we would program it and call it software.”); *see also* 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/QNG2-XSYJ>] (“Information from the vehicle’s sensors goes straight into a huge network of artificial neurons that process the data and then deliver the commands required to operate the steering wheel, the brakes, and other systems. The result seems to match the responses you’d expect from a human driver. But what if one day it did something unexpected—crashed into a tree, or sat at a green light? As things stand now, it might be difficult to find out why. The system is so complicated that even the engineers who design it may struggle to isolate the reason for any single action. And you can’t ask it: there is no obvious way to design such a system so that it could always explain why it did what it did.”).

179. *See* Bathaee, *supra* note 7, at 906.

180. *See, e.g.*, Daniel Susskind, *AlphaGo Marks Stark Difference between AI and Human Intelligence*, FIN. TIMES (Mar. 21, 2016), <https://www.ft.com/content/8474df6a-ed0b-11e5-bb79-2303682345c8> [<https://perma.cc/QZ7F-U69Z>] (“When researchers sat down with grandmasters and asked them to explain how they played such fine chess, the answers were useless. Some players appealed to ‘intuition,’ others to ‘experience.’ Many said they did not really know at all.”).

181. Of course, humans have developed heuristics to assess a person’s credibility and trustworthiness. *See* Knight, *supra* note 179 (“Sure we humans can’t always truly explain our thought processes either—but we find ways to intuitively trust and gauge people.”).

182. Human experts or even institutions can also signal credibility or authority, thus provoking epistemic deference. *See* M. Neil Browne & Ronda R. Harrison-Spoerl, *Putting Expert Testimony in Its Epistemological Place: What Predictions of Dangerousness in Court Can Teach Us*, 91 MARQ. L. REV. 1119, 1132–33 (2008) (“When the court hears the testimony of an ‘expert,’ especially someone recognized as a ‘scientific expert,’ the jury may be overly impressed by the credentials presented and terminology used by this individual, hindering the jury’s ability to

Indeed, it is not unusual for a human expert to explain their methodology by providing the broad strokes of the basis for the opinion.<sup>183</sup> In some cases, formalized mathematical or scientific models can be explained by analogy to something lay people can understand.<sup>184</sup> Human experts can also provide rankings of what they deemed to be most important to their decisions.

Consider a judge. While true that a judge may make a legal decision based on their intuition or expertise, they will also be able to provide a justification for their decision.<sup>185</sup> It may be the case that factors frequently used by

---

fully understand and evaluate the evidence presented by the expert.”). Such signaling may be nothing more than an implicit appeal to the authority of the person or institution that stated a particular opinion, which may eliminate any inquiry into the underlying rationale for the cited opinion. *Cf.* *Rostker v. Goldberg*, 453 U.S. 57, 112 (1981) (Marshall, J., dissenting).

183. This is precisely what an expert witness attempts to do when explaining a scientific opinion to a jury. In most cases, the jury cannot directly evaluate the expert’s scientific analysis but will instead determine whether the expert appears to be credible on the subject—whether he is to be believed. *See* H.L.A. HART, *ESSAYS ON BENTHAM: STUDIES IN JURISPRUDENCE AND POLITICAL THEORY* 261–62 (1982). The essays state:

To be an authority on some subject matter a man must in fact have some superior knowledge, intelligence, or wisdom which makes it reasonable to believe that what he says on the subject is more likely to be true than the results reached by others through their independent investigations, so that it is reasonable for them to accept the authoritative statement without such independent investigation or evaluation of his reasoning.

*Id.* In other words, the jury will believe the person and therefore the proposition. *See* Scott Brewer, *Scientific Expert Testimony and Intellectual Due Process*, 107 YALE L.J. 1535, 1583 (1998) (“Where S is some speaker offering testimony that p and H is a hearer of that testimony, it is the distinction between H’s believing that p and H’s believing S that p.”).

184. An expert that explains his view with analogies or simplifications often implicitly reveals what facts were material to his opinion. This is because to draw an analogy—that is, to engage in any analogical reasoning—one must generally make determinations as to which facts in one context are similar or different from those in an analogous context. *See* Frederick Schauer & Barbara A. Spellman, *Analogy, Expertise, and Experience*, 84 U. CHI. L. REV. 249, 253 (2017) (“[T]here remains a core position according to which the first move in the analogical process is the recognition of a relevant similarity between some previous set of facts and the set of facts that now calls for decision.”); LARRY ALEXANDER & EMILY SHERWIN, *DEMISTIFYING LEGAL REASONING* 76–83 (Cambridge ed. 2008) (“Similarities are infinite; therefore some rule or principle is necessary to identify important similarities.”); *cf.* RICHARD A. POSNER, *THE PROBLEMS OF JURISPRUDENCE* 91 (Harvard ed. 1990) (“A set of cases can compose a pattern. But when lawyers or judges differ on what pattern it composes, their disagreement cannot be resolved . . . by an appeal to an intuitive sense of pattern.”).

185. The requirement to write a legal opinion allows for a system where analogical reasoning is possible—otherwise, cases cannot look backwards to compare their reasoning about the facts before them to the facts and reasoning from previous cases. James Boyd White, *What’s an Opinion For?*, 62 U. CHI. L. REV. 1363, 1368 (1995) (“The judicial opinion is a claim of meaning: it describes the case, telling its story in a particular way; it explains or justifies the result; and in the process it connects the case with earlier cases, the particular facts with more general concerns. It translates the experience of the parties, and the languages in which they

laypersons, such as moral judgments or life experiences, may not be explicit in such a decision,<sup>186</sup> but many of the important facts of the case—akin to parameters fed into an AI model—will be identified and some rank order can be discerned from the opinion. Even in the case of non-dispositive and multi-factor tests, legal opinions often provide insight into the factors and facts most important to the analysis. In any event, the number of factors in most cases are often few enough that crude empirical or statistical analysis will often provide insight as to what factors were dispositive.<sup>187</sup>

For the first time, AI presents us with the power of the expert but at the expense of transparency.<sup>188</sup> Common legal heuristics, such as witness examination or the use of contemporaneous evidence to determine intent, simply will not work when AI is involved.<sup>189</sup> The AI may have decision-making agency, so the intent of its user or creator may not matter.<sup>190</sup> In most cases, there will simply be no relevant human intent to apply an intent heuristic to.<sup>191</sup>

#### IV. A FRAMEWORK FOR AI OPINION LIABILITY

##### A. BETTER FACTUAL HEURISTICS FOR AI OPINION LIABILITY

If traditional heuristics such as intent may not work with AI, then a new set of factual heuristics is needed for liability—heuristics tailored to machine-learning models, not to human beings. The liability question should not turn on what the creator or user of the AI intended, but instead on how the creator or user of the AI trained the AI,<sup>192</sup> what data was used, what biases in the data

---

naturally speak of it, into the language of the law, which connects cases across time and space; and it translates the texts of the law—the statutes and opinions and connotational provisions—into the terms defined by the facts of the present case.”).

186. Some studies have shown that judges use cognitive decision-making processes no different than laypersons when making decisions and therefore fall prey to cognitive biases, just as laypersons do. See Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 829 (2001) (“Our study demonstrates that judges rely on the same cognitive decision-making process as laypersons and other experts, which leaves them vulnerable to cognitive illusions that can produce poor judgments. Even if judges have no bias or prejudice against either litigant, fully understand the relevant law, and know all of the relevant facts, they might still make systematically erroneous decisions under some circumstances simply because of how they—like all human beings—think.”).

187. See, e.g., Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions*, 156 U. PENN. L. REV. 549, 585 (2008) (empirically determining that the first and fourth fair use factors in copyright were the most important).

188. See *supra* Section III.B.

189. See *supra* Sections III.B & III.D.

190. See *supra* Sections III.B & III.D.

191. See *supra* Section III.E.

192. See *infra* Section IV.A.2.

were, or could have been, detected, and most importantly, how the AI was tested, validated, and deployed.<sup>193</sup> This Section discusses some factual heuristics that will be useful in assessing liability, particularly because they allow us to make some basic determinations about the intent of the AI's user or creator.

To be sure, because of the Black Box Problem, each of these heuristics may be as ineffective as a traditional scienter heuristic, because the AI user or creator's intent or conduct may be completely decoupled from the AI's decision-making process.<sup>194</sup> These heuristics, however, get at the heart of whether humans using or deploying the AI should be responsible for the AI's conduct. These heuristics only go so far, but they are the beginning of any analysis, even if they are ultimately not dispositive.

### 1. *Deference and Autonomy*

The first heuristic for liability should be the degree of autonomy the AI was given and how much the AI's opinion was relied upon.<sup>195</sup> Just as one would assign a high degree of autonomy to a trusted agent, the user or creator of AI may defer to powerful and accurate AI.<sup>196</sup> The key question is whether deference to an AI's opinion was reasonable under the circumstances. The threshold factual question will therefore often be the extent to which the creator or user of AI deferred to the AI's decisions.

Here, deference refers to reliance on the AI opinion. For example, if a manufacturer relies on an AI's assessment of product safety without any human intervention or check, it may be a telltale sign that the user or creator of the AI believed that the AI was adequately tested and was worthy of deference. In cases where a high degree of deference is not justified, such as when the AI has not been adequately tested or lacked sufficient unbiased training data, it will be a basis upon which to assign liability.<sup>197</sup> That is, deference or reliance on AI without any human supervision may be evidence that the creator or user of the AI fell below a given standard of care; this will

---

193. *See infra* Section IV.B.

194. *See supra* Sections III.B, III.D, & III.E.

195. *See* Bathaee, *supra* note 7, at 936.

196. This notion is similar to the doctrine of negligent supervision, which posits that an employer should be responsible for a failure to exercise ordinary care in supervising an employee. *See* Jackson v. Ivory, 120 S.W.3d 587, 598 (Ark. 2003).

197. The question will often be whether the degree of supervision fell short of a standard of ordinary care, just as in the case of a negligent supervision action. *See supra* note 196.

be the case when there is a great amount of uncertainty as to how the AI is making its decisions or as to how it will perform in the real world.<sup>198</sup>

Examining deference and autonomy provides vital context in AI opinion cases. Deference to an AI's medical diagnosis may not be justified, even if it is accurate more than half of the time. The risk of loss if the AI fails to properly diagnose a patient will be too high in the case of some patients. Deference when the model is less than 50% accurate at diagnosing a disease, however, may be perfectly adequate if what is being diagnosed is a relatively benign malady, such as a cold.

A focus on deference and autonomy immediately converts a technical problem—the opacity of an AI model<sup>199</sup>—into a classic question of fact that a judge or jury can assess using battle-tested legal constructs, including the rules of evidence and standards of care from the law of negligence.<sup>200</sup> Of course, in many cases, the creator or user of the AI may subjectively and reasonably believe that deference was justified, and they may be wrong because of the Black Box Problem. In such cases, the deference and autonomy heuristic may not be dispositive. In other words, it may not be appropriate to excuse the creator or user of the AI from liability simply because they could not foresee the AI's decision boundary or the effects of the AI's opinions.

## 2. *Training, Validation, and Testing*

The next useful factual heuristic will be a focus on the training, validation, and testing of the AI model. Most of the most powerful AI are built on machine-learning algorithms that learn from data.<sup>201</sup> The most important question when examining such models is whether they “generalize,” meaning whether they have seized on patterns that generally exist in a particular sort of dataset.<sup>202</sup>

---

198. This is similar to the situation where an employee's history of conduct alerts an employer to a potential risk of wrongful conduct by the employee. *See Leftwich v. Gaines*, 521 S.E.2d 717, 726 (N.C. 1999) (noting that some cases find liability where “the employee's wrongdoings were forecast to the employer and took place while working”).

199. *See supra* note 178.

200. *See supra* note 107.

201. *See supra* note 113.

202. *See* GOODFELLOW ET AL., *supra* note 113, at 107 (“The central challenge in machine learning is that our algorithm must perform well on new, previously unseen inputs—not just on those which our model was trained. The ability to perform well on previously unobserved inputs is called generalization.”); YASER S. ABU-MOSTAFA ET AL., *LEARNING FROM DATA* 39–40 (2012) (noting that generalization, “a key issue in learning,” is the degree of error on data not used to train a model—that is, on “out of sample” data).

A model may be trained to fit too closely to the data upon which it was trained.<sup>203</sup> This concept is referred to as overfitting.<sup>204</sup> The easiest example of overfitting would be memorizing a math textbook and then taking a final exam with problems you have never seen before. Memorizing the textbook without understanding what is in it will not get the student very far. In such a case, the student has overfit to the material in the textbook but is not capable of generalizing based on the data they have studied.<sup>205</sup>

Patterns in data may also be spurious, meaning that a generalization from those patterns may not assist with new data inputs.<sup>206</sup> In that case, the model may have been trained to make very crude distinctions, which is what may be causing the inaccuracy. It may provide some baseline accuracy to predict mile run times based on the height of a runner—and perhaps it will work well at the extremes of the height distribution—but it will generally not provide a very good model for differentiating among close cases (e.g., two runners with similar heights in the middle of the distribution). The error rate will be too high for the model to generalize in a meaningful way.

It is important to note that any data-driven mathematical or statistical model will only work if there are patterns in the underlying data used to train them. Patterns in a particular period of stock market returns may not hold in the future when market dynamics and fundamentals change. Patterns in the medical records of certain genetically similar patients may not exist at all in others who do not share any genetic similarity.

To AI researchers, mathematicians, and economists, this notion is a familiar one—it is often referred to as the “No Free Lunch Theorem.”<sup>207</sup> Put

---

203. In such a case, the model has overfitted to the training data—that is, the model correctly predicts training data, but it has a high error rate on data it has not yet seen. *See id.* (“Overfitting occurs when the gap between the training error and test error is too large.”).

204. *Id.*

205. *See* ABU-MOSTAFA ET AL., *supra* note 184, at 119 (“Overfitting is the phenomenon where fitting the observed facts (data) well no longer indicates that we will get a decent out-of-sample error, and may actually lead to the opposite effect. You have probably seen cases of overfitting when the learning model is more complex than is necessary to represent the target function. The model uses its additional degrees of freedom to fit idiosyncrasies in the data (for example, noise), yielding a final hypothesis that is inferior.”).

206. This corresponds to underfitting—when there is no learnable pattern in the training data other than, perhaps, a crude pattern with little predictive power. *See* GOODFELLOW ET AL., *supra* note 95, at 108 (“Underfitting occurs when the model is not able to obtain a sufficiently low error value on the training set.”).

207. *See* GOODFELLOW ET AL., *supra* note 113, at 95 (“The no free lunch theorem for machine learning states that, averaged over all possible data generating distributions, every classification algorithm has the same error rate when classifying previously unobserved points. In other words, in some sense, no machine learning algorithm is universally any better than any other.”).



simply, if an algorithm performs well on a particular dataset, it does so at the expense of performing poorly on another.<sup>208</sup> It essentially posits that there is no uniform set of patterns across all possible datasets.<sup>209</sup> In other words, AI models—or any mathematical or statistical model, for that matter—must fit to the data upon which they have been trained. If that data contains patterns that are not in other datasets, the model will not be effective in making predictions when it is shown new data.<sup>210</sup>

This is why it matters how the AI has been trained, validated, and tested.<sup>211</sup> To begin with, it is important to note whether best practices were followed when training the AI. For example, it is worth asking whether the dataset had been separated into a subset for training and a subset for testing or validation.<sup>212</sup> This prevents the AI model from knowing any information about the test data.<sup>213</sup> This allows a more accurate determination of whether the AI model is appropriately generalizing.<sup>214</sup> The model is trained on one subset of data, and if the accuracy rate holds on the testing subset, then the model has successfully been trained to recognize patterns in the data.<sup>215</sup> Conversely, if the accuracy rate is high in training but is poor in testing, then the model may be overfitting on the training data.<sup>216</sup>

Moreover, because models require tuning during the training process, using subsets of data for validation during training provides further assurances that none of the test set information was used to train the model.<sup>217</sup> There are many best practices for training machine-learning models, which are beyond

---

208. *See id.*

209. *See id.*

210. *See id.* at 115 (“This means that the goal of machine learning research is not to seek a universal learning algorithm or the absolute best learning algorithm. Instead, our goal is to understand what kinds of distributions are relevant to the ‘real world’ that an AI agent experiences, and what kinds of machine learning algorithms perform well on data drawn from the kinds of data-generating distributions we care about.”).

211. Validation involves creating a subset of the training data and holding it out for tuning of the model. *See id.* at 118. Testing would be performed on a dataset that was not used for training or tuning. *See id.*

212. *See id.*

213. *See id.* at 119 (“It is important that the test examples are not used in any way to make choices about the model, including its hyperparameters. For this reason, no example from the test set can be used in the validation set.”).

214. *See id.*

215. *See id.* at 109–10.

216. *See id.*

217. *See id.* at 119 (“More frequently, the setting must be a hyperparameter because it is not appropriate to learn that hyperparameter on the training set. This applies to all hyperparameters that control model capacity. If learned on the training set, such hyperparameters would always choose the maximum possible model capacity, resulting in overfitting.”).

the scope of this Article, and many of these best practices are likely to change, but the extent to which the model was trained, validated, and tested according to some relevant standard of care will be imperative for assigning liability.

It is important to note that the training, validation, and testing heuristic is closely connected with the deference and autonomy heuristic. If a model is poorly validated and tested, then it may not have been reasonable to provide the model autonomy or deference.<sup>218</sup>

### 3. *Constraint Policies and Conscientiousness*

The extent to which constraints are provided to the AI model will also be an important heuristic. A defendant that takes great care to prevent opinions based on improper or spurious bases will be less culpable than one who provides the AI model an unbounded degree of freedom to achieve a particular accuracy, result, or reward.<sup>219</sup> The existence or lack of constraints says something about the creator or user of the AI's conscientiousness.<sup>220</sup> Did they attempt to exercise some care when deploying the AI? Because conscientiousness is something that must be incentivized, it makes sense to provide safe harbors for those who impose extensive constraints on AI decision making.<sup>221</sup>

Without a focus on constraints imposed on the AI, existing liability rules may perversely incentivize reckless behavior. If the imposition of constraints on the AI does not mitigate liability, it may only serve to establish that the user of the AI was aware of a particular risk.<sup>222</sup> For example, an AI model tasked with making hiring decisions that includes software safeguards against race or gender-based discrimination may prove that the defendant was aware of the

---

218. *See supra* Section IV.A.1.

219. Indeed, the failure to impose reasonable safeguards may allow the inference of recklessness or other forms of scienter, such as willful blindness. *See* Bathaee, *supra* note 7, at 933–34.

220. *See id.* at 933.

221. Some commentators have argued that safe harbors are an effective means of incentivizing laudable corporate conduct. *See* Elizabeth F. Brown, *No Good Deed Goes Unpunished: Is There a Need for a Safe Harbor for Aspirational Corporate Codes of Conduct*, 26 YALE L. & POL'Y REV. 367, 402 (2008) (“How can the law be amended in order to encourage businesses to seek to achieve higher standards of behavior than the bare legal minimum? One possible solution might be for states or the federal government or both to enact laws that limit the ability of corporations to be sued if they make good faith efforts to achieve aspirational standards of behavior but fails as long as their conduct still met the legal standards embodied in statutes, regulations, and the common law.”).

222. *Cf. id.* at 401–02 (“[T]he market forces may encourage some businesses to try to get as close to the line of what is legally permissible behavior in order to maximize profits. The danger with that sort of behavior is that businesses frequently misjudge where the line is and end up owing large penalties and legal bills for violations of the law.”).

risk that such discrimination would occur if the AI's opinion was relied upon. In other words, by imposing safeguards, the creator or user acknowledges consciousness of risk.

If constraints do not mitigate liability, then it may be better for the person using the AI not to impose any constraints on the AI at all and argue that any harm arising from the AI's opinions were completely unintended and perhaps unforeseeable. In other words, a conscientiousness heuristic that relieves a conscientious actor from liability may not only help assess the degree of culpability for deploying the AI, it would also incentivize reasonable care and risk mitigation.<sup>223</sup>

#### B. EXAMINING DATA BIAS, NOT DECISION BASIS IN OMISSIONS CASES

The Supreme Court has announced a rule in the securities law context that allows courts to examine the basis for an opinion statement only when there is a claim that material information was omitted from the opinion statement.<sup>224</sup> Such a rule, however, would be ineffective in many AI opinion contexts. If the AI is opaque, meaning that it suffers from the Black Box Problem, it will likely be nearly impossible to determine the basis for the AI's opinions.<sup>225</sup> In many cases, even a weak rank ordering of input parameters will not be possible.<sup>226</sup>

The basis of an opinion in an omissions case allows a court or factfinder to determine whether the speaker of the opinion was justified in holding the opinion.<sup>227</sup> If, for example, the speaker of the opinion considered information that undermined the opinion but failed to disclose that information, it may be fair to hold him liable for the omission.<sup>228</sup> The question of liability in such a circumstance will turn on whether it was reasonable of the person to hold the opinion notwithstanding the inconsistent or contradictory information in the speaker's possession.<sup>229</sup> The rule also safeguards against an entirely reckless or uninformed opinion.<sup>230</sup> In cases where the speaker never made any investigation as to any of the relevant facts or ignored gaping deficiencies in information required for the opinion, it may also be entirely fair to hold them

---

223. *See id.* at 402.

224. *See Omnicare, Inc. v. Laborers Dist. Council Constr. Indus. Pension Fund*, 135 S. Ct. 1318, 1327–29.

225. *See Bathaee, supra* note 7, at 916–17.

226. *See id.*

227. *See supra* note 78.

228. *See supra* note 47.

229. *See Omnicare*, 135 S. Ct. at 1327–29.

230. *See supra* note 63.

liable for the statement.<sup>231</sup> A reckless opinion in such cases is the analog of a false statement.<sup>232</sup>

In the AI context, the models will likely be trained on large amounts of information. If the model has been properly trained, validated, and tested, the question will seldom be about whether any factual investigation was done to justify the AI's opinion.<sup>233</sup> The question will instead be about whether the model was trained with data that implicitly expressed some form of bias.<sup>234</sup>

To AI systems, data bias can be the same as blindness.<sup>235</sup> If AI is trained with data that contains a prevalent pattern, it will undoubtedly leverage that pattern in its decision making—it may even do so with too much emphasis.<sup>236</sup> It may also be the case that the underlying data narrowly covers only a subset of possible data points and leads the AI to make decisions in all contexts as it would in a narrow unrepresentative one.<sup>237</sup>

Consider an AI designed to provide an opinion on recidivism in prisoners being considered for parole. If the underlying data is trained on a population of past inmates that were subject to widespread racial discrimination by the police while released on parole, then the data may use race as a proxy for its decision making.<sup>238</sup> Of course, one would say that if race is not included as an input to the model, then it cannot be considered. But even then, a model may be capable of using proxies for race, such as their economic backgrounds or even zip codes.<sup>239</sup> Because the data bias is so overwhelming, other factors will correlate with the bias and infiltrate the AI's opinions.<sup>240</sup>

---

231. *See id.*

232. This is in part because the implicit statement that the speaker has a basis for his opinion has proven false. *See supra* note 47.

233. The data used to train the model is the definition of factual information; it is the very basis upon which the model is built. In other words, with respect to AI that learns from data, if there were no foundation, there would be no data and therefore no AI model.

234. *See supra* Section IV.A.

235. This is because the AI cannot learn from data it never observed during training. If the out-of-sample data is too different from the training data the model has encountered, the model may overfit on the training data and fail to generalize with respect to the new data it has not yet encountered. *See supra* note 205.

236. *See id.*

237. This phenomenon arises from sampling bias. It is axiomatic that “[i]f the data is sampled in a biased way, learning will produce a similarly biased outcome.” *See* ABU-MOSTAFA ET AL., *supra* note 202, at 172.

238. *See* Bathaee, *supra* note 7, at 920.

239. *See id.*

240. The model may be “snooping” at the prohibited data, meaning that the data has made its way into the model even though it was not explicitly provided. Accordingly, it is generally assumed that “[i]f a data set has affected any step in the learning process, its ability

The question is thus (a) whether there exists a bias in the training data, (b) whether the bias is strong, and (c) whether other input parameters correlate with that bias. If the bias is improper, the opinion may also be improper.<sup>241</sup> These three determinations can be made even when the model exhibits the Black Box Problem. There are myriad mathematical and statistical methods that can demonstrate the existence and strength of biases in underlying data, including correlation and covariance among variables.<sup>242</sup> Such a showing, coupled with a showing that such bias is legally improper, may be sufficient in some cases to establish opinion liability.

Take for instance the example of an AI designed to diagnose a particular medical condition based on the existence of a particular set of genetic traits. If the AI fails to diagnose a large number of patients and those patients relied on a negative diagnosis, then the question will be whether the AI was properly trained and tested.<sup>243</sup>

Assuming that it was properly tested and trained, the next question is whether bias in the data would explain why the AI's accuracy decreased in real-world application. A showing that all of the data came from the health records of a large interrelated population of patients with some other common genetic trait, then that bias may be the problem with the model.<sup>244</sup> It may be that the diagnosis is highly accurate in that biased population, but not so among the general population.<sup>245</sup>

The question for liability is therefore whether that bias in the underlying data was detectable, and if so, how strong that bias was.<sup>246</sup> It may be obvious

---

to assess the outcome has been compromised." See ABU-MOSTAFA ET AL., *supra* note 202, at 173.

241. *See id.*

242. *See, e.g.*, GOODFELLOW ET AL., *supra* note 113, at 119–21.

243. *See supra* Section IV.A.

244. *See* Robert David Hart, *If You're Not a White Male, Artificial Intelligence's Use in Healthcare Could Be Dangerous*, QUARTZ (July 10, 2017), <https://qz.com/1023448/if-youre-not-a-white-male-artificial-intelligences-use-in-healthcare-could-be-dangerous/> [<https://perma.cc/Q9WK-PJLP>] ("The highly selective nature of trials systematically disfavor women, the elderly, and those with additional medical conditions to the ones being studied pregnant women are often excluded entirely. AIs are trained to make decisions using skewed data, and their results will therefore factor the biases contained within. This is especially concerning when it comes to medical data, which weighs heavily in the favor of white men.").

245. *See id.*

246. In medical cases, privacy constraints may compound the Black Box Problem's barriers to determining how the AI made its decision. This means that data biases may be especially difficult to detect in the medical context. *See id.* ("AI systems often function as black boxes, which means technologists are unaware of how an AI came to its conclusion. This can make it particularly hard to identify any inequality, bias, or discrimination feeding into a particular decision. The inability to access the medical data upon which a system was trained—

that drawing from a narrow population of patients from which to train the AI model was an error—well below the standard of care for diagnosis. In such a case, liability would fairly attach. And in the case where it can be shown that the creator of the AI knew about the data bias but deployed the model anyway, the omission of the data bias would also be grounds for liability.<sup>247</sup>

C. HIGH RISK / HIGH VALUE APPLICATIONS AND STRICT LIABILITY

There may be some cases where it is never safe to defer to an AI opinion, and in those cases, strict liability may be appropriate.<sup>248</sup> Medical applications will likely be riddled with such circumstances. Indeed, it may be that it will always be reckless to leave cancer diagnosis entirely to AI.<sup>249</sup> The same can be said about AI designed as weapons for police or military functions.<sup>250</sup> It may always be unreasonable to rely on an armed robot powered by an AI system for crowd control—indeed, the AI's opinion as to whether a person poses a

---

for reasons of protecting patients' privacy or the data not being in the public domain—exacerbates this.”).

247. See *supra* Part II.

248. Some commentators have pointed out that an important precondition for a strict-liability regime to be viable is adjudicability, meaning that there is a defined set of facts or conduct to which strict liability will apply. See James A. Henderson, Jr., *Why Negligence Dominates Tort*, 50 UCLA L. REV. 377, 391 (2002) (“For disputes under strict liability to be adjudicable, the boundaries of the liability system—the descriptions of harm-causing activities for which the system holds enterprises strictly responsible—must be relatively specific and must not depend on fact-sensitive risk-utility calculations.”). Because an AI's reasoning or conduct may be unpredictable due to the Black Box problem, it is the application of the AI that must be the defined trigger for a strict liability regime to be viable, not the particular conduct or risks involved in an AI's deployment.

249. Section 402A of the Restatement (Second) of Torts, which was drafted by William Prosser, was one of the first significant movements toward strict liability for unreasonably dangerous products, and most states eventually adopted some form of the rule stated in the Restatement. James A. Henderson, Jr. & Aaron D. Twerski, *A Proposed Revision of Section 402A of the Restatement (Second) of Torts*, 88 CORNELL L. REV. 1512, 1512–13 (1992); see also George L. Priest, *The Invention of Enterprise Liability: A Critical History of the Intellectual Foundations of Modern Tort Law*, 14 J. LEGAL STUD., 461, 512, 518 (1985). The rule has been applied to tobacco and cigarette products that cause cancer—see, for example, *Hearn v. R.J. Reynolds Tobacco Co.*, 279 F. Supp. 2d 1096, 1103 (D. Ariz. 2003)—and it is certainly conceivable that the rule would also apply to an AI diagnosis product that improperly diagnoses (or fails to diagnose) cancer.

250. See, e.g., Joseph A. Page, *Of Mace and Men: Tort Law as a Means of Controlling Domestic Chemical Warfare*, 57 GEO. L.J. 1238, 1258 (1969) (arguing that spray weapons could give rise to liability “[w]hen the plaintiff is the intended target and suffers more than transitorily disabling harm as a result of a construction or design defect in the spray or inadequate warnings and instruction in its use”).

threat may be 98% accurate, but a 2% failure rate may result in death or injury. These are high risk applications, which may justify a strict liability regime.<sup>251</sup>

There is another class of cases where deference to an AI opinion may be improper—high value applications. For example, AI opinions are unlikely to be a fair replacement for juries or judges. Constitutional and democratic norms may require a human being to make factual determinations at trial.<sup>252</sup> A judge that entirely delegates legal decision making to an AI system that has been trained to predict how they would decide cases may also be doing so improperly.<sup>253</sup> Sometimes, a human check is required, even if it is a slight one.<sup>254</sup> This may be because due process requires it or it may be because the AI is opining on a function that as a society we would prefer humans to perform.

---

251. See Bathaee, *supra* note 7, at 931 (arguing that strict liability may be appropriate in some limited cases but that a blanket strict liability rule for AI would not be appropriate); cf. Yesha Yadav, *The Failure of Liability in Modern Markets*, 102 VA. L. REV. 1031, 1039 (2016) (arguing against strict liability for determinative algorithms). Strict liability in tort for abnormally dangerous conduct is, for the most part, circumscribed in application, as it applies only to a short list of abnormally dangerous activities, which have not been significantly expanded over the years. See John C.P. Goldberg & Benjamin C. Zipursky, *The Strict Liability in Fault and the Fault in Strict Liability*, 85 FORDHAM L. REV. 743 (2016) (“Each of the three Restatements of tort law has recognized a special domain of strict liability under the labels ‘ultrahazardous’ or ‘abnormally dangerous’ activities. This domain is quite narrow, applying only to injuries caused by blasting, escaped wild animals, bursting reservoirs, and a few other activities. Plaintiffs’ lawyers, courts, and commentators have at times suggested that the particular form of liability that attaches to abnormally dangerous activities should occupy more of the torts landscape. But no such expansion has occurred.”).

252. See Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, 105 GEO. L.J. 1147, 1178 (2018) (“Yet the nondelegation doctrine, still a fixture in American constitutional and administrative law, places some theoretical limits on those delegations, which must, for example, be accompanied by an intelligible principle. Although this doctrine has long accepted even broad delegations of authority to administrative agencies, the law has always assumed that the recipient of that authority would be a human being, such as an officer of the United States, or on occasion, a private individual or group of individuals. . . . Yet if government actions should be undertaken by humans, then delegation to autonomously learning machines could potentially transfer governmental power outside the bounds that the Constitution permits.”).

253. See *id.*

254. In Federal Courts and tribunals, the constraint may be a result of Article III of the United States Constitution, which vests the judicial power in the courts, which consist of (human) judges appointed for life. See U.S. CONST. art. III. Indeed, certain cases and controversies cannot be decided even by humans who are not appropriately appointed, and do not operate, within the requirements of Article III. See *N. Pipeline Constr. Co. v. Marathon Pipe Line Co.*, 458 U.S. 50, 87 (1982) (“We conclude that . . . the Bankruptcy Act of 1978, has impermissibly removed most, if not all, of ‘the essential attributes of the judicial power’ from the Art. III district court, and has vested those attributes in a non-Art. III adjunct.”); cf. *Freytag v. Commissioner*, 501 U.S. 868, 870 (1991) (holding that the appointment of special trial judge in Article I tax court did not violate separation of powers).

Strict liability may make sense for high-value applications. It may also make sense to bar AI from high-value applications entirely.

It is notable that strict liability has generally been rejected when opinion statements are involved. Indeed, in the Securities Act of 1933 Act context, false statements in prospectuses give rise to strict liability, but even in that context, courts have imposed a scienter-like requirement that the opinion statement be both subjectively and objectively false,<sup>255</sup> meaning that the speaker of the opinion intended to mislead with the opinion or did not genuinely believe the opinion.<sup>256</sup>

To be sure, there are several good reasons to reject strict liability in the opinion context. Opinions are often based on contradictory or incomplete information, so it is generally not enough that the speaker of an opinion know of information that contradicted his opinion to render the opinion false.<sup>257</sup> A doctor may know that some percentage of patients with a particular set of symptoms may have a completely different, perhaps more serious, diagnosis, but one would not say that the doctor should necessarily be liable if the diagnosis proves incorrect (and the alternative, more serious diagnosis turned out to be correct).

The determinative factor for liability is whether the doctor's judgment was reasonable under the circumstances.<sup>258</sup> Perhaps the probability of the alternative diagnosis was relatively low. Perhaps the patient had other characteristics that affected the relative probabilities. Or the doctor simply may have relied on his own experience to make a decision. None of these circumstances describe a doctor who has behaved improperly.<sup>259</sup> In other words, being wrong does not necessarily mean negligent or malicious—especially when individual judgment is involved.

---

255. See *supra* note 42 and accompanying text.

256. See *supra* notes 42, 58, and accompanying text.

257. See *supra* note 2.

258. This is because most malpractice cases will be negligence cases, which require some showing that the doctor breached the relevant standard of care. See *supra* note 197. For the most part, this will be a standard of care that is relative to other physicians and not the standard of care that applies to a lay person, as it may never be reasonable for a lay person to attempt to practice medicine without any acquired skill or training. See Charles R. Korsmo, *Lost in Translation: Law, Economics, and Subjective Standards of Care in Negligence Law*, 118 PENN. ST. L. REV. 285, 327 (2013) (“The illusion that skilled professionals are held to a ‘higher’ standard of care for a given activity is maintained only by ignoring the requirement that unskilled laypeople avoid the professional activity altogether.”).

259. That is, unless the doctor relying on his own experience does so without having adequate experience. See, e.g., *Andersen v. Khanna*, 913 N.W.2d 526, 537 (Iowa 2018) (“We conclude the district court erred when it found, as a matter of law, there is no duty to disclose personal characteristics, such as experience and training, under Iowa law.”).



A strict liability rule for a wrong diagnosis opinion in this hypothetical case would be oppressive. It would be impossible for doctors to operate under such conditions, because they practice in a field where incomplete or probabilistic information is sometimes all that is available.<sup>260</sup> It would also incentivize the doctor to defensively attempt to rule out improbable diagnoses, slowing down treatment and perhaps in many cases increasing the costs.<sup>261</sup> The net effect would be to cripple the exercise of judgment by an expert—but most of the time, the expert’s judgment and trained intuition is precisely what a patient seeks from the expert.

AI opinions are differently situated. Although AI shares some characteristics with a trained human expert, such as the ability to make judgments based on intuition, experience, and training, an AI’s incentives to be thorough do not change with liability rules. For example, an AI will likely not practice defensive medicine. That is, AI operating in a strict liability regime will make the same predictions—based on data—regardless of whether it will be subject to strict liability. This eliminates some of the major policy problems with strict liability.

And the person deploying the AI can decide whether the potential of being strictly liable is worth it.<sup>262</sup> Under a strict liability regime, a person or company deploying the AI may think twice before using the AI to autonomously make valuation decisions for high-value assets, but may decide that the AI’s opinions on low value items are worth the tradeoff of being strictly liable. Notably, the focus will be on the decision to use the AI, not on how the AI arrives at its opinions.<sup>263</sup>

---

260. Although medical malpractice or misdiagnosis is an insurable risk, which satisfies one of the preconditions for a viable strict liability regime, there can be no legitimate set of physician conduct in the ordinary course of care that could be a priori defined as prohibited. See Henderson, *supra* note 248, at 391. A diagnosis or medical test, for example, cannot be deemed to give rise to strict liability simply because in some cases it causes injury. The context in which any given medical test or diagnosis is used will vary greatly from case to case.

261. See *The Medical Malpractice Threat: A Study of Defensive Medicine*, 1971 DUKE L.J. 939, 943 (1971) (noting that in response to a ruling finding liability, “a physician may go far beyond the court-established standard by performing procedures which are neither legally nor medically required in order to guarantee that no hidden problems have been overlooked which might otherwise have become the basis of a malpractice suit”).

262. This assumes that risks are independent and ascertainable, such that the amount of insurance necessary can be determined with some regularity. See Mark A. Geistfeld, *Interpreting the Rules of Insurance Contract Interpretation*, 68 RUTGERS U. L. REV. 371, 383–91 (2015) (noting that insurable risks must be independent across policyholders for the insurer to be able to predict and distribute risk across a pool of policies).

263. This is because how the AI makes decisions may be off limits due to the Black Box Problem, so specific conduct cannot be defined as a priori subject to strict liability. Rather,

D. WHY DISCLOSURE RULES ARE LESS EFFECTIVE IN THE CASE OF AI OPINIONS

In many opinion cases, disclosure of the basis of the opinion will preclude liability because it will be difficult to contend material information has been omitted or that an affirmative statement is misleading. With respect to Black Box AI, a disclosure cannot include the basis of the opinion or even a set of material facts,<sup>264</sup> so any disclosure will be limited to characteristics of the AI, such as how it was trained.<sup>265</sup> Accordingly, this Section argues that disclosure rules will generally not be effective in the case of AI opinions.

1. *Disclosure in the Non-AI Opinion Context*

All opinion statements suffer from the risk that some piece of contradictory information was improperly weighed when the opinion was made.<sup>266</sup> There is also the risk that incentives or biases played an improper role in the decision-making process.<sup>267</sup> That risk is precisely why a disingenuously held opinion can give rise to liability—it's usually the sign of bias or some incentive contrary to providing a truthful opinion.<sup>268</sup>

One way to fix the problem is to disclose everything about the decision-making process.<sup>269</sup> A valuation opinion that explicitly states what was important to the decision-making process and why will generally be more valuable than one that simply states a conclusion.<sup>270</sup> The reason is that the person hearing the opinion can evaluate the soundness of the opinion and determine what aspects of the opinion's model of the world are more or less correct.

---

only the specific use or application of the AI can be deemed subject to strict liability. *See supra* note 248.

264. *See supra* Part III.

265. *See supra* Section IV.A.2.

266. *See Omnicare, Inc. v. Laborers Dist. Council Constr. Indus. Pension Fund*, 135 S. Ct. 1318, 1328 (holding that opinion is not false simply because one can “second-guess inherently subjective and uncertain assessments”).

267. *See supra* note 67 and accompanying text.

268. *See supra* note 2.

269. *See, e.g., In re Donald J. Trump Casino Sec. Litig.*, 7 F.3d 357, 371 (3d Cir. 1993) (“[C]autionary language, if sufficient, renders the alleged omissions or misrepresentations immaterial as a matter of law.”).

270. In some cases, the opinion may be on a matter that is of such importance or complexity that a reasonable person would have expected to exercise some diligence as to the basis for the opinion before the opinion was stated. *See Omnicare*, 135 S. Ct. at 1330 (in some cases a reasonable person would not expect the opinion “to reflect baseless, off-the-cuff judgments, of the kind that an individual might communicate in daily life”).

A judicial opinion, for example, articulates not only the outcome the judge has reached, but also explains how the judge reached that outcome. A real estate appraisal will often have the most pertinent facts set out within a report. If there has been disclosure, it is simply much less likely that the person hearing the opinion has been misled.

2. *Disclosure Will Be Less Effective in the AI Context*

When AI is concerned, disclosure of all of the parameters used by the AI may be of little value. Unlike humans, AI can simultaneously weigh significantly more information,<sup>271</sup> but disclosure of what information is provided to the AI will not likely make much of a difference to a person considering the AI's decision or opinion.<sup>272</sup> For example, medical AI that evaluates 3,000 different patient characteristics is not any less opaque if those 3,000 characteristics are disclosed.

There is also no way to succinctly explain how each parameter has been weighed by the AI if the AI suffers from the Black Box Problem.<sup>273</sup> There will generally be no way of strictly rank ordering the model's inputs in terms of their effect on the ultimate opinion.<sup>274</sup> Indeed, a particular parameter may only be relevant if hundreds of others bear some threshold characteristic, and may be much less relevant when those other parameters are not above that threshold.<sup>275</sup> Was the patient's diet relevant to the diagnosis? The answer may be that it depends on a host of other factors, so is diet more important than say, liver function? There will often be no strict rank ordering.<sup>276</sup>

Assuming that the AI model suffers from the Black Box Problem, disclosure will generally not mitigate the opaqueness of the overall opinion. This is completely different than in the case where a human exercises judgment—the human can provide some explanations and provide a rough ordering of what factors were most important.<sup>277</sup> That may not be possible when AI is involved.<sup>278</sup> This is why disclosure rules are likely to be ineffective when AI is concerned.

---

271. *See supra* Section III.B.

272. *See id.*

273. *See id.*

274. *See id.*

275. This may result on a neural-layer-scale, as each artificial neuron layer is coupled with a non-linear activation function. One popular non-linearity used as an activation function is a rectifier, which passes a scaled output from the neural layer onto the next layer if the signal is above some threshold; if the signal is below some threshold, the activation function passes no signal to the next layer. *See* GOODFELLOW ET AL., *supra* note 113, at 187.

276. *See supra* Section III.B.

277. *Id.*

278. *Id.*

E. WHEN CAN YOU INFER USER OR CREATOR INTENT FROM AN AI MODEL'S OPINION?

Given that there are several factual heuristics that can still be used to evaluate culpability on the part of the person deploying the AI—(1) training, validation and testing,<sup>279</sup> (2) deference and autonomy,<sup>280</sup> and (3) constraints and conscientiousness<sup>281</sup>—is it therefore possible that one can infer scienter based on these heuristics? The answer is likely yes in many cases, but with the important caveats described in this Section:

- Data bias will be more important in omissions cases;<sup>282</sup>
- Certain applications should be subject to strict liability;<sup>283</sup> and
- Disclosure is likely irrelevant in solving the liability issue.<sup>284</sup>

With these caveats in mind, it is possible to infer a very specific form of scienter if the heuristics imply culpability. At the extreme, an AI that received complete autonomy and little supervision without adequate training or validation while operating with no a priori constraints (not supervision, but deterministic constraints in its programming) will likely imply that the person who created or deployed the AI was at least reckless for doing so.<sup>285</sup> Thus, it may be that the factual heuristics described in this Part can give rise to an inference of scienter.

F. PUTTING IT ALL TOGETHER: SCIENTER SHOULD BE SUFFICIENT, NOT NECESSARY FOR OPINION LIABILITY

The caveats noted above, however, make clear why scienter should not always be required to impose liability. It may be that a model is correctly trained and tested, provided the right amount of human supervision, and given several constraints to address potential known risks, but that a latent bias in the data caused the AI to render improper opinions.<sup>286</sup> In such a case, none of the heuristics will allow a factfinder to infer scienter. And because of the Black Box Problem, there will generally be no evidence that the person deploying the AI intended to mislead or subjectively disbelieved the AI's opinions.<sup>287</sup>

---

279. See *supra* Section IV.A.2.

280. See *supra* Section IV.A.1.

281. See *supra* Section IV.A.3.

282. See *supra* Section IV.B.

283. See *supra* Section IV.C.

284. See *supra* Section IV.D.

285. See Bathaee, *supra* note 7, at 932–38.

286. See *supra* Section IV.C.

287. See *supra* Section III.B.

In such a case, the important question is one of negligence—why did the person deploying the AI miss the data bias? Should the data bias have been studied? Should the dataset upon which the AI was trained have been described or even disclosed? All of this depends on context.

Worse yet, none of these questions should matter at all if the risk of loss is sufficiently high.<sup>288</sup> If an AI malfunction could cause hundreds of deaths, perhaps it was a poor use case for the AI and no amount of precaution, conscientiousness, or supervision should absolve the person deploying the AI of liability.<sup>289</sup>

In other words, liability should certainly attach when there is scienter or where scienter can be inferred from precise heuristics (such as those described in this Article). But it is clear that requiring scienter as a necessary element would completely insulate a wide swath of AI from liability entirely.<sup>290</sup> It would also allow AI to sanitize human conduct that would otherwise give rise to liability.<sup>291</sup>

It is important to separate AI opinions from human opinions when liability is concerned and to apply a set of specialized rules and heuristics to AI opinions. Scienter may work well for humans, but it cannot be the requirement for AI, as that would mean virtual immunity from liability when AI is involved.

## V. CONCLUSION

Opinion statements are generally not provably true or false.<sup>292</sup> They are functionally models of reality and are based on a set of facts—actual or assumed—that are considered together.<sup>293</sup> That is why the law has focused on intent-based heuristics, such as scienter, to assign liability based on opinion statements.<sup>294</sup> However, AI potentially decouples the connection between the opinion and the speaker of the opinion.<sup>295</sup> It also obfuscates the factual basis and fact weighting that is the basis for the opinion.<sup>296</sup> This obfuscation arises from AI's Black Box problem, which stems from the inherent connective and complex structure of the machine-learning algorithms used to build AI systems.<sup>297</sup> All of this means that intent will often not be inferable by simply

---

288. *See supra* Section IV.C.

289. *See id.*

290. *See supra* Part III.

291. *See id.*

292. *See supra* Part I.

293. *See id.*

294. *See id.*

295. *See supra* Part III.

296. *See id.*

297. *See id.*

examining the AI, as it would be in the case of a deterministic, instruction-based computer program.<sup>298</sup>

A new set of heuristics are necessary to determine whether a person who deploys an AI system should be responsible for the harm caused by it.<sup>299</sup> Those heuristics are more precise than conventional intent-based heuristics—they address how the model was constructed, constrained, and supervised.<sup>300</sup> These heuristics may not point towards an improper intent on the part of the AI's creator or user, but they are more precise—that is, they are heuristics that assume extensive training and testing of the AI, context-driven decision making as to the necessary level of supervision for the AI, and whether the appropriate constraints were put in place *a priori*.<sup>301</sup>

There are also contexts in which these heuristics do not strongly suggest that any person intentionally designed the AI in a manner that is at all culpable.<sup>302</sup> In these contexts, there may still be a case for liability.<sup>303</sup> When an AI is deployed in a context that has a high risk of harm or in which societal norms demand human judgment, strict liability may be appropriate—even though strict liability regimes prove problematic in the case of human opinion statements.<sup>304</sup>

It may also be the case that although the AI has been appropriately trained, deployed, and supervised, there was a significant amount of bias in the data used to train it.<sup>305</sup> In such a case, none of the scienter-like heuristics may point towards liability, but liability may nonetheless be appropriate.<sup>306</sup>

To be sure, courts must wrestle with a new set of liability heuristics as well as the significant policy judgments they implicitly come with and will need to gain significant amounts of experience with AI systems before a viable opinion liability regime emerges, but what is almost certain is that existing scienter rules for opinion liability would excuse a wide swath of AI and AI-assisted opinions from liability.<sup>307</sup> This is a result worse than a lack of regulation—it is tantamount to an unintended immunity from opinion liability entirely.

---

298. *See id.*

299. *See supra* Section IV.A.

300. *See id.*

301. *See id.*

302. *See supra* Sections IV.B–F.

303. *See supra* Section IV.F.

304. *See supra* Section IV.C.

305. *See supra* Section IV.B.

306. *See supra* Sections IV.B & IV.F.

307. *See supra* Part III.

# TARPITS: THE STICKY CONSEQUENCES OF POORLY IMPLEMENTING TECHNOLOGY-ASSISTED REVIEW

David Dowling<sup>†</sup>

## TABLE OF CONTENTS

I.	INTRODUCTION .....	172
II.	TAR'S BACKGROUND.....	175
III.	LEGAL CONSEQUENCES OF MISUSE.....	179
A.	CIVIL SANCTIONS.....	179
1.	<i>Federal Rule of Civil Procedure 37(c) Sanctions</i> .....	179
2.	<i>Spoliation Sanctions</i> .....	180
3.	<i>Sanctions for Intentional Misuse</i> .....	182
B.	OTHER LEGAL CONSEQUENCES OF TAR IN CIVIL LITIGATION .....	183
1.	<i>Administrative Agencies' Control of TAR</i> .....	183
2.	<i>TAR's Expansion of Discovery</i> .....	184
C.	SANCTIONS IN CRIMINAL PROSECUTION.....	185
1.	<i>Background Criminal Law</i> .....	185
2.	<i>Sanctions for Brady Violations</i> .....	186
3.	<i>Federal Rule of Criminal Procedure 16 Sanctions</i> .....	187
4.	<i>Sanctions for Jencks Act Violations</i> .....	190
D.	PERSONAL LIABILITY FOR MISUSE .....	191
IV.	ECONOMIC CONSEQUENCES OF MISUSE .....	192
V.	CONCLUSION.....	194

---

DOI: <https://doi.org/10.15779/Z38222R65N>

© 2020 David Dowling.

<sup>†</sup> B.S., Auburn University, 2016; J.D., Cornell Law School, 2020. I would like to thank the Honorable Jonathan W. Feldman, U.S. Magistrate Judge for the Western District of New York and Professor Keir Weyble of Cornell Law School for their help and encouragement, and for their commitment to their students. I would also like to express my appreciation to Harrison Geron and Allaa Mageid for their excellent work editing this piece. Finally, thank you to my parents for their support throughout my academic career, especially the 12-year investment in my homeschooling.

## I. INTRODUCTION

“The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”<sup>1</sup> Society is increasingly incorporating technology into more and more areas of life as the phenomenon of the Internet of Things gains momentum.<sup>2</sup> It is no secret that technology is advancing rapidly—a laptop computer today is 96% cheaper and 1,000 times better than a 1994 model.<sup>3</sup>

Law, by contrast, is an anachronistic profession, from the Socratic method of law school classes to the shelves of outdated books in law offices. Indeed, the very thought of having to adapt to technological advances may invoke visceral responses from some attorneys, particularly those who receive holiday cards from legal publishers. An affinity for the old ways, though, will not shield the profession of law from the rising tide of technological change. For example, artificial intelligence is being incorporated into the profession to postulate hypotheses, conduct legal research, and write legal memoranda.<sup>4</sup> In recent years, the legal profession has seen the rise of technology-assisted review (TAR) in discovery. Experts define TAR as “[a] process for [p]rioritizing or [c]oding a [c]ollection of [d]ocuments using a computerized system that harnesses human judgments of one or more [s]ubject [m]atter [e]xpert(s) on a smaller set of [d]ocuments and then extrapolates those judgments to the remaining [d]ocument [c]ollection.”<sup>5</sup> Although other types of TAR exist,<sup>6</sup> this Article focuses on TAR that uses supervised machine

---

1. ISAAC ASIMOV & JASON A. SHULMAN, ISAAC ASIMOV’S BOOK OF SCIENCE AND NATURE QUOTATIONS 281 (1988).

2. See Jacob Morgan, *A Simple Explanation Of ‘The Internet of Things’*, FORBES (May 13, 2014, 12:05 AM), <https://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/#1b3efdf51d09> [https://perma.cc/P2NM-93U5] (“[The Internet of Things] is the concept of basically connecting any device with an on and off switch to the internet (and/or to each other). This includes everything from cellphones, coffee makers, washing machines . . . and almost anything else you can think of.”).

3. Mark J. Perry, *Technology Has Advanced So Rapidly That a Laptop Computer Today is 96% Cheaper Than a 1994 Model and 1,000x Better*, AEIDEAS (May 25, 2016), <http://www.aei.org/publication/technology-has-advanced-so-rapidly-that-a-laptop-computer-today-is-96-cheaper-than-a-1994-model-and-1000x-better/> [https://perma.cc/ZZY7-JRGP].

4. Katherine Medianik, *Artificially Intelligent Lawyers: Updating the Model Rules of Professional Conduct in Accordance with the New Technological Era*, 39 CARDOZO L. REV. 1497, 1498 (2018).

5. Maura R. Grossman & Gordon V. Cormack, *The Grossman-Cormack Glossary of Technology-Assisted Review*, 7 FED. CTS. L. REV. 1, 32 (2013).

6. Namely, TAR programs that rely on rules-based processes to emulate human decision-making processes. See generally *id.* To put it in lawyerly terms, these TAR programs effectively operate as syllogism machines.



learning,<sup>7</sup> harnessing judgments from human administrators,<sup>8</sup> to distinguish relevant documents for production in discovery.<sup>9</sup> Specifically, “supervision” means that the program is taught to distinguish relevant documents from irrelevant documents through the use of a training set prepared by a human administrator.<sup>10</sup> The training set is a packet of documents, which the administrator has coded as relevant/irrelevant.<sup>11</sup> The program infers how to distinguish between the two categories by reference to the training set’s examples—literally learning from them.<sup>12</sup>

Often, electronic data is stored with vague descriptors or in “generic, co-mingled folders such as an e-mail system’s ‘inbox’ or ‘outbox.’”<sup>13</sup> In these circumstances, and in cases dealing with large amounts of electronic data, automated search methods like TAR are a reasonable approach for lawyers to take.<sup>14</sup> Industry experts use several metrics to evaluate TAR’s success, from commonly understood terms like “accuracy”<sup>15</sup> to mathematical expressions like “Area Under the ROC Curve,” or AUC.<sup>16</sup> This Article evaluates TAR’s success as its ability to effectively comply with a user’s discovery obligations. This definition is better suited for the focus of this work because a success metric like “90% recall of relevant documents” (i.e., 90% of the time, the program correctly labels documents as relevant)<sup>17</sup> misses that the 10% of false negatives may be not just relevant, but crucial, even to the point of being more worthwhile than the other 90% altogether.<sup>18</sup>

Because TAR is fundamentally dependent on human judgment, it stands to reason that flaws in human judgment could be incorporated into the

---

7. *See id.* at 22.

8. *See id.* at 32.

9. *See id.* at 15.

10. *See id.* at 26.

11. *See id.* at 32.

12. *See id.* at 33.

13. SHIRA A. SCHEINDLIN & DANIEL J. CAPRA, ELECTRONIC DISCOVERY AND DIGITAL EVIDENCE CASES AND MATERIALS 242 (2009).

14. *See id.*

15. Grossman & Cormack, *supra* note 5, at 8 (“[Accuracy is] [t]he fraction of [d]ocuments that are correctly coded by a search or review effort . . . [H]igh accuracy is commonly advanced as evidence of an effective search or review effort.”).

16. *Id.* at 87 (“[AUC is] a summary measure used to assess the quality of [p]rioritization. AUC is the [p]robability that a randomly chosen [r]elevant [d]ocument is given a higher priority than a randomly chosen [n]on-[r]elevant [d]ocument. An AUC score of 100% indicates a perfect ranking, in which all [r]elevant [d]ocuments have higher priority than all [n]on-[r]elevant [d]ocuments.”).

17. *See id.* at 106.

18. Interview with Fernando Delgado, Ph.D. student of Info. Sci., Cornell Univ. (Nov. 14, 2018).

implementation of TAR and made manifest in TAR's product. Specifically, if TAR's parameters are flawed—for example, by being over- or under-inclusive—then TAR will produce unsatisfactory (either overly-exposing or incomplete) discovery and the human administrators will be responsible. As the adage goes: “Computers don’t make mistakes. Only people make mistakes.”<sup>19</sup> One of the fundamental principles of computer science follows from this adage: “garbage in, garbage out,” the idea that flawed inputs produce flawed outputs.<sup>20</sup> This Article relies on the principle of ‘garbage in, garbage out’ in describing both the problems that actors face with TAR and the solutions to those problems. TAR’s consequences analyzed in this Article stem from misuse (user error in training/implementing TAR) rather than malfunction (breakdown in the TAR program itself), so the first and major part of the solution is for actors to replace bad inputs.<sup>21</sup> The second part of the solution is for actors to think carefully about when and how thoroughly to use TAR. Thus, by improving TAR’s inputs and carefully implementing the program, actors minimize the sticky consequences that follow from misuse.

This Article further explores the legal consequences of poorly implemented TAR in discovery. Its goal is to identify the issues that arise from using TAR, the impact those issues can cause, and to suggest solutions to specific problems. For the purposes of this Article, the legal consequences of poorly implementing TAR fall into three major categories: (1) Federal Rules of Civil Procedure (Fed. R. Civ. P.) sanctions and spoliation sanctions in the civil context; (2) due process violations in the criminal context; and (3) personal liability for the consequences of misusing TAR, like the inadvertent disclosure of privileged or confidential documents. These consequences have been properly dealt with in this piece. Part I of this Article has introduced TAR and discussed the integration of technology into law generally. Part II then discusses the background of TAR in discovery, including a small body of case law, and explores the existing literature on TAR. Part III analyzes the three primary legal consequences of poorly implementing TAR: (1) civil litigation consequences in the form of Fed. R. Civ. P. sanctions, spoliation sanctions for

---

19. Nishant Shah, *Computers Don’t Make Mistakes?*, INDIAN EXPRESS (May 24, 2015), <https://indianexpress.com/article/technology/social/computers-dont-make-mistakes/> [https://perma.cc/3NP3-LUEQ] (internal quotations omitted).

20. Margaret Rouse, *Garbage In, Garbage Out*, SEARCHSOFTWAREQUALITY, <https://searchsoftwarequality.techtarget.com/definition/garbage-in-garbage-out> [https://perma.cc/RRR5-WUGJ] (last updated Mar. 2008).

21. See COAL. OF TECH. RES. FOR LAWYERS, 2016 GUIDELINES REGARDING THE USE OF TECHNOLOGY-ASSISTED REVIEW 19 (2016) (“To establish defensibility [of a TAR protocol], counsel must accurately determine the prevalence of responsive information, ensure that its training process yields acceptable levels of recall and precision for its production of documents, and validate its final production results.”).

a party's failure to produce the full amount of required disclosure, agency control, and discovery expansion; (2) criminal prosecution consequences in the form of *Brady* violations and other due process concerns when the government fails to produce the full amount of required disclosure; and (3) personal liability for those responsible for TAR's accidental inclusion of privileged or confidential information for disclosure to the requesting party. Part IV analyzes the economic costs that companies risk from the misuse of TAR in corporate law. Part V concludes with a brief summary and a look towards the future.

## II. TAR'S BACKGROUND

Existing scholarship has not explored these three legal consequences of poorly-implemented TAR. Rather, existing scholarship falls into three general categories: (1) those defining TAR and related terms;<sup>22</sup> (2) those predicting how and where TAR may be used;<sup>23</sup> and (3) those discussing whether attorneys are required to turn over their seed sets (the information on which the TAR program is trained) in discovery to the opposing party under Fed. R. Civ. P. 26(g).<sup>24</sup> One author suggests that using TAR may actually be a requirement for attorneys because the Model Rules of Professional Conduct "obligate plaintiff attorneys to educate themselves about new technologies and to encourage their use when their clients can benefit."<sup>25</sup> There is strong evidence that suggests that the use of TAR can benefit clients; TAR is economically efficient, able to "reduce the cost of document review by up to 75 percent."<sup>26</sup> Additionally, research has revealed the inaccuracy and inconsistency of traditional manual document review, highlighting the need for TAR.<sup>27</sup> Some

---

22. See, e.g., Grossman & Cormack, *supra* note 5, at 32.

23. See, e.g., Annika K. Martin, *How to Stop Worrying and Love Predictive Coding*, 52 TRIAL 36 (2016).

24. See, e.g., Karl Schieneman & Thomas C. Gricks, III, *The Implications of Rule 26(g) on the Use of Technology-Assisted Review*, 7 FED. CTS. L. REV. 247, 259–63 (2013).

25. Martin, *supra* note 23, at 37; see Steven M. Puiszis, *A Lawyer's Duty of Technological Competence*, AM. BAR ASS'N 1, 2 (2017), [https://www.americanbar.org/content/dam/aba/events/professional\\_responsibility/2017%20Meetings/Conference/conference\\_materials/session4\\_information\\_governance/puiszis\\_lawyers\\_duty\\_technological\\_competence.authcheckdam.pdf](https://www.americanbar.org/content/dam/aba/events/professional_responsibility/2017%20Meetings/Conference/conference_materials/session4_information_governance/puiszis_lawyers_duty_technological_competence.authcheckdam.pdf) [<https://perma.cc/5Z4T-PLKQ>] (stating that twenty-seven states have adopted the amendment to the ABA's Model Rule 1.1 requiring lawyers to keep "abreast of changes in the law and its practice," which includes knowing "the benefits and risks . . . associated with relevant technology") (internal quotations omitted).

26. Martin, *supra* note 23, at 38.

27. *Id.* (stating that "one study found that 'human reviewers missed between 20 percent and 75 percent of all relevant documents, and 90 percent of those mistakes resulted from inarguable human error'").

firms, however, may prefer manual document review: document review can be a great profit center for law firms, who are able to bill clients for many hours of junior associates' time-intensive labor.<sup>28</sup> The junior associates, on the other hand, undoubtedly would not be sorry to see document review assigned elsewhere.<sup>29</sup> Some firms have very successfully navigated the changing electronic discovery ("e-discovery") landscape by building an internal e-discovery practice offering everything from manual review by contract attorneys to complex review by technologists and e-discovery attorneys.<sup>30</sup> This in-house strategy requires a significant capital investment,<sup>31</sup> however, and artificial intelligence like TAR threatens to undermine the effectiveness of complex in-house manual review.<sup>32</sup> In any case, the future is not one of machines replacing humans.<sup>33</sup> Rather, TAR requires a great deal of human interaction, relying ultimately on user judgment.<sup>34</sup> As described above, supervised machine learning relies on the human administrator's judgment in creating the right training set, and humans have the final word in pronouncing a document's relevance or irrelevance. Therefore, careful human judgment is crucial to TAR's success.

Similarly, courts are just beginning to grapple with TAR. The first federal court decision approving the use of TAR in discovery was handed down in 2012.<sup>35</sup> The Southern District of New York, in *Da Silva Moore v. Publicis Groupe*, noted "that computer-assisted review is not perfect," but the "Federal Rules

---

28. Interview with Jim Noles, Partner, Barze Taylor Noles Lowther LLC (Dec. 20, 2018).

29. Mary Kate Sheridan, *The Truth About Doc Review*, VAULT (Sept. 4, 2018), <https://www.vault.com/blogs/vaults-law-blog-legal-careers-and-industry-news/the-truth-about-doc-review> [<https://perma.cc/WUW2-85PJ>] ("Nothing seems to invite as much disdain from junior associates as document review.").

30. See ALM Media, *How a Few Savvy Law Firms Turned E-Discovery Into a Cash Cow*, YAHOO FIN. (Nov. 27, 2017), <https://finance.yahoo.com/news/few-savvy-law-firms-turned-060001034.html> [<https://perma.cc/CDY9-9MSH>].

31. *Id.* ("It requires a dedicated internal staff of lawyers and technologists, data review centers and data hosting centers that can run 24/7 at scale . . . . [I]t's been hard for that investment to make sense for a lot of other firms.'").

32. See *id.* Morgan Lewis' Blair "notes the threat of artificial intelligence is real . . . . Much more of the process will be automated, perhaps requiring smaller groups to handle the work." *Id.*

33. See, e.g., *THE TERMINATOR* (Orion Pictures 1984) (portraying a post-apocalyptic future where machines have become sentient and are determined to destroy the last remnants of humanity).

34. Martin, *supra* note 23, at 38. (stating that "[w]hile predictive coding effectively transfers the drudgery of review to the machine, the judgment remains entirely the lawyer's").

35. *Moore v. Publicis Groupe*, 287 F.R.D. 182, 193 (S.D.N.Y. 2012) (stating that "[t]his Opinion appears to be the first in which a Court has approved of the use of computer-assisted review).

of Civil Procedure do not require perfection.”<sup>36</sup> The opinion also considered the evidentiary implications of TAR.<sup>37</sup> Despite the relatively small body of TAR-related case law, the case law since *Da Silva Moore* “has developed to the point that it is now black letter law that courts will permit where the producing party wants to utilize TAR for document review.”<sup>38</sup> When TAR is used to cull documents—to decide which documents to review and which to ignore—the court will ask whether the culling decision was reasonable and proportionate.<sup>39</sup> To determine whether the culling decision was reasonable and proportionate, the court will ask whether the cost of reviewing the culled documents was greater than their expected value.<sup>40</sup>

In deciding whether a culling decision was reasonable and proportionate, one important question to ask is what happens when search terms are applied before TAR is used, reducing the volume of produced documents? If producing relevant documents is like searching for a needle in a haystack, the pre-TAR application of search terms is like reducing the number of searchable haystacks, lowering one’s chances of finding the needle. The Northern District of Indiana, in *Biomet M2a Magnum Hip Implant Products Liability*, held that using search terms to reduce the number of documents in discovery from 19.5 million to 2.5 million before applying TAR “complie[d] fully with the requirements of Federal Rules of Civil Procedure 26(b) and 34(b)(2).”<sup>41</sup> The pre-TAR application of search terms might be unreasonable in a case with fewer documents, though, and future cases will likely explore this. Three distinct factors help explain the dearth of case law on the use of TAR in discovery. First, litigators have an absence of litigation vehicles to advance discovery disputes, as most disputes are resolved through Fed. R. Civ. P. motions<sup>42</sup> and are not worth litigating after trial (i.e., on appeal). Consequently,

---

36. *Id.* at 191.

37. *Id.* at 189 (holding that Federal Rule of Evidence 702 and *Daubert* are trial rules that do not apply to “how documents are searched for and found in discovery”).

38. *Rio Tinto PLC v. Vale S.A.*, 306 F.R.D. 125, 127 (S.D.N.Y. 2015); *see also* *Dynamo Holdings L.P. v. Comm’r*, 143 T.C. 183, 192 (2014) (holding that the court will allow a party to use predictive coding where the party reasonably requests to use predictive coding and represents to the court that it will retain experts to conduct a search acceptable to the opposing party).

39. *See* THE ELECTRONIC DISCOVERY INSTITUTE, THE FEDERAL JUDGES’ GUIDE TO DISCOVERY 167 (3d ed. 2017).

40. *Id.* Note that this inquiry focuses on economic efficiency (where expected cost is greater than expected value, the efficient option is to forego the conduct), equating economic efficiency with reasonableness.

41. *In re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*, NO. 3:12-MD-2391, 2013 U.S. Dist. LEXIS 84440, at \*5 (N.D. Ind. Apr. 18, 2013).

42. *See, e.g.*, FED. R. CIV. P. 37(a) (providing a motion for an order compelling disclosure or discovery).

there is not much opportunity to create a body of binding case law on discovery disputes.<sup>43</sup> Second, discovery is mostly managed by the parties, frequently with the help of a magistrate judge. Parties are unlikely to spend the resources necessary to substantially litigate discovery disputes, and even if they did, the magistrate judge's opinion on the discovery dispute would likely not be published, because the dispute is ancillary to the underlying action.<sup>44</sup> Third, some judges are behind the learning curve with TAR, because it is an emerging technology; therefore, judges may be loath to write extensive opinions on a topic they are unfamiliar with.<sup>45</sup>

Courts may influence the way parties implement TAR by limiting the scope<sup>46</sup> of electronically-stored information (ESI) discovery at the pretrial discovery conference and by issuing evidentiary orders ahead of trial, such as Federal Rule of Evidence 502(d) orders.<sup>47</sup> Judges may decide to appoint special masters to help navigate TAR issues “as technology grows increasingly complex.”<sup>48</sup> Delegating responsibility for TAR oversight by appointing a

---

43. Interview with Keir Weyble, Professor, Cornell Law School (Oct. 9, 2018).

44. See Ellen Platt, *Unpublished vs. Unreported: What's the Difference?* 5 PERSP.: TEACHING LEGAL RES. & WRITING 26, 26–27 (1996) (stating that “[b]y 1994, all federal circuit courts and the majority of state courts had adopted some sort of policy to limit publication of opinions” and that “West’s editors typically exclude short memorandum decisions, orders, and other routine housekeeping items from both print sources and WESTLAW”); see also *Submission Guidelines for Court Opinions*, THOMSON REUTERS LEGAL, <https://legal.thomsonreuters.com/en/solutions/government/court-opinion-submission-guidelines> [https://perma.cc/JHH3-GK3M] (last visited Jan. 4, 2020) (instructing that judges should submit opinions that are “of general interest and importance to the bench and the bar, such as those that: [d]eal with an issue of first impression[.] [e]stablish, alter, modify, or explain a rule of law[.] . . . [p]resent a unique holding[.] [or] [i]nvolve newsworthy cases”). Because TAR is an emerging technology in discovery, though, magistrate judges may write future discovery opinions concerning TAR that meet the qualifications for publication (e.g., because they deal with an issue of first impression) or that get affirmed or adopted by the district courts and get published for that reason.

45. See *Proverbs* 17:28 (NIV) (“Even fools are thought wise if they keep silent, and discerning if they hold their tongues.”). Of course, the lawyers are obligated to educate the judge on the subject matter before the court, and this is not meant to suggest that judges would not decide cases on subject matter they are ignorant about—merely that their opinions in those cases would likely be brief and lack significant discussion of the details.

46. See FED. R. CIV. P. 26(b)(2) (detailing the proportionality factors judges use to define the scope of discovery, including the importance of the issues at stake in the action, the amount in controversy, and the parties’ relative access to relevant information).

47. See RONALD J. HEDGES ET AL., *MANAGING DISCOVERY OF ELECTRONIC INFORMATION* 39 (3d ed. 2017) (stating that judges may order, under Federal Rule of Evidence 502(d), that the production of materials in discovery will not waive privilege or work-product protection and that the order is enforceable against third parties, as well).

48. Jenny Le, *Technology-Assisted Review: Insight into Lingering Questions*, 23 PRETRIAL PRAC. & DISCOVERY 4, 6 (2015).

special master or assigning the dispute to a magistrate judge places a level of separation between the judge and the litigants, freeing the judge to focus on the case instead of ancillary issues like discovery. Courts have declined, however, to force a party to use TAR when the party objects to it.<sup>49</sup> This Article will discuss in Part III why it nevertheless might be wise to require prosecutors to use TAR over objection.

### III. LEGAL CONSEQUENCES OF MISUSE

#### A. CIVIL SANCTIONS

##### 1. *Federal Rule of Civil Procedure 37(c) Sanctions*

This Section discusses the ways in which a party's misuse of TAR can incur Fed. R. Civ. P. 37(c) sanctions, the significance of those sanctions, and how to avoid them.

The civil consequences of a failure to produce less than the full amount of required disclosure can be severe. Imagine that the required disclosure includes documents a<sub>1</sub>–a<sub>5,000</sub>. The producing party uses a TAR program that uses too narrow a definition of relevance. As a result, the program produces only documents a<sub>1</sub>–a<sub>4,500</sub>. This series of events exposes the producing party to the threat of sanctions. Fed. R. Civ. P. 37(c) prevents a party from using evidence after failing to disclose it as required by Fed. R. Civ. P. 26(a) and (e) (required and supplemental disclosures, respectively).<sup>50</sup> The court may also impose monetary sanctions, inform the jury of the party's failure to disclose, and impose other sanctions.<sup>51</sup> If a party's failure to disclose is harmless, then the court is not required to impose Fed. R. Civ. P. 37(c) sanctions.<sup>52</sup> If the party made an honest mistake in failing to disclose because the party did not know of the existence of the evidence, for example, and the requesting party does know of the evidence's existence, the mistake is a strong indication that the violation was harmless.<sup>53</sup> The burden of showing that the violation was harmless remains, however, with the party facing sanctions.<sup>54</sup> Judges have wide

---

49. THE ELECTRONIC DISCOVERY INSTITUTE, *supra* note 39, at 171.

50. FED. R. CIV. P. 37(c)(1).

51. *Id.* at (A)–(C). Additionally, when a party's discovery production is overly broad, the court may modify the scheduling order. *See* Casey C. Sullivan, *AI-Driven Discovery Process Produces Millions of Unresponsive Docs*, LOGIKCULL (Sept. 20 2018), <https://blog.logikcull.com/discovery-on-autopilot-crashes-and-burns> [<https://perma.cc/T3DR-FQTY>]. This Article also shows that significant problems arise when a party misuses TAR (overproduction of unresponsive documents created delay).

52. *See* *Roberts v. Galen of Va., Inc.*, 325 F.3d 776, 782–83 (7th Cir. 2003).

53. *Id.* at 783.

54. *Molly, Ltd. v. Deckers Outdoor Corp.*, 259 F.3d 1101, 1107 (9th Cir. 2001).

discretion in the imposition of Fed. R. Civ. P. 37(c) sanctions and are able to dismiss the suit entirely for disclosure violations, even where the offending party has not violated a court order.<sup>55</sup> Judges are given such wide latitude because the Supreme Court intended to strictly enforce compliance with the discovery rules when it adopted Fed. R. Civ. P. 37(c).<sup>56</sup> Therefore, lawyers using TAR should give special consideration to ensuring that the TAR code is trained to produce the full amount of relevant documents to which the opposing party is entitled. This can be done on the front end by rigorously developing the training set, and on the back end by reviewing TAR's output for accuracy. The "should" here is a normative statement; to comply with the spirit of Fed. R. Civ. P. 37(c), and with the moral obligation<sup>57</sup> inherent in turning over documents in discovery, lawyers should use TAR appropriately to robustly fulfill the discovery requirements.

## 2. *Spoliation Sanctions*

This Section discusses how a party's misuse of TAR in civil discovery can expose the party to spoliation sanctions, the possible consequences of those sanctions, and how a party can avoid incurring them.

Spoliation is "[t]he intentional destruction, mutilation, alteration, or concealment of evidence."<sup>58</sup> A party's failure to preserve evidence that results in the destruction or loss of relevant information is spoliation, which is negligent and may even be grossly negligent or willful, depending on the circumstances.<sup>59</sup> For example, a party's failure to collect records from key players constitutes gross negligence or willfulness.<sup>60</sup> Therefore, if a party uses TAR to decide which documents to collect for disclosure and TAR does not identify documents held by key players, the party will be liable for gross negligence that possibly amounts to spoliation, especially if the evidence is subsequently lost or destroyed by those players.

---

55. *Ortiz-Lopez v. Sociedad Espanola de Auxilio Mutuo y Beneficiencia de P.R.*, 248 F.3d 29, 33–34 (1st Cir. 2001).

56. Lisa Stockholm, *The Duty to Disclose: Rule 37(c) and Self-Executing Sanctions*, AM. BAR ASS'N (June 30, 2011), <http://apps.americanbar.org/litigation/committees/trialevidence/articles/summer2011-self-executing-sanctions-rule-37c.html> [https://perma.cc/4CBY-9U26].

57. In other words, fairness in avoiding "trial by ambush." See *How Courts Work*, AM. BAR ASS'N (Dec. 2, 2013), [https://www.americanbar.org/groups/public\\_education/resources/law\\_related\\_education\\_network/how\\_courts\\_work/discovery/](https://www.americanbar.org/groups/public_education/resources/law_related_education_network/how_courts_work/discovery/) [https://perma.cc/CNH7-WYTA].

58. *Spoliation*, BLACK'S LAW DICTIONARY (10th ed. 2014).

59. *Pension Comm. of the Univ. of Montreal Pension Plan v. Banc of Am. Sec., LLC*, 685 F. Supp. 2d 456, 464 (S.D.N.Y. 2010).

60. *Id.* at 465.



To prevail on a claim of spoliation, the complaining party must show that:

[T]he spoliating party (1) had control over the evidence and an obligation to preserve it at the time of destruction or loss; (2) acted with a culpable state of mind upon destroying or losing the evidence; and that (3) the missing evidence is relevant to the innocent party's claim or defense.<sup>61</sup>

The factfinder may presume relevance and prejudice where the spoliating party acted in bad faith or was grossly negligent.<sup>62</sup> Any presumption is rebuttable, though, regardless of the spoliating party's degree of culpability.<sup>63</sup>

The court may impose sanctions for a party's spoliation of evidence, but only to the degree "necessary to redress conduct which [sic] abuses the judicial process."<sup>64</sup> The court's sanction should deter spoliation, place the risk of erroneous judgment on the offending party, and restore the prejudiced party to the position it would have occupied in the absence of the other's spoliation.<sup>65</sup> The court's power to impose spoliation sanctions originates in the court's inherent power to control the judicial process and the litigation before it.<sup>66</sup> Possible spoliation sanctions include, from least to most severe: mandating "further discovery, cost-shifting, fines, special jury instructions, preclusion, and the entry of default judgment or dismissal."<sup>67</sup>

Because a party's poor implementation of TAR can result in a spoliation sanction of an adverse default judgment, parties should carefully use TAR to avoid spoliation by identifying key documents and flagging them for preservation once the duty to preserve attaches.<sup>68</sup> The "should" here is both a normative and an economic statement: normative in that an attorney could be professionally accountable<sup>69</sup> for failing the client if the court dismissed the case

---

61. *Id.* at 467.

62. *Id.*

63. *Id.* at 468.

64. *Id.* at 465 (internal quotations omitted); *see* *Anderson v. Beatrice Foods Co.*, 900 F.2d 388, 395 (1st Cir. 1990) (stating that a judge, in crafting sanctions, "should take pains neither to use an elephant gun to slay a mouse nor to wield a cardboard sword if a dragon looms").

65. *Pension Comm. of the Univ. of Montreal Pension Plan*, 685 F. Supp. 2d at 469 (quoting *West v. Goodyear Tire & Rubber Co.*, 167 F.3d 776, 779 (2d Cir. 1999)).

66. *Id.* at 465.

67. *Id.* at 469.

68. *See id.* at 466 (stating that the duty to preserve evidence arises when a party reasonably anticipates litigation, requiring a party to put a litigation hold in place to ensure preservation of relevant documents).

69. *Compare* AM. BAR ASS'N, MODEL RULES OF PROFESSIONAL CONDUCT 1.1 (2019) (stating that "[a] lawyer shall provide competent representation to a client") *with* AM. BAR ASS'N, MODEL RULES OF PROFESSIONAL CONDUCT 8.4(a) (2019) (stating that it is professional misconduct for a lawyer to "violate or attempt to violate the Rules of Professional Conduct").

as a spoliation sanction, and economic in that cost-shifting sanctions and fines are expenses that can be avoided.

### 3. *Sanctions for Intentional Misuse*

This Section discusses how courts may respond to a party's intentional misuse of TAR as distinct from the accidental misuse described in the previous two Sections. This Section discusses the limits that the technology imposes on the court's insight into the TAR process and possible solutions for deterring parties from intentionally misusing TAR.

To this point, this Article has discussed unintentional acts or omissions: mistakes in a party's implementation of TAR, either in training TAR or in using it, that can lead to sanctions. But intentional acts and omissions also merit examination. Because courts are split on the degree of transparency required in the use of TAR,<sup>70</sup> and because of the complicated nature of machine learning, parties have ample opportunity to intentionally restrict the TAR program's code to produce less than the required amount of discovery, or to improperly cull the dataset by applying overbroad search terms to the dataset before running TAR. The Fed. R. Civ. P. provide meager deterrents to this kind of behavior. Fed. R. Civ. P. 26(g)(1) requires a party's attorney to sign the disclosure, certifying that it is complete and correct.<sup>71</sup> Fed. R. Civ. P. 26(g)(3) requires the court to impose an appropriate sanction for unjustified improper certification, regardless of whether the opposing party moves for the court to do so.<sup>72</sup> These sanctions may include an order to pay the reasonable expenses (including attorneys' fees) caused by the improper certification.<sup>73</sup> While Fed. R. Civ. P. 26(g)(3) sanctions may deter an actor from intentionally misusing TAR, they are not a complete deterrent. Because TAR's processes are complicated, an actor may feel quite certain that intentional misuse will not be perceptible, or if it is perceptible, that misuse is unlikely to be caught by the other party due to a lack of transparency concerning the actor's training of TAR via the seed set. The actor may be tempted to see the low risk of apprehension as no risk at all. Thus, there is no deterrent in these circumstances if the actor lacks moral fortitude.

Yet there are some ways to deal with the intentional misuse of TAR. One route that the law could take to deter intentional misuse is to amend the Fed. R. Civ. P. to include a more robust sanction for intentional misuse of TAR,

---

70. *Rio Tinto PLC v. Vale S.A.*, 306 F.R.D. 125, 128 (S.D.N.Y. 2015).

71. FED. R. CIV. P. 26(g)(1)(A). For a thorough discussion of the procedures that an actor should implement with TAR to comply with FED. R. CIV. P. 26(g), see Schieneman & Gricks, *supra* note 24, at 260–63.

72. FED. R. CIV. P. 26(g)(3).

73. *Id.*

perhaps by requiring an offending party to redo the discovery process with a TAR program specified by the court and the requesting party. Fed. R. Civ. P. 26(g)(3) already gives judges wide latitude<sup>74</sup> to craft sanctions for offending behavior, though, so amending the Fed. R. Civ. P. to include a specific sanction seems unnecessary. Additionally, imposing harsher sanctions will not deter the behavior if the actor continues to see little to no risk of apprehension. An alternative solution is to better police such misconduct so that the immoral actor perceives the probability of apprehension as much greater and is consequently deterred from engaging in the behavior. This solution also proves unsatisfactory, for two reasons. First, because the courts will likely be unwilling to become more involved in scrutinizing a party's methods for complying with discovery obligations, and second, because the technology does not lend itself to such scrutiny due to its complexity. These solutions are unsatisfactory mainly because the root of the problem lies within the actor. The best solution, then, is to remind attorneys of their ethical obligations<sup>75</sup> and encourage them to become sophisticated in their understanding of TAR so that they are both willing to use TAR appropriately and able to identify when they (or another) are misusing TAR.

#### B. OTHER LEGAL CONSEQUENCES OF TAR IN CIVIL LITIGATION

##### 1. *Administrative Agencies' Control of TAR*

This Section briefly describes how administrative agencies can control the ways in which parties use TAR to respond to subpoenas. This Section also examines the costs that administrative agencies' control imposes on litigants. The purpose of this Section is to alert attorneys to another player (in addition to courts and opposing parties) that may constrain a party's use of TAR.

Administrative agencies like the SEC can control the way litigants use TAR, affecting the cost of a party's compliance with discovery obligations. When the SEC issues a judicially-enforced administrative subpoena, the SEC can require that the defendant's use of any computer-assisted review or TAR be pre-approved by the legal and technical staff of the Division of Enforcement.<sup>76</sup> Therefore, administrative agencies can restrict a party's use of

---

74. *See id.* (using the language "may include" to signal that the list is not exhaustive).

75. AM. BAR ASS'N, MODEL RULES OF PROFESSIONAL CONDUCT 8.4(c) ("[I]t is professional misconduct for a lawyer to] engage in conduct involving dishonesty, fraud, deceit or misrepresentation."); AM. BAR ASS'N, MODEL RULES OF PROFESSIONAL CONDUCT 3.4(d) ("[A lawyer shall not] in pretrial procedure . . . fail to make reasonably diligent effort to comply with a legally proper discovery request by an opposing party.").

76. *See SEC v. Elite Advisors Sports Mgmt.*, No. CV 14-9690 PA (RZx), 2015 U.S. Dist. LEXIS 55847, at \*47 (C.D. Cal. Apr. 27, 2015).

TAR in discovery, potentially leading to greater costs for the producing party (if the party must instead rely on manual review, for example).<sup>77</sup>

Even if the Division of Enforcement subsequently approves all TAR requests, the additional burden of seeking approval increases the transaction costs<sup>78</sup> that the producing party must suffer. Another example is the Commodity Futures Trading Commission (CFTC), which requires a producing party who wishes to use TAR to consult with the CFTC attorney to define and agree upon the technology used and the requirements it must meet.<sup>79</sup> As with the SEC, the CFTC attorney may dictate the technology to be used and how it is to be employed, which can affect the costliness of TAR.<sup>80</sup> In the energy industry, regulated entities are pushing the Federal Energy Regulatory Commission (FERC) to allow them to use TAR to comply with discovery obligations in enforcement investigations.<sup>81</sup> If FERC decides to allow the use of TAR, it, too, would likely wish to prescribe when and how TAR can be used.<sup>82</sup> Consequently, litigants should be aware of administrative agencies' ability to control the litigants' use of TAR so that they can properly prepare to meet their discovery obligations if the agencies prohibit them from using TAR.

## 2. *TAR's Expansion of Discovery*

This Section describes how TAR's successes may expand the scope of discovery by making data access cheaper and easier. TAR may also expand discovery by changing the definition of what qualifies as "inaccessible" data. Fed. R. Civ. P. 26(b)(2)(B) states that parties need not provide ESI discovery from sources that are not reasonably accessible because access carries an undue burden or cost.<sup>83</sup> Because TAR can sort through millions of documents<sup>84</sup> much

---

77. See *supra* notes 26–27 and accompanying text.

78. See *Cost—transaction cost*, BLACK'S LAW DICTIONARY (10th ed. 2014) ("A cost connected with a process transaction, such as . . . the cost involved in litigating a dispute.").

79. COMMODITY FUTURES TRADING COMM'N, CFTC DATA DELIVERY STANDARDS 3 (2016).

80. See *id.*

81. DAVID A. APPLEBAUM & TODD L. BRECHER, AM. GAS ASS'N ET AL., ENHANCING THE TRANSPARENCY, EFFICIENCY, AND FAIRNESS OF THE FEDERAL ENERGY REGULATORY COMMISSION'S ENFORCEMENT PROGRAM 13–14 (2019).

82. See *id.* at 14 (claiming that FERC has resisted the use of advanced technology like TAR in enforcement).

83. *FDIC v. Bowden*, No. CV413–245, 2014 U.S. Dist. LEXIS 77890, at \*36 (S.D. Ga. June 6, 2014) (applying Fed. R. Civ. P. 26(b)(2)(B)).

84. See, e.g., *In re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*, NO. 3:12-MD-2391, 2013 U.S. Dist. LEXIS 84440, at \*2 (N.D. Ind. Apr. 18, 2013) (applying TAR to 2.5 million documents).

more efficiently than manual review,<sup>85</sup> however, the availability of TAR may reduce the credibility of claims that some particular ESI is inaccessible. Consequently, courts may require parties to provide ESI that would not otherwise have been required, at least not at the expense of the producing party.<sup>86</sup> Therefore, the legal world's adoption of TAR may expand the discovery obligation by increasing parties' access to data.

### C. SANCTIONS IN CRIMINAL PROSECUTION

#### 1. *Background Criminal Law*

This Section provides an overview of the various sources of law that entitle criminal defendants to government disclosure, how prosecutors' failure to comply with disclosure requirements violates due process of law, and what the consequences of those violations are.

The Fourteenth Amendment provides that no citizen shall be deprived of "life, liberty, or property without due process of law."<sup>87</sup> Case law entitles defendants to government disclosure of exculpatory evidence, and the government's failure to turn over such evidence is a violation of due process.<sup>88</sup> Additionally, Federal Rule of Criminal Procedure (Fed. R. Crim. P.) 16 entitles criminal defendants to a great deal of disclosure from the government, including books, papers, documents, data, photographs, and tangible objects if the item is material to preparing the defense, the government plans to use the item in trial, or the item was taken from or belongs to the defendant.<sup>89</sup> Finally, statutory law provides criminal defendants with mandatory government disclosure, and provides sanctions for failure to turn over such evidence.<sup>90</sup> Therefore, if the government employs TAR to determine which evidence should be turned over to the defendant as part of the government's disclosure, and TAR is poorly implemented such that less than the full amount of relevant evidence is produced, the defendant's Fourteenth Amendment due process rights may be violated, possibly resulting in the reversal of some convictions.

In state prosecutions, a defendant is entitled to an appeal in federal court (known as habeas corpus relief) if the defendant's constitutional rights have

---

85. See *supra* notes 26–27 and accompanying text.

86. See *Zubulake v. UBS Warburg LLC*, 216 F.R.D. 280, 284 (S.D.N.Y. 2003) (stating that cost-shifting is potentially appropriate where the requesting party seeks inaccessible data).

87. U.S. CONST. amend. XIV, § 1.

88. *Brady v. Maryland*, 373 U.S. 83, 87 (1963).

89. FED. R. CRIM. P. 16(a)(1)(E), (i)–(iii).

90. See, e.g., Jencks Act, 18 U.S.C. § 3500(b), (d) (2018) (providing mandatory government disclosure with regards to testimonial statements of witnesses in the possession of the United States).

been violated and the defendant has exhausted all state court remedies.<sup>91</sup> The defendant must “show cause for . . . failure to develop the facts in state-court proceedings and actual prejudice resulting from that failure.”<sup>92</sup> In the context of a *Brady* claim, (i.e., a claim that evidence substantially favorable to the defense failed to be disclosed), a defendant meets that standard by showing that:

(a) [T]he prosecution withheld exculpatory evidence; (b) petitioner reasonably relied on the prosecution’s open file policy as fulfilling the prosecution’s duty to disclose such evidence; and (c) the State confirmed petitioner’s reliance on the open file policy by asserting during state habeas proceedings that petitioner had already received everything known to the government.<sup>93</sup>

If the defendant proves that the prosecutor’s misuse of TAR resulted in the satisfaction of each of these elements, the defendant is entitled to federal review which may well result in a reversed conviction. Therefore, because habeas corpus relief provides defendants in state prosecutions an additional layer of scrutiny in the form of federal review, state prosecutors have another incentive to ensure that they employ TAR properly to avoid due process violations.

## 2. *Sanctions for Brady Violations*

This Section analyzes the operation of the government’s *Brady* obligations to a criminal defendant and discusses how a prosecutor’s misuse of TAR can deny the defendant the due process of law. This Section also discusses how prosecutors should consider using TAR in the context of *Brady*.

A *Brady* violation has three components: (1) the evidence at issue is favorable to the accused (because it is exculpatory or impeaching evidence), (2) the evidence was suppressed by the prosecution, and (3) prejudice ensued.<sup>94</sup> *Brady* requires the government in a criminal prosecution to disclose favorable evidence to the defense, and the government’s failure to do so is a violation of due process where the evidence is material to guilt or punishment, regardless of the government’s good or bad faith.<sup>95</sup> *Brady* is functionally a rule of imputed

---

91. *Banks v. Dretke*, 540 U.S. 668, 683, 690 (2004).

92. *Id.* at 690–91 (quoting *Keeney v. Tamayo-Reyes*, 504 U.S. 1, 11 (1992)).

93. *Id.* at 692–93 (quoting *Strickler v. Greene*, 527 U.S. 263, 289 (1999)) (internal brackets and quotation marks omitted).

94. *Strickler v. Greene*, 527 U.S. 263, 281–82 (1999).

95. *Brady v. Maryland*, 373 U.S. 83, 87 (1963).

liability,<sup>96</sup> because it charges the prosecutor with knowledge of favorable evidence known to others acting on the government's behalf, including police officers.<sup>97</sup> Additionally, defendants need not request the evidence in order to be entitled to it.<sup>98</sup>

TAR is a double-edged sword in criminal prosecutions: on the one hand, the prosecutor will be responsible for failure to disclose evidence resulting from someone on the prosecution team's improper use of TAR; on the other hand, TAR can help the prosecutor more efficiently and thoroughly sort and evaluate evidence to determine whether it is material and exculpatory or useful for impeachment.<sup>99</sup> Nonetheless, *Brady* violations are a serious problem—they can amount to official misconduct, and the National Registry of Exonerations estimates that official misconduct is a contributing factor in “more than 48%” or “about 52%” of exonerations.<sup>100</sup> Therefore, because of their ethical duties, prosecutors should take care to ensure that TAR is appropriately implemented to protect the innocent and convict the guilty.<sup>101</sup>

### 3. *Federal Rule of Criminal Procedure 16 Sanctions*

This Section discusses the requirements of Federal Rule of Criminal Procedure 16 disclosures, how TAR should be used to comply with those requirements, and how courts can impose sanctions for failing to comply with the requirements. This Section also suggests that to bridge the power gap between the government and criminal defendants, courts should impose an affirmative duty on prosecutors to use TAR in discovery to sort and categorize documents.

Fed. R. Crim. P. 16 imposes on the prosecution a continuing duty to disclose the existence of evidence discovered before or during trial if that evidence is subject to discovery and the other party requested or the court ordered the prosecution to produce it.<sup>102</sup> Discovery, then, is not a one-shot endeavor; rather, the prosecution has an ongoing obligation to disclose certain

---

96. See Daniel S. Medwed, *Brady's Bunch of Flaws*, 67 WASH. & LEE L. REV. 1533, 1538 (2010) (explaining that even evidence known only to police officers is imputed to the prosecutor).

97. See *Kyles v. Whitley*, 514 U.S. 419, 437 (1995).

98. *Id.* at 433.

99. See Martin, *supra* note 23, at 38.

100. % *Exonerations by Contributing Factor*, NAT'L REGISTRY OF EXONERATIONS, <https://www.law.umich.edu/special/exoneration/Pages/ExonerationsContribFactorsByCrime.aspx> [<https://perma.cc/5XYZ-F4YG>] (last visited Jan. 4, 2020).

101. See AM. BAR ASS'N, CRIMINAL JUSTICE STANDARDS FOR THE PROSECUTION FUNCTION 3–1.2(b) (4th ed. 2016) (“The prosecutor should seek to protect the innocent and convict the guilty . . . and respect the constitutional and legal rights of all persons . . .”).

102. FED. R. CRIM. P. 16(c).

evidence to the defense. Because the prosecution's duty to disclose is continuous, the prosecution should use TAR multiple times throughout the progression of the litigation to thoroughly satisfy that duty as efficiently as possible. Of course, cost and time considerations present practical limitations on the extent to which prosecutors should follow this recommendation. Because most disclosure happens early,<sup>103</sup> the marginal benefits of an additional TAR cycle are diminished as trial approaches, so the use of TAR should be frontloaded in the litigation cycle. The exact number of times prosecutors should use TAR and the depths to which the program should probe in a given case will, of course, vary depending on the circumstances.<sup>104</sup>

Fed. R. Crim. P. 16(d)(2) empowers the district court to impose sanctions on a party for that party's failure to comply with Fed. R. Crim. P. 16.<sup>105</sup> The possible sanctions include ordering the party to permit the discovery,<sup>106</sup> specifying the methods by which the discovery will be conducted,<sup>107</sup> granting a continuance,<sup>108</sup> prohibiting the introduction of the undisclosed evidence at trial,<sup>109</sup> or entering "any other order that is just under the circumstances."<sup>110</sup> Although the court has wide discretion in deciding what order is just under the circumstances, courts of appeal have held that a district court should impose the least severe sanction that will effectively accomplish prompt compliance with discovery orders.<sup>111</sup> Even so, the least severe sanctions could prove to be significant setbacks for the prosecution, given the time and resources necessary to litigate a case.<sup>112</sup> Therefore, prosecutors have a significant incentive to use TAR appropriately in discovery to avoid the court's imposition of sanctions.

---

103. *See id.* (stating that the government's disclosure is at the defendant's request, which usually happens early).

104. *See* Dept. of Justice, Justice Manual § 9–5.002 (stating that because ultimate responsibility for disclosure rests with the prosecutor, "the prosecutor's decision about how to conduct this review is controlling")

105. *See* FED. R. CRIM. P. 16(d)(2).

106. *Id.* at 16(d)(2)(A).

107. *See id.*

108. *Id.* at 16(d)(2)(B).

109. *Id.* at 16(d)(2)(C).

110. *Id.* at 16(d)(2)(D).

111. *United States v. Martinez*, 455 F.3d 1127, 1130 (10th Cir. 2006); *United States v. Gee*, 695 F.2d 1165, 1169 (9th Cir. 1983); *see United States v. De La Rosa*, 196 F.3d 712, 715 (7th Cir. 1999) (stating that although trial courts have discretion to fashion sanctions under Fed. R. Crim. P. 16(d)(2), "a new trial is warranted only after all other, less drastic remedies are inadequate").

112. *See First Estimates of Judicial Costs of Specific Crimes, From Homicide to Theft*, RAND (Sept. 12, 2016), <https://www.rand.org/news/press/2016/09/12.html> [<https://perma.cc/5PVH-RTQX>] (finding that the judicial costs of a homicide can range from \$22,000–\$44,000).



The Fed. R. Crim. P. do not provide defendants with the same quality of disclosure to which civil litigants are entitled.<sup>113</sup> For example, in civil litigation, a party must produce ESI as it is kept in the usual course of business or must organize and label the ESI to correspond to categories in the discovery request, but no parallel rule exists in criminal prosecutions.<sup>114</sup> This creates a major problem called “data dumps,” where the prosecution “drives up the dump truck, dumps off all of the discovery, five thousand documents, fifty tapes, and now we have to go through all of that to make a determination as to what’s relevant or not.”<sup>115</sup> In addition to producing discovery in this unorganized manner, the prosecution also sometimes delivers the ESI in a format that is inaccessible to the defense.<sup>116</sup> Some judges believe it is inappropriate to involve themselves in ESI discovery, and they do little to help the defense obtain usable discovery.<sup>117</sup> There are few other resources for criminal defendants—although the Defender Services Office has funded the National Litigation Support Team (NLST) to help criminal defense lawyers through training and direct assistance with ESI, the NLST has only four members, a “woefully inadequate” number to support the federal judiciary, federal and community defenders, and nearly 10,000 panel attorneys across the nation.<sup>118</sup> The government, of course, has infinitely more resources than public defenders do.<sup>119</sup>

Two possible solutions emerge to solve the ESI data dump problem. The first solution is to subsidize/train criminal defense attorneys in the purchase and use of TAR. Leaving the sorting burden with the defendant may be the most efficient option because the defendant has the strongest incentive to uncover exculpatory evidence. The second solution is to require prosecutors to use TAR to sort and categorize the ESI for discovery prior to turning the data over to the defense. The second solution is superior for a variety of reasons: (1) it will be easier to implement uniformly, (2) some of the capital is likely already in place, and (3) it addresses the problem at the source. Subsidizing TAR for defense attorneys is a more piecemeal approach that raises difficult questions, such as how often an attorney must act as a federal defender before being entitled to a TAR subsidy. By contrast, requiring

---

113. Interview with the Honorable Jonathan W. Feldman, U.S. Magistrate Judge, Western District of New York (Oct. 4, 2018).

114. Compare FED. R. CIV. P. 34(b)(2)(E)(i) with FED. R. CRIM. P. 16.

115. CRIMINAL JUSTICE ACT REVIEW COMM., 2017 REPORT OF THE AD HOC COMMITTEE TO REVIEW THE CRIMINAL JUSTICE ACT 230 (2017).

116. *Id.* at 228–30.

117. *Id.* at 232.

118. *Id.*

119. Interview with the Honorable Jonathan W. Feldman, *supra* note 112.

prosecutors to use TAR to sort and categorize the discovery before turning it over to the defense reduces duplication of efforts and solves the data dump problem. This solution will also be easier to implement uniformly. Because prosecutors are government agents, regulations regarding their behavior and processes are easier to standardize and deploy uniformly throughout the country. Furthermore, as TAR becomes more and more prevalent, it will become commoditized—cheaper and more widely available.<sup>120</sup> On the other hand, putting this burden on the government might be viewed as shifting the defense's responsibilities to the prosecution. The prosecution has a responsibility, however, to disclose usable evidence, and a strong incentive to uncover exculpatory evidence, too, because *Brady* holds the prosecutor responsible for exculpatory evidence known to anyone on the prosecution team.<sup>121</sup> Therefore, the best solution to the data dump problem is to require prosecutors to use TAR to sort and standardize ESI before turning it over to the defense. This solution will help safeguard defendants' due process rights by obviating the need for defense counsel to undertake the daunting (and probably neglected) task of sorting through mountains of unlabeled data, freeing them to focus on more important aspects of the case.

A defendant's need of usable data is, in many ways, much greater in a criminal prosecution than in a civil suit, because the defendant stands to lose liberty or possibly even life, and the defendant will in any event have a criminal record if found guilty. Therefore, courts and policymakers should carefully consider imposing a duty on prosecutors to use TAR in a way that does not accomplish the bare minimum, but instead assists defendants by sorting data in a more tailored fashion. This solution will help bridge the gap between the quality of civil disclosure and the quality of criminal disclosure.

#### 4. *Sanctions for Jencks Act Violations*

This Section discusses the Jencks Act as another source of law providing criminal defendants with government disclosure, and how a prosecutor's misuse of TAR can expose the prosecutor to Jencks Act sanctions.

The Jencks Act provides that in a federal criminal prosecution, after a prosecution witness has testified on direct examination, the court shall, on motion of the defendant, order the United States to produce any statement of the witness that (1) the United States possesses and (2) relates to the subject matter of the witness' testimony.<sup>122</sup> If the prosecution fails to comply with the

---

120. See *supra* note 3 and accompanying text (suggesting that technology becomes commoditized as it advances).

121. See *Kyles v. Whitley*, 514 U.S. 419, 437 (1995).

122. 18 U.S.C. § 3500(b) (2018).

order, the court shall strike the witness' testimony or declare a mistrial if the interests of justice so require.<sup>123</sup> Therefore, if the prosecution uses TAR to retrieve the witness' statements for compliance with a Jencks order, and TAR fails to produce all of the relevant statements, the prosecution is subject to these mandatory sanctions, which could cause the prosecution to lose the trial (if the witness' testimony is stricken and that testimony is crucial evidence) or expend resources to litigate a new trial (if the court declares a mistrial). However, the government does have the privilege of withholding the identity of informants.<sup>124</sup> For that reason, the prosecution should employ TAR carefully to avoid accidentally disclosing informants' identities when retrieving witnesses' statements for disclosure under the Jencks Act.<sup>125</sup>

#### D. PERSONAL LIABILITY FOR MISUSE

This Section examines how attorneys may be personally liable for misusing TAR and how they may contract some of that liability to others, reducing the attorneys' exposure to risk.

An attorney's poor implementation of TAR may inadvertently uncover privileged or confidential material, such as work product or material protected by attorney-client privilege. If TAR's seed set is coded broadly enough to capture this data, and TAR is not trained to exclude the data from TAR's net, then TAR creates a risk that privileged or confidential material will be inadvertently disclosed to the requesting party.<sup>126</sup> As a matter of corporate reality, TAR's failure will ultimately be attributed to some person or persons. Consequently, the wrath of employers and the possibility of civil liability in tort will descend on those employees responsible for the inadvertent disclosure of privileged material inappropriately collected and disseminated (or at least marked as relevant and included in a packet for disclosure) by the TAR program.

Civil liability may not be restricted solely to the attorneys who implement TAR and code its seed set but may extend to the creator and seller of the TAR program in question, as well, under a breach of warranty liability theory. It is likely, however, that lawyers who purchase TAR programs will insist on contracting liability to the seller in the case of TAR failure.<sup>127</sup> Because law firms

---

123. 18 U.S.C. § 3500(d) (2018).

124. *McCray v. Illinois*, 386 U.S. 300, 308 (1967).

125. Note that informants could testify, subjecting their statements to the Jencks Act, without revealing their identities.

126. However, privilege checks typically happen independently of TAR processes. For example, a team may review for privilege all of the documents that TAR marks as relevant. Interview with Fernando Delgado, *supra* note 18.

127. Interview with Keir Weyble, *supra* note 43.

and TAR sellers are sophisticated parties, there is nothing to suggest that the parties cannot bargain for such contracts or that courts will not enforce the contracts.<sup>128</sup> Therefore, while it is unlikely that TAR sellers will be directly liable for TAR's accidental disclosure of privileged or confidential information, it is entirely plausible that the sellers would be indirectly liable for such damage through a contractual provision. Of course, even if sellers agree to assume contractual liability for program failure, sellers are unlikely to assume responsibility for user error. Consequently, attorneys still have an incentive to deploy TAR properly so that they avoid personal liability for user error, even if they can contract other liability away. Additionally, discretion is an inescapable aspect of discovery, regardless of whether an actor uses manual review or TAR. An actor must exercise discretion in deciding which documents to mark as relevant, and so must TAR. Those who use TAR should take care to train the program to exercise discretion appropriately and ensure that they are not attempting to outsource responsibility to TAR, because ultimately, they will have to answer for TAR's failures, potentially in the form of personal liability.

While the court may mitigate some of the sting of accidental disclosure—for example, by issuing a Federal Rule of Evidence 502(d) order that accidental disclosure during discovery does not amount to a waiver of privilege or work-product protection—the reality is that the damage is already done as soon as the disclosure is made. Even if privileged and work-product evidence is not admissible at trial, the receiving party can still use it to understand the producing party's trial strategy, or as a guidepost to direct the receiving party where to look for more favorable evidence.<sup>129</sup> Therefore, protecting confidentiality and privileged documents is critical and should be a primary concern of those involved in developing and implementing TAR.

#### IV. ECONOMIC CONSEQUENCES OF MISUSE

This Part discusses the economic consequences of misusing TAR in corporate law. Although these costs result from legal actions, they come from delays, rather than sanctions, and so they are distinct from the legal consequences described in the preceding Sections.

Although this Article has mostly focused on parties' use of TAR in litigation, TAR can also be helpful in corporate law. For example, TAR may be very useful in mergers because mergers require analysis of vast numbers of

---

128. *Cf. Williams v. Walker-Thomas Furniture Co.*, 350 F.2d 445, 449 (D.C. Cir. 1965) (stating that contracts should not be enforced where the element of unconscionability is present).

129. Interview with Keir Weyble, *supra* note 43.

documents.<sup>130</sup> Recently, Lighthouse (an e-discovery facilitation company) assisted its client in a pending merger when the Department of Justice's (DOJ) discovery request required the client to analyze 1.1 million documents.<sup>131</sup> Lighthouse employed TAR to reduce the amount of data requested by 87% and "easily meet" the DOJ's seventy-five-day production deadline.<sup>132</sup> This was not the first time that the DOJ approved the use of TAR in analyzing merger documents. In 2013, the DOJ and Constellation Brands agreed to use TAR to review millions of documents related to Constellation's purchase of Corona and other brands from Anheuser-Busch Inbev.<sup>133</sup> The review was conducted within two weeks.<sup>134</sup> TAR saved Constellation "significant costs," while the DOJ received "a targeted production that included a high percentage of information that was helpful to its analysis of the proposed merger," which the DOJ ultimately approved.<sup>135</sup>

But the use of TAR in corporate transactions, if carelessly done, can be costly. Mistakes in using TAR are costly because they have rippling effects—TAR that uses predictive coding will rely on the information it learns to make decisions about which documents are relevant, multiplying mistakes by duplicating them.<sup>136</sup> So the risk of creating enormous problems by overly restricting the seed set is obvious. One potential risk of using TAR in these circumstances, though, is perhaps not obvious: the DOJ may require access to a sample set of both the responsive and non-responsive datasets, which would require the parties to turn over documents that the government otherwise would never have seen (that is, those in the non-responsive dataset).<sup>137</sup> This overexposure may give the DOJ reasons to reject the merger, and will be more costly because a greater number of documents must be produced.

It hardly bears mentioning that failure to comply with the DOJ's requirements will result in a merger's delay or abandonment, which would be a very significant cost. So, while TAR can generate great savings,<sup>138</sup> it can also

---

130. See, e.g., LIGHTHOUSE, A COMPLEX SECOND REQUEST COMES INTO FOCUS 1, 2 (2018) (analyzing 1.1 million documents for a merger).

131. *Id.*

132. *Id.* at 4.

133. Geoffrey Vance & Alison Silverstein, *McDermott and DOJ Embrace Predictive Coding*, LEGALTECH NEWS (July 9, 2013, 12:00 AM), <https://advance.lexis.com/search?crid=c81f8908-8667-4f0b-8259-a2af3f849db5&pdsearchterms=LNSDUID-ALM-LAWTNW-1202609909310&pdbyypasscitator docs=False&pdmfid=1000516&pdisurlapi=true> [https://perma.cc/XJ5B-4SV4].

134. *Id.*

135. *Id.*

136. *Id.*

137. *Id.*

138. See *supra* note 26 and accompanying text.

create unnecessary economic costs caused by delays, duplication of efforts, and overexposure. Lawyers, of course, are obligated to understand the risks involved with using TAR in a given matter.<sup>139</sup> Therefore, lawyers must consider economic costs like these when deciding how to counsel their clients, and companies must consider economic costs when deciding how to move forward with (or whether to abandon) mergers and other business deals.

## V. CONCLUSION

Lawyers should be considering the four possible pitfalls of civil sanctions, due process violations, personal liability, and economic costs when training and implementing TAR in discovery in civil litigation, criminal prosecution, and corporate law. Courts should not impose strict rules governing its use, and TAR should not require extensive or intensive court supervision. Because TAR will likely soon be deployed in all sizeable<sup>140</sup> law firms, it may expose a great number of firms and prosecutors to the four pitfalls this Article has outlined. Therefore, the actors involved should be thinking about these issues now. The lawyers and firms who start properly implementing TAR early will be the most successful at navigating the changing legal landscape and surviving the increasing integration of technology into the profession.<sup>141</sup> The firms that fail to adopt TAR, or adopt TAR but poorly implement it, will not fare well.

The developing implementation of TAR has many ramifications which attorneys, judges, TAR vendors, and others should consider. The phenomenon of technology being integrated into the legal profession parallels the agricultural revolution that took place in this country beginning in the late 20th century. In 1950, the United States had 5,388,437 farms;<sup>142</sup> by 1982, that

---

139. Vance & Silverstein, *supra* note 132.

140. For the purposes of this Article, “sizeable” is synonymous with “BigLaw.” While definitions differ, BigLaw is typically defined as a firm employing 100+ lawyers. Sally Kane, *The Definition of the BigLaw Nickname*, BALANCE CAREERS (May 7, 2018), <https://www.thebalancecareers.com/biglaw-nickname-definition-2164198> [https://perma.cc/BKS4-JS5F]. The phrase “BigLaw” also captures other characteristics of law firms that make them influential, such as the amount of revenue they generate.

141. See Blair Janis, *How Technology Is Changing the Practice of Law*, GPSOLO MAG. (May/June 2014), [https://www.americanbar.org/groups/gpsolo/publications/gp\\_solo/2014/may\\_june/how\\_technology\\_changing\\_practice\\_law/](https://www.americanbar.org/groups/gpsolo/publications/gp_solo/2014/may_june/how_technology_changing_practice_law/) [https://perma.cc/SE53-LCN6] (“The key to our future success as legal service providers lies in our ability to identify the specific lawyering areas in which we can be replaced and those in which we cannot be replaced. The most prosperous law practices in 2020 will be those that are able to successfully adjust their business models to use artificial intelligence-type tools while at the same time promoting and delivering the part of the legal service value proposition that the machines are not able to provide.”).

142. U.S. DEP’T OF COMMERCE & BUREAU OF THE CENSUS, 1982 CENSUS OF AGRICULTURE 1 (Oct. 1984), <http://usda.mannlib.cornell.edu/usda/AgCensusImages/1982/01/51/1982-01-51.pdf> [https://perma.cc/8VPW-HVWY].

number had decreased by more than half to 2,240,976<sup>143</sup> and has remained at roughly that level since.<sup>144</sup> At the same time, the market value of crops has increased exponentially.<sup>145</sup> In the same way, the market value of lawyers' work product will increase as technology replaces lower-level tasks like document review and lawyers specialize accordingly.<sup>146</sup> Lawyers should carefully and thoughtfully implement TAR according to the methods described in this Article to avoid the pitfalls of misusing TAR. The law should leave plenty of room for parties to experiment with implementing TAR by not imposing strict punishments for poorly implemented TAR, at least in the nascent stage of development and deployment.<sup>147</sup> At the same time, judges should take care to write detailed opinions about TAR protocols and disputes so that the actors in the legal community can learn from each other's successes and mistakes.<sup>148</sup> Additionally, the companies that create TAR programs should think about confidentiality issues that could arise, and take care to develop programs that can properly identify and handle privileged and confidential material, preventing inadvertent disclosure and avoiding contractual liability. TAR vendors should also teach consumers how to train TAR with an effective seed set and how to monitor TAR's performance. Finally, of course, companies and lawyers should consider the economic and normative consequences of employing TAR. TAR is an amazing tool that can greatly benefit those who use it, and now is the time to consider the sticky consequences of poorly implementing TAR.

---

143. *Id.*

144. In 2012, the number of U.S. farms was 2,109,303. U.S. DEP'T OF AGRIC. & NAT'L AGRIC. STATISTICS SERV., 2012 CENSUS OF AGRICULTURE 7 (May 2014), [https://www.nass.usda.gov/Publications/AgCensus/2012/Full\\_Report/Volume\\_1\\_Chapter\\_1\\_US/usv1.pdf](https://www.nass.usda.gov/Publications/AgCensus/2012/Full_Report/Volume_1_Chapter_1_US/usv1.pdf) [<https://perma.cc/N5TX-5KWW>].

145. *Id.* (showing that the value of crops increased from 62,256,087 (times \$1,000) in 1982 to 212,397,074 (times \$1,000) in 2012).

146. *See Specialization*, INVESTOPEDIA, <https://www.investopedia.com/terms/s/specialization.asp> [<https://perma.cc/UB4S-ZDXC>] (last visited Oct. 31, 2018) (explaining that specialization is "a method of production whereby an entity focuses on the production of a limited scope of [services] to gain a greater degree of efficiency"). Greater efficiency creates value because it lowers costs. *See Efficiency*, INVESTOPEDIA, <https://www.investopedia.com/terms/e/efficiency.asp> [<https://perma.cc/YM6W-4SDA>] (last visited Oct. 31, 2018) (explaining that efficiency uses the least amount of inputs to achieve the highest amount of output).

147. *See* THE ELECTRONIC DISCOVERY INSTITUTE, *supra* note 39, at 173 (stating that ESI protocols "should be short and allow flexibility to each party for purposes of conducting their own reviews").

148. *See* *Rio Tinto PLC v. Vale S.A.*, 306 F.R.D. 125, 129 (S.D.N.Y. 2015) ("The Court has written this Opinion, rather than [sic] merely signing the parties' stipulated TAR protocol, because of the interest within the e-discovery community about TAR cases and protocols.").





# MEASURING AND PROTECTING PRIVACY IN THE ALWAYS-ON ERA

*Dan Feldman<sup>†</sup> & Eldar Haber<sup>††</sup>*

## ABSTRACT

Data-mining practices have greatly advanced in the interconnected era. What began with the internet now continues through the Internet of Things (IoT)—whereby users can constantly be connected to the internet through various means, like televisions, smartphones, wearables, and computerized personal assistants, among other “things.” As many of these devices constantly receive and transmit data, the increased use of IoT devices might lead society into an “always-on” era, where individuals are “datafied”—constantly quantified and tracked.

This situation leads to difficult policy choices. On the one hand, the current sectorial regulatory approach, which protects privacy through regulating information gathering or use only in pre-defined industries or specified cohorts, greatly risks individuals’ privacy. On the other hand, strict privacy regulations might diminish data utility, which is crucial for technological development and innovation. There is a tradeoff between data utility and privacy protection, and the sectoral approach to privacy does not strike the right balance. This Article proposes a technological solution that might help. Relying on a method called differential privacy, this Article suggests adding “noise” to data deemed sensitive *ex ante*. In short, combining computational solutions with formulas that measure the probability of data sensitivity will better protect privacy in the always-on era.

This Article introduces legal and computational methods that could be used by IoT service providers and can optimally balance the tradeoff between data utility and privacy. Part II discusses the protection of privacy under the sectoral approach and estimates what values this approach embeds. Part III discusses privacy protection in the “always-on” era. This Part assesses how technological changes have shaped the sectoral regulation regime, then discusses why IoT devices negatively impact privacy, and finally explores the potential regulatory mechanisms that might meet the challenges of the “always-on” era. After concluding that the current regulatory framework is severely limited in protecting individuals’ privacy, this Article discusses technology as a solution in Part IV. This Part proposes a new computational model that relies on differential privacy and a modern invention called private coresets. This proposed model introduces “noise” to users’ data according to the probability that the IoT device collects sensitive data, in order to preserve individuals’ privacy and ensure service providers can utilize the data at the same time.

---

DOI: <https://doi.org/10.15779/Z38HH6C63R>

© 2020 Dan Feldman & Eldar Haber.

<sup>†</sup> Senior Lecturer, Computer Science Department, University of Haifa; Director, Robotics & Big Data Lab, University of Haifa; Faculty member, Center for Cyber, Law and Policy (CCLP), University of Haifa.

<sup>††</sup> Senior Lecturer, Faculty of Law, University of Haifa; Visiting Professor, Bocconi University, Italy; Faculty member, Center for Cyber, Law and Policy (CCLP), and Haifa Center for Law and Technology (HCLT), University of Haifa. This work was supported by the Center for Cyber Law & Policy at the University of Haifa in conjunction with the Israel National Cyber Directorate in the Prime Minister’s Office.

## TABLE OF CONTENTS

<b>I.</b>	<b>INTRODUCTION .....</b>	<b>198</b>
<b>II.</b>	<b>THE SECTORAL PRIVACY PUZZLE.....</b>	<b>200</b>
	A. PRIVACY AND DATA PROTECTION IN THE UNITED STATES .....	201
	B. SECTORAL PRIVACY IN CONTEXT .....	211
<b>III.</b>	<b>PROTECTING PRIVACY IN AN ALWAYS-ON ERA .....</b>	<b>213</b>
	A. SECTORAL PRIVACY AND TECHNOLOGY .....	214
	B. PRIVACY IN THE “ALWAYS-ON” ERA .....	216
	C. ALWAYS-ON REGULATIONS .....	220
<b>IV.</b>	<b>REGULATING THE ALWAYS-ON ERA THROUGH TECHNOLOGY .....</b>	<b>227</b>
	A. TECHNOLOGY AS A SOLUTION .....	227
	B. DIFFERENTIAL PRIVACY USING CORESETS.....	233
	C. MEASURING NOISE VIA THE PROBABILITY OF SENSITIVITY .....	243
<b>V.</b>	<b>CONCLUSION.....</b>	<b>249</b>

### I. INTRODUCTION

Technology has posed many threats to individuals’ privacy throughout history. Digitization further expanded the risks to privacy by facilitating an increase in data mining and storage capacities. Yet, the internet is not the most threatening technological innovation to privacy, as a newer technological innovation might increase such risks substantially. In what is termed the Internet of Things (IoT)—where ordinary household objects become computerized and connected to the internet—data collection and retention capabilities increase dramatically. This change has enabled service providers to collect massive amounts of sensitive data about their users. IoT devices might capture, to name but a few examples, conversations, imagery, videos, geolocation, biometric data, and even vital signs (e.g., blood pressure or heart rate).<sup>1</sup>

Historically, Congress was quite responsive to technological inventions and digitization that potentially threatened individuals’ privacy. Beginning in the 1970s, Congress reacted to privacy threats by crafting a series of federal laws that protect privacy within specified industries or cohorts. These laws can be categorized as protecting financial privacy, educational privacy, health

---

1. See Paul Ohm, *Sensitive Information*, 88 S. CAL. L. REV. 1125, 1143–44 (2015).

privacy, children's privacy, and consumer data privacy.<sup>2</sup> Under this so-called sectoral approach to privacy, regulation applies to a specific context of information gathering or use and is directed at specific pre-defined industries or specific cohorts.<sup>3</sup>

However, the emergence of IoT might make such a sectoral approach to privacy obsolete. IoT is driving society into an "always-on" era in which, more so than ever before, individuals are constantly surrounded by devices that capture their daily routines, including highly sensitive data. Consequently, it is problematic that the current sectoral privacy protection approach does not generally apply to the service providers that offer IoT services; rather, it applies based on the type of data or sector. This leaves IoT users no proper safeguards against such datafication. Thus, IoT exacerbates the limitations of the sectoral approach in the "always-on" era.

Protecting privacy in the always-on era necessitates rethinking the sectoral approach altogether. But before implementing non-sectoral, strict privacy regulatory interventions, policymakers must carefully balance the legitimate interests of IoT companies and users. While marketing is often cited as one of the main reasons for comprehensive data mining, IoT companies rely on data for various other purposes, such as the development of their services. But users' privacy should not be abandoned to accommodate the companies' needs. In other words, decision makers must find a proper way to ensure both data utility and users' privacy.

Technology can be the panacea for a proper tradeoff. While technological solutions, such as de-anonymization or encryption, proved insufficient to protect privacy in the past, other solutions could prove otherwise. This Article proposes a new mathematical model, relying mostly on a method called differential privacy.<sup>4</sup> This new model introduces "noise" that hides information about individual users in data deemed "sensitive" *ex ante*, depending on various parameters, such as the type of the IoT device, the sensors on the device, the types of data gathered, and the ways data is used. In other words, using technology, this Article proposes a mathematical solution that can both aid in protecting the values embedded in the sectoral approach<sup>5</sup> and ensure extensive privacy protection across IoT devices without sacrificing the utility of the data.

---

2. *See infra* Part II.

3. *Id.*

4. *See infra* Section IV.B.

5. The sectoral approach particularly protects financial privacy, educational privacy, health privacy, children's privacy, and consumer data privacy. *See infra* Section II.B.

Part II discusses the protection of privacy under the sectoral approach and extracts the values embedded in that approach. Part III discusses privacy protection in the always-on era. It assesses how technological changes have shaped sectoral regulation, why privacy is negatively impacted by IoT devices, and whether new regulatory mechanisms to solve the challenges arising in the always-on era are viable. After showing that the current regulatory framework is severely limited in protecting individuals' privacy, Part IV discusses the use of technology as a panacea and presents a new mathematical model that relies mostly on differential privacy. The proposed model introduces "noise" into users' data to preserve individuals' privacy—based on the probability of data sensitivity of the IoT device—while enabling service providers to utilize the data. Part V ends by suggesting that any privacy model, including the proposed model in this Article, must be further examined and recalibrated to embed the values that society wishes to protect.

## II. THE SECTORAL PRIVACY PUZZLE

Many scholars have attempted to articulate the need to protect privacy and what it should stand for,<sup>6</sup> but privacy has no clear or single definition.<sup>7</sup> The modern view traces back to Samuel Warren and Louis Brandeis, who defined the right to privacy as the "right to be let alone."<sup>8</sup> In time, privacy literature came to deal extensively with forming a theoretical conception of privacy that furnished a better understanding of that right. Alan Westin offered the "control theory," which conceptualizes privacy as the right to control information about oneself.<sup>9</sup> Ruth Gavison and Anita Allen conceptualized privacy within a "limited access theory," which posits that privacy is "related to our concern over our accessibility to others."<sup>10</sup> Finally, Helen Nissenbaum proposed a conceptual framework of privacy as contextual integrity that links the protection of personal information to the norms of specific contexts.<sup>11</sup>

---

6. See, e.g., William L. Prosser, *Privacy*, 48 CALIF. L. REV. 383, 389 (1960) (offering four types of privacy invasions); Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477 (2006) (offering a framework for a better understanding of privacy).

7. See Daniel J. Solove, *Conceptualizing Privacy*, 90 CALIF. L. REV. 1087, 1090 (2002).

8. Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 193 (1890).

9. See ALAN F. WESTIN, *PRIVACY AND FREEDOM* 7 (1967) ("[T]he claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about themselves is communicated to others.").

10. Ruth Gavison, *Privacy and the Limits of Law*, 89 YALE L.J. 421, 523 (1980). See generally ANITA L. ALLEN, *UNEASY ACCESS: PRIVACY FOR WOMEN IN A FREE SOCIETY* (1988).

11. See generally HELEN NISSENBAUM, *PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE* (2010).

Although the scholarly debate on how best to articulate the right to privacy continues today, American policymakers formally acknowledged the existence of this right only after Warren and Brandeis' outcry. Upon calling attention to privacy, policymakers inspired broader interest in recognizing a right to it, and along with the influence of William Prosser, common law tort actions that protect privacy emerged.<sup>12</sup> Subsequently, courts also acknowledged privacy rights in various areas of decision making, usually related to intimately personal matters, such as certain reproductive rights.<sup>13</sup> States have also enacted their own privacy laws.<sup>14</sup> But privacy rights remained unrecognized at the federal level until the 1970s, when Congress began enacting legislations aimed at protecting privacy in particular industries or specific contexts. This regulatory approach has come to be known as the "sectoral approach."

#### A. PRIVACY AND DATA PROTECTION IN THE UNITED STATES

The American protection of information privacy—also known as data protection—mainly adopts the sectoral approach.<sup>15</sup> Unlike an omnibus approach to privacy, embraced by many jurisdictions such as the European Union,<sup>16</sup> a sectoral approach generally protects information privacy only within a specific context of information-gathering or use and is usually directed only to specific pre-defined industries or specific cohorts.

---

12. See Prosser, *supra* note 6, at 386–89; see also Solove, *supra* note 7, at 1100; RESTATEMENT (SECOND) OF TORTS § 652 (AM. LAW INST. 1965). The four privacy torts that are generally recognized in the United States are intrusion, public disclosure of private facts, false light, and appropriation. Notably, however, tort law is primarily state-legislated, so the recognition of privacy torts could differ between states. See DANIEL J. SOLOVE & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW 77–231 (3d ed. 2009).

13. See, e.g., *Griswold v. Connecticut*, 381 U.S. 479, 486 (1965) (acknowledging a right to privacy for married couples' use of contraceptives); *Eisenstadt v. Baird*, 405 U.S. 438, 443 (1972) (protecting the right of unmarried individuals to possess contraception); *Roe v. Wade*, 410 U.S. 113 (1973) (acknowledging a right to privacy in a woman's decision to have an abortion under the Due Process Clause of the Fourteenth Amendment); *Lawrence v. Texas*, 539 U.S. 558 (2003) (invalidating sodomy laws due to sexual privacy).

14. See Daniel J. Solove & Paul M. Schwartz, *An Overview of Privacy Law*, in PRIVACY LAW FUNDAMENTALS 145–47 (2015).

15. The conventional concept of information privacy refers to protecting a right to control one's personal data. Beyond information rights, privacy rights could also be spatial, regarding individual's physical sphere of control, or decisional, regarding control over personal choices. See Joel R. Reidenberg, *Privacy Wrongs in Search of Remedies*, 54 HASTINGS L.J. 877, 878–89 (2003); see also Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1202–05 (1998).

16. An omnibus approach refers to "one overarching law that regulates privacy consistently across all industries." See Daniel Solove, *The Growing Problems with the Sectoral Approach to Privacy Law*, TECHPRIVACY (Nov. 13, 2015), <https://teachprivacy.com/problems-sectoral-approach-privacy-law> [<https://perma.cc/X5HV-D3YC>].

However, American privacy protection is not entirely sectoral. The right to privacy was interpreted as being embedded in the Bill of Rights, perhaps most evidently in the First, Third, Fourth, and Fifth Amendments.<sup>17</sup> In addition, states sometimes protect privacy within their constitutions or simply legislate state privacy statutes.<sup>18</sup> For example, states have used tort law, or legislated specific privacy or data breach notification statutes, to protect privacy.<sup>19</sup> On both state and federal levels, laws to some extent protect the privacy of data flowing through specific channels of communication, regardless of the data's potential sensitivity.<sup>20</sup> For example, some laws set boundaries on engaging in wiretapping, accessing stored communication, and obtaining data from pen register devices.<sup>21</sup> Some laws also regulate when private entities must keep records for investigatory purposes<sup>22</sup> or facilitate governmental investigations.<sup>23</sup> In other instances, privacy is regulated based on the notion of protecting the public from governmental intrusion with respect

---

17. The First Amendment protects privacy by permitting the right to speak anonymously and freedom of association. The Third Amendment protects privacy by restricting the government from requiring soldiers to reside in people's houses. The Fourth Amendment prevents the government from conducting "unreasonable searches and seizures," which may implicate individuals' privacy interests. The Fifth Amendment protects individuals against self-incrimination. See U.S. CONST. amends. I, III–V; see also Daniel J. Solove, *A Brief History of Information Privacy Law*, in PROSKAUER ON PRIVACY, 1, 1–5 (2006); Solove & Schwartz, *supra* note 14, at 41.

18. See Scott A. Sundstrom, *You've Got Mail! (And the Government Knows It): Applying the Fourth Amendment to Workplace E-Mail Monitoring*, 73 N.Y.U. L. REV. 2064, 2076–77 (1998).

19. See Solove & Schwartz, *supra* note 14, at 44; Priscilla M. Regan, *Federal Security Breach Notifications: Politics and Approaches*, 24 BERKELEY TECH. L.J. 1103, 1108–12 (2009).

20. See Ohm, *supra* note 1, at 1136 (articulating these laws as "protected channel laws").

21. See The Communications Act of 1934, Pub. L. No. 73-416, § 605, 48 Stat. 1064, 1103–04 (1934) (regulating the practice of wiretapping limitedly); Omnibus Crime Control and Safe Streets Act of 1968, Pub. L. No. 90-351, 82 Stat. 197, 211 (1968) (codified as amended at 18 U.S.C. §§ 2510–2522 (2012)) (regulating wiretapping); Electronic Communications Privacy Act of 1986, Pub. L. No. 99-508, 100 Stat. 1848 (1986) (codified as amended at scattered sections of 18 U.S.C.) (revising the Wiretap Act and adding two other acts to deal with technological developments: The Stored Communications Act (SCA), which regulates access to both the content and metadata stored by electronic communications services; and the Pen Register Act, which regulates devices that obtain information about calls).

22. See, e.g., Bank Secrecy Act of 1970, Pub. L. No. 91-508, 84 Stat. 1114 (1970) (mandating federally insured banks and other financial institutions to aggregate financial data and report in order to assist law enforcement agencies in conducting financial investigations); see also Solove, *supra* note 17, at 1–29.

23. See, e.g., Communications Assistance for Law Enforcement Act (CALEA) of 1994, Pub. L. No. 103-414, 108 Stat. 4279 (1994) (requiring telecommunication providers to facilitate government interceptions of communications and surveillance under some circumstances).

to what data the state is entitled to collect<sup>24</sup> or how state agencies should handle acquired data.<sup>25</sup>

Some forms of privacy protections, although important, are not part of this Article's general evaluation of privacy. First, Constitutional protection is excluded, as it does not relate to the practices of private actors.<sup>26</sup> Second, the states' protection of privacy rights is also excluded, as it is incoherent and would apply inconsistently depending on the individual policymaker. Finally, public data collection and retention protections—in contrast to data collection and retention by private parties—are excluded because they are less relevant to the discussion on IoT, which is a field currently controlled mainly by private companies.

---

24. *See, e.g.*, Foreign Intelligence Surveillance Act (FISA) of 1978, Pub. L. No. 95-511, 92 Stat. 1783 (1978) (regulating foreign intelligence gathering within the United States); Privacy Protection Act of 1980, Pub. L. No. 96-440, 94 Stat. 1879 (codified at 42 U.S.C. § 2000(a)(a) (2012)) (restricting the government's ability from conducting unlawful searches and seizures of work product of the press and media); Electronic Communications Privacy Act of 1986, Pub. L. No. 99-508, 100 Stat. 1848 (1986) (codified as amended at 18 U.S.C. §§ 2510–2522 (2012)) (regulating electronic storage and surveillance); Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act of 2001, Pub. L. No. 107-56, 115 Stat. 272 (2001) (codified as amended at scattered sections of U.S.C. (2012)) (amending various acts such as the ECPA and FISA, loosening requirements for data gathering by law enforcement agencies under some circumstances); *see also* DANIEL J. SOLOVE & PAUL M. SCHWARTZ, *PRIVACY LAW FUNDAMENTALS* 4 (2017). Individuals' privacy interests in personally identifiable information in the possession of federal agencies received further protection. *See* Privacy Act of 1974, Pub. L. No. 93-579, 88 Stat. 1896 (codified as amended at 5 U.S.C. § 552(a) (2012)). The Federal Trade Commission (FTC) is also tasked with protecting consumers from “unfair or deceptive acts” under § 5 of the Federal Trade Commission Act and it generally regulates commercial collection, use, and release of data under some circumstances. *See* The Federal Trade Commission Act (FTC Act), Pub. L. No. 63-203, 38 Stat. 717, 719 (codified at 15 U.S.C. §§ 45(a), 6505(a) (2012)). For more on the FTC's role in the field of data protection, see generally FED. TRADE COMM'N, *PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE, RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS* 30 (Mar. 2012), <http://www.ftc.gov/os/2012/03/120326privacyreport.pdf> [<https://perma.cc/MHB5-R8TX>].

25. *See, e.g.*, Computer Matching and Privacy Protection Act of 1988, Pub. L. No. 100-503, 102 Stat. 2507 (1988) (codified at 5 U.S.C. § 552(a) (2012)); *see also* Driver's Privacy Protection Act of 1994, Pub. L. No. 103-322, 108 Stat. 2099 (codified at 18 U.S.C. §§ 2721–2725 (2012)) (regulating the states' authority to disclose personal driver records by prohibiting, with exceptions, the disclosure or sale of drivers records without obtaining prior consent from the individual); Solove, *supra* note 17, at 1–37. Notably, the law also governs privacy through the lens of government records, granting individuals rights regarding their personal data stored in government records systems. *See, e.g.*, Privacy Act of 1974, Pub. L. No. 93-579, 88 Stat. 1896 (codified at 5 U.S.C. § 552(a) (2012)) (regulating certain kinds of collection and use of records by certain federal agencies, excluding the private sector, state, and local agencies); Solove & Schwartz, *supra* note 14, at 26.

26. *See* Reidenberg, *supra* note 15, at 879.

Instead, this Article focuses on the sectoral approach at its core, which is the federal protection of privacy as applied to the private sector. To better understand the sectoral approach, Section II.B first reviews and analyzes the five categories of sectoral privacy according to federal statutes regulating private parties: financial privacy, educational privacy, health privacy, children's privacy, and consumer data privacy.<sup>27</sup>

The first category is financial privacy, where financial information is granted federal protection under some circumstances.<sup>28</sup> The first federal law that regulated private use and dissemination of information was the Fair Credit Reporting Act of 1970 (FCRA).<sup>29</sup> The FCRA generally regulates the use of credit reports, supervising “the collection, maintenance and dissemination of ‘consumer reports’ ”<sup>30</sup> and require “consumer reporting agencies to maintain procedures to ensure ‘maximum possible accuracy.’ ”<sup>31</sup> It was enacted due to privacy concerns regarding “exclusion, secondary use, and disclosure” of data gathered by credit bureaus.<sup>32</sup> Essentially, the FCRA imposes obligations on consumer reporting agencies and provides individuals with certain rights and control over personal financial records held by credit reporting companies.<sup>33</sup>

In 1978, the Right to Financial Privacy Act (RFPA) further regulated financial data.<sup>34</sup> RFPA sets limits on financial institutions' disclosure of

---

27. It should be further noted that there are federal acts that apply to the private sector, but focus on conducting an activity by a specific entity, which will be excluded from this analysis. One example is the Employee Polygraph Protection Act of 1988 (EPPA), which regulates the use of polygraphs by private employers. *See* Employee Polygraph Protection Act of 1988, Pub. L. No. 100-347, 102 Stat. 646 (1988) (codified at 29 U.S.C. §§ 2001–2009 (2012)). For more on federal privacy acts, see Solove & Schwartz, *supra* note 14, at 42–44.

28. Beyond the sectoral laws discussed in this Part, other federal laws also relate to financial regulation. *See, e.g.*, Bank Secrecy Act of 1970, Pub. L. No. 91-508, 84 Stat. 1114 (1970).

29. *See* Daniel J. Solove & Chris Jay Hoofnagle, *A Model Regime of Privacy Protection*, 2006 U. ILL. L. REV. 357, 359–60 (2006).

30. *Id.* (citing 15 U.S.C. § 1681 (2012)).

31. *Id.* (citing 15 U.S.C. § 1681(e)(b)). The FCRA was amended by the Fair and Accurate Credit Transaction Act of 2003 (FACTA) with the aim to prevent identity theft and promote accurate credit rating by requiring credit reporting agencies to provide consumers with an annual credit report. *See* Fair and Accurate Credit Transaction Act of 2003, Pub. L. No. 108-159, 117 Stat. 1952, 1968 (2003); *see also* Solove & Schwartz, *supra* note 14, at 42–43.

32. 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, 145 (2014).

33. One example is the right of individuals to request a copy of their credit report. *See* 15 U.S.C. § 1681(b) (2012) (noting more permissible purposes of consumer reports).

34. *See* 12 U.S.C. §§ 3401–3422 (2012); Solove, *supra* note 17, at 1–30.



financial records to a government authority without a warrant or subpoena.<sup>35</sup> Under this act, an unauthorized disclosure by a financial institution or by any government agency obtaining financial records could result in civil penalties.<sup>36</sup> In addition, under the financial privacy category one might also include the Financial Services Modernization Act of 1999, also known as the Gramm-Leach-Bliley Act (GLBA).<sup>37</sup> Although the GLBA does not necessarily relate directly to financial data, it generally regulates financial institutions' processing of personal information,<sup>38</sup> concerning the collection, use, and disclosure of personally identifiable financial information.<sup>39</sup> With some exceptions, the GLBA obliges financial services entities to secure customer records, provide notice and opt-out procedures to consumers before sharing their information with some third parties, and disclose their privacy practices.<sup>40</sup>

The second category is educational privacy, which affords legal protection of student information privacy. A key example is the Family Educational Rights and Privacy Act of 1974 (FERPA).<sup>41</sup> FERPA protects the privacy of school records by regulating access to educational records, students' private records, and other information maintained by educational institutes, such as health records, psychological evaluations, and additional information directly related to students.<sup>42</sup> With some exceptions, students or their parents must consent before an institution may hand over personally identifiable information.<sup>43</sup> FERPA also grants parents and students access to students' files, in order to challenge false or harmful information contained in them.<sup>44</sup>

---

35. See 12 U.S.C. §§ 3401–3422; George B. Trubow & Dennis L. Hudson, *The Right to Financial Privacy Act of 1978: New Protection from Federal Intrusion*, 12 J. MARSHALL J. PRAC. & PROC. 487, 497 (1979); see also Solove, *supra* note 17, at 1–30.

36. 12 U.S.C. § 3417.

37. See The Financial Services Modernization Act (Gramm-Leach-Bliley) Act, Pub. L. No. 106-102, 113 Stat. 1338 (1999) (codified as amended at scattered sections of 12 & 15 U.S.C.).

38. See Solove, *supra* note 17, at 1–39.

39. Defined as “nonpublic personal information.” See 15 U.S.C. §§ 6801–6802 (2012).

40. See Theodore Rostow, *What Happens When an Acquaintance Buys Your Data: A New Privacy Harm in the Age of Data Brokers*, 34 YALE J. ON REG. 667, 677 (2017). This rationale arose partly from “privacy concerns regarding consumer financial information.” Jolina C. Cuaresma, *The Gramm-Leach-Bliley Act*, 17 BERKELEY TECH. L.J. 497, 497 (2002).

41. See Family Educational Rights and Privacy Act of 1974, Pub. L. No. 93-380, 88 Stat. 57 (1974) (codified as amended at 20 U.S.C. § 1232(g) (2012)); 34 C.F.R. § 99 (2018).

42. See Solove & Schwartz, *supra* note 14, at 42; see also Dalia Topelson et al., *Privacy and Children's Data—An Overview of the Children's Online Privacy Protection Act and the Family Educational Rights and Privacy Act*, 23 BERKMAN CTR. RES. PUB. 1, 2 (Nov. 14, 2013), <http://dx.doi.org/10.2139/ssrn.2354339> [<https://perma.cc/87PP-UUX7>].

43. 34 C.F.R. § 99.31 (2018).

44. See Ohm, *supra* note 1, at 1157–58.

Notably, the scope of educational privacy on the federal level is rather limited. FERPA only applies to educational institutions or agencies that receive federal funds from the U.S. Department of Education (DoE).<sup>45</sup> While educational privacy on the federal level is limited in scope, the rationale behind FERPA can readily be understood as contextual. This law seeks to protect the confidentiality of certain records accumulated in educational facilities because these institutions collect information that might be highly sensitive.

The third category is health privacy, where health information deserves stronger data protection than other ‘regular’ data. The first federal acknowledgment of the importance of health data was embedded in the Freedom of Information Act (FOIA).<sup>46</sup> This law generally mandates disclosure of information upon FOIA requests. But it exempts public access to (1) government records for “personnel and medical files and similar files the disclosure of which would constitute a clearly unwarranted invasion of personal privacy,”<sup>47</sup> and (2) “records or information compiled for law enforcement purposes . . . [that] could reasonably be expected to constitute an unwarranted invasion of personal privacy,” potentially including health data.<sup>48</sup>

Apart from exempting health information from FOIA requests, Congress afforded more federal protection to health privacy in 1996 due to a perceived need to digitize health information and preserve the confidentiality of such information. The Health Insurance Portability and Accountability Act (HIPAA) of 1996, which regulates privacy in health records, mandates the Secretary of the Department of Health and Human Services (HHS) to promulgate rules that govern how states must protect the confidentiality of certain health information.<sup>49</sup> The initial rationale was not the protection of privacy per se but the creation of new standards with the goal of “reducing administrative costs.”<sup>50</sup>

Today, health information is mainly protected on the federal level under what is collectively termed the HIPAA Privacy Rule.<sup>51</sup> Because of a perceived

---

45. See 20 U.S.C. § 1232(g)(a)(1)(A) (2012); Topelson et al., *supra* note 42, at 3. The entities covered by FERPA include elementary and secondary schools, school districts, colleges and universities, and state educational agencies, along with other institutions that provide educational services. See 34 C.F.R. § 99.1(a)(1–2) (2018).

46. See Helm & Georgatos, *supra* note 32, at 147; see also Freedom of Information Act (FOIA), Pub. L. No. 90-23, 81 Stat. 54 (1966) (codified at 5 U.S.C. § 552(a)(3)(A) (2012)).

47. See 5 U.S.C. § 552(b)(6); Helm & Georgatos, *supra* note 32, at 147.

48. 5 U.S.C. § 552(b)(7)(C).

49. See Solove, *supra* note 17, 1–38.

50. H.R. REP. NO. 104-497, at 61 (1996); Ohm, *supra* note 1, at 1150.

51. See Ohm, *supra* note 1, at 1150–51; see also Standards for Privacy of Individually Identifiable Health Information, 67 Fed. Reg. 53, 182 (Aug. 14, 2002) (codified at 45 C.F.R. pts. 160 & 164 (2018) (establishing national standards for protecting certain health

need to strengthen enforcement and expand patient rights, Congress further revised HIPAA in 2009 under the Health Information Technology for Economic and Clinical Health Act (HITECH).<sup>52</sup> The HITECH Act broadened HHS's authority to encompass "business associates" and all businesses that receive information from entities covered by HIPAA.<sup>53</sup> It also added a security breach notification provision<sup>54</sup> and dramatically increased the penalties for HIPAA violations. The HIPAA final rule (or Omnibus Rule) was released in 2013.<sup>55</sup>

HIPAA regulation applies to health data held by covered entities or business associates.<sup>56</sup> Health data is any information, oral or recorded, in any form or medium, that is created or received by various defined entities,<sup>57</sup> which

[r]elates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual.<sup>58</sup>

---

information); Final Omnibus HIPAA Rule Preamble, 78 Fed. Reg. 5567 (Jan. 25, 2013) (expanding the definition of business associate and the reach of HIPAA). HIPAA was set to come into force in 2000, but did so only on April 14, 2003. Accordingly, the HIPAA Security Rule was finalized in 2003, but compliance was set for April 21, 2005. *See* Daniel J. Solove, *HIPAA Turns 10: Analyzing the Past, Present, and Future Impact*, 84 J. AM. HEALTH INFO. MGMT. ASS'N 22, 24–25 (2013).

52. Health Information Technology for Economic and Clinical Health Act (HITECH), Pub. L. No. 111-5, 123 Stat. 226 (2009) (codified as amended at scattered sections of 42 U.S.C.). This Act "was passed as a subsection of the American Recovery and Reinvestment Act (ARRA) of 2009." Kimberly L. Rhodes & Brian Kunis, *Walking the Wire in the Wireless World: Legal and Policy Implications of Mobile Computing*, 16 J. TECH. L. & POL'Y 25, 40 (2011); *see also* Solove, *supra* note 17, at 26–28.

53. HITECH also promulgated a data breach notification requirement. *See* HITECH, *supra* note 52; *see also* Solove, *supra* note 17, at 26.

54. *See* 45 C.F.R. §§ 164.400–414 (2018).

55. Modifications to the HIPAA Privacy, Security, Enforcement, and Breach-Notification Rules under the HITECH and the Genetic Information Nondiscrimination Act; Other Modifications to the HIPAA Rules, 78 Fed. Reg. 5566 (Jan. 25, 2013). For more on the HIPAA rule, *see* Frank Pasquale & Tara Adams Ragone, *Protecting Health Privacy in an Era of Big Data Processing and Cloud Computing*, 17 STAN. TECH. L. REV. 595, 608–20 (2014).

56. *See* 45 C.F.R. §§ 160.102–103 (2018).

57. These entities include health care provider, health plan, public health authority, employer, life insurer, school or university, or health care clearinghouse. *See* 45 C.F.R. § 160.103.

58. *Id.*

HIPAA “require[s] stronger protections and affirmative consumer consent for certain uses of financial and health data, like marketing.”<sup>59</sup>

The HIPAA Privacy Rule governs protected health information (PHI), which includes any “individually identifiable health information” that these entities hold, such as demographic data and information relating to a patient’s medical background and care.<sup>60</sup> It requires, *inter alia*, the anonymization of health data by removal of various types of identifiers.<sup>61</sup> The HIPAA Security Rule provides standards for protecting PHI in electronic form that the covered entity “creates, receives, maintains, or transmits.”<sup>62</sup> Under HIPAA, covered entities are required to: (1) designate a privacy official and develop and implement privacy policies; (2) ensure that only the “minimum necessary PHI be accessed and used” and that people authorize disclosure of their PHI (with few exceptions); (3) provide patients with a set of rights; and (4) mandate security safeguards.<sup>63</sup>

The most recent health privacy legislation was passed in 2008: the Genetic Information Nondiscrimination Act (GINA).<sup>64</sup> GINA regulates the use of genetic predisposition to disease by group health plans and health insurers when basing coverage decisions or setting premiums.<sup>65</sup> It also restricts employers from using genetic information when making personnel decisions affecting their employees.<sup>66</sup>

The fourth category is children’s privacy, wherein Congress acknowledged the importance of protecting children’s privacy online.<sup>67</sup> The Children’s Online

---

59. See Andrea Reichenbach, *Defining ‘Sensitive’ in World of Consumer Data*, ACXION (July 27, 2015), <https://www.acxiom.com/blog/defining-sensitive-world-consumer-data/> [<https://perma.cc/RWQ4-SJ33>].

60. See HIPAA Privacy Rule, 45 C.F.R. § 164.514(b)–(c) (2018); Rostow, *supra* note 40, at 676–77.

61. See 45 C.F.R. § 164.514(b)–(c). More closely, the HIPAA Privacy Rule requires that the information will neither identify an individual nor provide “a reasonable basis to believe that the information can be used to identify an individual.” § 164.514(a). This could be achieved either by a statistical or a safe harbor standard (which for the latter, must include the suppression or generalization of eighteen enumerated identifiers). See Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1737 (2010).

62. 45 C.F.R. § 160.103(4)(i).

63. See Solove, *supra* note 17, at 26.

64. Genetic Information Nondiscrimination Act (GINA) of 2008, Pub. L. No. 110-233, 122 Stat. 881 (2008) (codified in scattered sections of 26, 29, and 42 U.S.C.).

65. *Id.* §§ 101–105.

66. See Ohm, *supra* note 1, at 1137; GINA, §§ 201–205. For more on GINA, see generally Jennifer J. Lee, Note, *The First Civil Rights Act of the 21st Century: Genetic Information Nondiscrimination Act of 2008*, 4 I/S: J.L. & POL’Y FOR INFO. SOC’Y 779 (2008).

67. It is worth mentioning that Congress also sought to regulate the exposure of children to inappropriate materials online by enacting the Child Online Protection Act, but it eventually

Privacy Protection Act (COPPA) of 1998 regulates the use of children's personal information on the internet.<sup>68</sup> COPPA is supplemented by a rule issued by the FTC known as the COPPA Rule.<sup>69</sup> Both forms of regulation apply to Online Service Providers (OSPs)<sup>70</sup> that target children under age thirteen or knowingly collect personal information from them.<sup>71</sup> They were intended to "prohibit unfair or deceptive acts or practices in connection with

---

failed to pass constitutional muster because it placed an "impermissible burden" on speech. *ACLU v. Reno*, 217 F.3d 162, 166, 168–69 (3d Cir. 2000) (referencing The Child Online Protection Act, Pub. L. No. 105-277, 112 Stat. 2681–736 (1998)).

68. See Kathryn C. Montgomery & Jeff Chester, *Data Protection for Youth in the Digital Age: Developing a Rights-based Global Framework*, 1 EUROPEAN DATA PROTECTION L. REV. 277, 279–84 (2015). It should be noted that the Family Educational Rights and Privacy Act also regulates children's informational privacy and family privacy. FERPA, however, applies only to the release of educational records to unauthorized persons by educational institutions. See The Family Educational Rights and Privacy Act (FERPA), Pub. L. No. 93-380, 88 Stat. 57 (1974) (codified at 20 U.S.C. § 1232(g) (2012)); *Family Educational Rights and Privacy Act (FERPA)* 3 U.S. DEP'T EDUC., <http://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html> [<https://perma.cc/9V2C-JBE2>] (last visited Jan. 9, 2020).

69. Children's Online Privacy Protection Rule, 64 Fed. Reg. 59,888 (Nov. 3, 1999) (codified at 16 C.F.R. § 312 (2018)) [hereinafter COPPA Rule]. The COPPA Rule took effect in April 2000 and was last updated in 2013. For the latest update, see 78 Fed. Reg. 3972 (Jan. 17, 2013).

70. COPPA refers to OSPs as "operators" and defines them as

any person who operates a website [or] online service and who collects or maintains personal information from or about the users of or visitors to such website or online services, or on whose behalf such information is collected or maintained, where such website or online service is operated for commercial purposes, including any person offering products or services for sale through that website or online service, involving commerce.

15 U.S.C. § 6501(2) (2012).

71. "Personal information" is defined as

individually identifiable information about an individual collected online, including: (1) A first and last name; (2) A home or other physical address including street name and name of a city or town; (3) Online contact information . . . ; (4) A screen or user name where it functions in the same manner as online contact information . . . ; (5) A telephone number; (6) A Social Security number; (7) A persistent identifier that can be used to recognize a user over time and across different Web sites or online services . . . ; (8) A photograph, video, or audio file where such file contains a child's image or voice; (9) Geolocation information sufficient to identify street name and name of a city or town; or (10) Information concerning the child or the parents of that child that [is collected] from the child and combine[d] with [one of the above identifiers].

16 C.F.R. § 312.2 (2018).

personally identifiable information from and about children on the internet,” and the FTC enforces both forms of these regulations.<sup>72</sup>

The fifth category is consumer data privacy, which provides that consumer data might be perceived as highly sensitive and hence need firmer protection in some contexts.<sup>73</sup> One example is the Cable Communication Policy Act (CCPA) of 1984.<sup>74</sup> The CCPA requires cable companies to maintain the confidentiality of cable subscribers’ records.<sup>75</sup> The law states that cable operators must inform subscribers about the use and assembly of personally

---

72. 15 U.S.C. §§ 6501–6505; Children’s Online Privacy Protection Rule, 64 Fed. Reg. 59,888 (Nov. 3, 1999) (codified at 16 C.F.R. § 312 (2018)); Danielle J. Garber, *COPPA: Protecting Children’s Personal Information on the Internet*, 10 J.L. & POL’Y 129, 153 (2002). An “unfair or deceptive” act or practice is a material “representation, omission or practice that is likely to mislead the consumer acting reasonably in the circumstances, to the consumer’s detriment” or a practice that “causes or is likely to cause substantial injury to consumers which is not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition.” See Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 599 (2014). Substantial injury, in this instance, could apply on both financial harms and unwarranted health and safety risks. *Id.*; 15 U.S.C. § 45 (2012) (outlawing unfair methods of competition); Fed. Trade Comm’n v. Info. Search, Inc., Civ. No. 1:06-cv-01099 (D. Md. Mar. 9, 2007) (“The invasion of privacy and security resulting from obtaining and selling confidential customer phone records without the consumers’ authorization causes substantial harm to consumers and the public, including, but not limited to, endangering the health and safety of consumers.”).

73. It should be noted that other federal laws protect consumers from abusive telecommunication or marketing practices and relate to privacy to some extent. These laws, however, are excluded from this Article’s analysis as they do not directly relate to information privacy. See, e.g., Telephone Consumer Protection Act (TCPA) of 1991, Pub. L. No. 102-243, 105 Stat. 2394 (1991) (codified at 47 U.S.C. § 227 (2012)) (regulating the collection and use of telephone numbers); Controlling the Assault of Non-Solicited Pornography and Marketing (CAN-SPAM) Act of 2003, Pub. L. No. 108-187, 117 Stat. 2699 (2003) (codified at 15 U.S.C. § 7701 (2012)) (regulating the collection and use of email addresses and restricts knowingly sending commercial messages to deceive or mislead recipients); The Junk Fax Prevention Act (JFPA) of 2005, Pub. L. No. 109-21, 119 Stat. 359 (2005) (codified at 47 U.S.C. § 609) (expanding the scope of liability for sending junk fax).

74. Cable Communications Policy Act of 1984, Pub. L. No. 98-549, 98 Stat. 2779 (1984) (codified at 47 U.S.C. § 551 (2012)). Notably, other federal and state laws also address consumer data privacy. Some types of customer information, termed “Customer Proprietary Network Information” (CPNI), regulate the use of information that relates to

the quantity, technical configuration, type, destination, location, and amount of use of a telecommunications service subscribed to by any customer of a telecommunications carrier, and that is made available to the carrier by the customer solely by virtue of the carrier-customer relationship; and information contained in the bills pertaining to telephone exchange service or telephone toll service received by a customer of a carrier.

47 U.S.C. § 222(h)(1) (2012); see also Paul M. Schwartz, *Preemption and Privacy*, 118 YALE L.J. 902, 924–25 (2009).

75. See Solove, *supra* note 17, at 1–33.

identifiable information to be collected by the company and the ways in which said information may be disclosed. The law also states specific purposes for which a cable operator may disclose or make use of personal information.<sup>76</sup>

Another example is the Video Privacy Protection Act (VPPA) of 1988.<sup>77</sup> Allegedly, the VPPA was enacted in response to the hearing of the Supreme Court nominee Robert Bork, after reporters tried to obtain the nominee's video rental history list.<sup>78</sup> The VPPA regulates the use of video rental information and generally prohibits "video tape service providers"<sup>79</sup> from disclosing personally identifiable information regarding rent or sale of video material of their customers to a third party.<sup>80</sup> Essentially, it limits some entities' disclosure of forms of video viewing habits.<sup>81</sup> Notably, the VPPA was amended in 2013 in light of new rental services, such as Netflix, ultimately weakening its privacy protections.<sup>82</sup>

In sum, the sectoral approach seeks to protect sensitive data within specified contexts. To understand how technological innovations might challenge the effectiveness of this approach, it must be broken down. A taxonomic analysis will shed some light on the core values and interests that Congress sought to protect with this approach and on new potential challenges to it.

#### B. SECTORAL PRIVACY IN CONTEXT

The rationales behind the sectoral privacy laws suggest that Congress legislates by seeking to identify which data could become sensitive in some industries or in a specific context and whether the risk of privacy harm could increase due to the context. While sensitivity of information is generally difficult to define, academic scholarship has focused mainly on four factors: "possibility of harm; probability of harm; presence of a confidential

---

76. Federal law also protects the privacy of satellite subscribers. *See* 47 U.S.C. § 338(i) (2012).

77. *See* Video Privacy Protection Act Amendments Act of 2012, Pub. L. No. 112-258, 126 Stat. 2414 (2012) (codified at 18 U.S.C. §§ 2710–2711 (2012)). For further information on the VPPA, see Schwartz, *supra* note 74, at 912.

78. Solove, *supra* note 17, at 1–34; Ohm, *supra* note 1, at 1140.

79. Notably, "video tape service provider" could be broadly interpreted to extend beyond the traditional video tape services, for example, to DVD service providers. Schwartz, *supra* note 74, at 912.

80. Personally identifiable information is defined as "information which identifies a person as having requested or obtained specific video materials or services from a video tape service provider." 18 U.S.C. § 2710(a)(3) (2012).

81. The sensitivity in this instance relates to the "title, description, or subject matter of any video tapes or other audio visual material." § 2710(b)(2)(D)(ii).

82. *See* Video Privacy Protection Act Amendments Act of 2012, Pub. L. No. 112-258, 126 Stat. 2414 (2012); Ohm, *supra* note 1, at 1141.

relationship; and whether the risk reflects majoritarian concerns.”<sup>83</sup> Without normatively evaluating whether and how American policymakers should define sensitive information, this Section strives to locate the current rationales behind the sectoral privacy regulations and the core values that Congress sought to protect using the sectoral approach.

As Section II.A above shows, Congress legislated to protect information privacy in five specific contexts. For instance, children’s data is afforded protection online because children are considered a cohort entitled to special care and assistance, and because the internet created new risks for them.<sup>84</sup> Similarly, financial transactions, educational records, health data, and some types of consumer data are considered sensitive enough to warrant protection when the probability of data retention by private parties is deemed high due to the context of its gathering.

Although Congress is not precluded from providing future federal information privacy protection in contexts other than the five sectors listed above, the current legal framework is rather limited in scope. Financial privacy, for instance, is only protected on the federal level when it is gathered by predefined private entities, all of which are, to some extent, related to the financial industry.<sup>85</sup> And health privacy is protected in the context of sensitive information usually collected by the protected institutions (e.g., health records and psychological evaluations). Similarly, educational privacy is protected only when there is a high probability of implicating sensitive data. Although children’s privacy is often vulnerable to websites’ data-mining practices, federal protection does not extend to offline activities, and even online, it is still limited in scope. Finally, consumer data privacy is only deemed sensitive in a specific context in some industries (e.g., cable subscribers’ records and the rent or sale of video materials), but not in other instances.

One might argue that, at root, the federal sectoral approach is reactive in nature, since Congress normally intervened to regulate information privacy when a new challenge arose. Especially when a new technology threatened information privacy, or perhaps when a new technology developed at an intolerable rate, Congress would search the context in which individuals’

---

83. See Ohm, *supra* note 1, at 1161.

84. For more on online risks to children, see John Palfrey et al., *Enhancing Child Safety and Online Technologies: Final Report of the Internet Safety Technical Task Force*, BERKMAN CTR. FOR INTERNET & SOC’Y (2008), [https://cyber.harvard.edu/sites/cyber.law.harvard.edu/files/ISTTF\\_Final\\_Report.pdf](https://cyber.harvard.edu/sites/cyber.law.harvard.edu/files/ISTTF_Final_Report.pdf) [<https://perma.cc/WDL4-XME4>].

85. See Solove & Schwartz, *supra* note 14, at 28; ROBERT ELLIS SMITH, BEN FRANKLIN’S WEB SITE: PRIVACY AND CURIOSITY FROM PLYMOUTH ROCK TO THE INTERNET 316–19 (2000); Solove & Hoofnagle, *supra* note 29, at 359.



privacy was at risk and then regulate the affected industries respectively.<sup>86</sup> Ultimately, Congressional legislations granted protection through the lens of the industry—or the cohort—most affected by a new technology at issue.

Naturally, the sensitive data that Congress sought to protect generally could be gathered by various industries simultaneously and was not fully protected by the current regulatory framework. Thus, it is false to assume that the protected information is only sensitive within these sectors, or that the potential harm to data subjects is only plausible within the specified contexts. And the risk of harm is even higher than before in multiple contexts due to rapid technological developments. So, the sensitivity of data must constantly remain on policymakers' agenda. IoT, perhaps the most dramatic technological innovation in recent years, especially necessitates overall scrutiny of data protection in the United States and a search for other viable solutions.

As the next Part shows, the types of information that should be deemed sensitive and the types of information privacy that deserves protection can swiftly change due to constant innovations in how data are gathered, processed, and stored. But before embarking on a normative evaluation of the new threats to information privacy posed by new technologies, it is essential to discern the role of technology in the development of sectoral privacy protection to date.

### III. PROTECTING PRIVACY IN AN ALWAYS-ON ERA

The evolution of technology in the twenty-first century is likely the most rapid in human history. But even prior to these new rapid developments, evolving technology at the time fulfilled a substantial role in establishing sectoral privacy protection. Today, the federal sectoral approach seems ever reactive to new technologies, and it has not responded to the rapid pace of technological development in recent years, including IoT.

To better understand the potential risks that IoT might impose on the right to information privacy and to see if legal intervention could substantially reduce such risks, this Part will proceed as follows. Section A briefly discusses technology's impact on shaping sectoral privacy prior to IoT. Section B introduces the so-called "always-on" era, and the challenges this era poses in the context of sectoral privacy. Section C examines whether, and to what extent, legal intervention might help protect individuals' privacy in this era.

---

86. See *supra* Section II.A.

## A. SECTORAL PRIVACY AND TECHNOLOGY

Technology and privacy protection go hand in hand in America. For example, since people began to communicate via mail and telegraph, concerns over insecurity, disclosure, and breach of confidentiality led many U.S. policymakers to strengthen mail security.<sup>87</sup> Another example is Warren and Brandeis's influential article "The Right to Privacy," which was allegedly inspired by the combination of relatively new technologies, such as the yellow press and cameras (instantaneous photography),<sup>88</sup> along with new business models of some industries that found profitability in publishing such data.<sup>89</sup> Likewise, the spread of telephone use raised various privacy concerns that eventually led to diverse legislative responses on both the state and the federal levels.<sup>90</sup>

Compared to other technological inventions to date, digitization has probably influenced privacy protection the most. In its early days, digitization created the need to develop standards of fair information practices in dealing with citizens' personal information.<sup>91</sup> When electronic communications presented privacy protection with new challenges,<sup>92</sup> Congress passed legislation relating to the interception and access of electronic communications

---

87. See Helm & Georgatos, *supra* note 32, at 141–42 (describing how a then-emerging technology (the mail) influenced privacy protection).

88. See Warren & Brandeis, *supra* note 8; Solove, *supra* note 17, at 1–11; Andreas Busch, *Privacy, Technology, and Regulation: Why One Size Is Unlikely to Fit All*, in *SOCIAL DIMENSIONS OF PRIVACY: INTERDISCIPLINARY PERSPECTIVES* 303, 304–05 (Beate Roessler & Dorota Mokrosinska eds., 2015).

89. Revealing photographs and gossip about individuals' personal lives was profitable mainly for the penny press. See Danielle Keats Citron, *Mainstreaming Privacy Torts*, 98 CALIF. L. REV. 1805, 1807 (2010); Warren & Brandeis, *supra* note 8, at 196.

90. See, e.g., The Communications Act of 1934, Pub. L. No. 73-416, 48 Stat. 1064 (1934); Omnibus Crime Control and Safe Streets Act of 1968, Pub. L. No. 90-351, 82 Stat. 211 (1968) (codified as amended at 18 U.S.C. §§ 2510–22 (2012)); Helm & Georgatos, *supra* note 32, at 142–43.

91. The first American acknowledgment of fair information practices standards was by the Department of Health and Human Services, which in 1973 elaborated a code of practice for the fair treatment of citizens' personal information. See U.S. DEP'T OF HEALTH, EDUC. & WELFARE, SECRETARY'S ADVISORY COMM. ON AUTOMATED PERSONAL DATA SYSTEMS, RECORDS, COMPUTERS AND THE RIGHTS OF CITIZENS (1973), reprinted in U.S. PRIVACY PROTECTIONS STUDY COMMISSION, PERSONAL PRIVACY IN AN INFORMATION SOCIETY 15 n.7 (1977); Reidenberg, *supra* note 15, at 879–80. For criticism on fair information practices in the United States, see Omer Tene, *Privacy Law's Midlife Crisis: A Critical Assessment of the Second Wave of Global Privacy Laws*, 74 OHIO ST. L.J. 1217, 1218–20 (2013); Joel R. Reidenberg, *Setting Standards for Fair Information Practice in the U.S. Private Sector*, 80 IOWA L. REV. 497, 499–500 (1995).

92. See Helm & Georgatos, *supra* note 32, at 143.

and computer tampering.<sup>93</sup> HIPAA, as a final example, was formed under a perceived need to digitize health information and better protect such information due to its potential sensitivity.<sup>94</sup>

As Section II.A implies, digital networks—especially the internet—led to the passage of several Acts designed to better protect individuals' privacy in specific contexts. For instance, COPPA protection for children's privacy online arose out of the potential risks to children's privacy when children surf the web.<sup>95</sup> Arguably, children's information needed as much protection before the internet era as it does now.<sup>96</sup> But before the internet, there were physical barriers to collecting information about children. Therefore, although the need to protect such information existed then, it was simply not on the agenda, either because it was somewhat tricky to violate children's privacy, or perhaps because any potential violation was at a socially tolerable level as conceived by policymakers. However, with the rise of the internet, the ease of conveying information to websites, especially those directed at children, has made the protection of children's information more relevant and crucial—so, COPPA was born.

Regarding the importance of digitization for information privacy, the internet has clearly influenced the development of sectoral privacy laws. But examining the vast amount of sensitive information extracted online,<sup>97</sup> one might conclude that Congress has done little to regulate it in this regard. Considering OSPs' capacity to harvest data online, sectoral privacy clearly has hardly imposed any obligations on OSPs online as it did in regulated sectors of the kinetic world. For instance, if the VPPA and the CCPA were crafted to protect consumers from revealing their preferences or habits regarding what they acquire because such information is sensitive, then online services like

---

93. See Electronic Communications Privacy Act of 1986, Pub. L. No. 100-618, 102 Stat. 3195 (1986) (codified as amended at 18 U.S.C. § 2710 (2012)); The Computer Fraud and Abuse Act, Pub. L. No. 99-474, 100 Stat. 1213–16 (1986) (codified at 18 U.S.C. § 1030 (2012)).

94. See *supra* Section II.A. Other instances are the collection and retention of sensitive consumer data like cable subscribers' records and video rental information as protected by the CCPA and VPPA, respectively. *Id.*

95. For further reading on the rationales behind COPPA, see generally Eldar Haber, *Toying with Privacy: Regulating the Internet of Toys*, 80 OHIO ST. L.J. 399 (2019).

96. It should, however, be further noted that various factors could have also affected the necessity to protect children's rights, and thus, the importance of protecting children might have changed throughout time. One example would be the international acknowledgment of granting such protection in 1989. See G.A. Res 44/25, Convention on the Rights of the Child (Nov. 20, 1989). For more on online risks to children, see Palfrey et al., *supra* note 84.

97. Almost everything end-users do on computerized networks is known to private parties. See, e.g., Daniel J. Solove, *Digital Dossiers and the Dissipation of Fourth Amendment Privacy*, 75 S. CAL. L. REV. 1083, 1084 (2002).

YouTube and Netflix must comply as well.<sup>98</sup> Moreover, scrutiny of internet usage, such as search query information, could reveal a great deal about individuals and thus represents a major threat to privacy under the sectoral approach, which does not generally apply to OSPs' harvesting personal data online.<sup>99</sup>

But the internet is simply the beginning in this context. Even the ability to collect sensitive information over the internet, no matter how massive it seems, might still be less powerful than that of impending technological innovations. As we move toward an era where ordinary devices, or "things," are becoming interconnected through the internet and are equipped with powerful sensors capable of capturing conversations, imagery, videos, geolocation, biometric data, and even vital signs, such as blood pressure or heart rate, information privacy is at great risk.<sup>100</sup>

Therefore, IoT must be further scrutinized for a grasp of its potential ramifications regarding information privacy. To gain a better understanding of these risks and potential solutions, the next Section will introduce the "always-on" era and analyze sectoral privacy within that context.

#### B. PRIVACY IN THE "ALWAYS-ON" ERA

Data collection and retention have been constantly on the rise since the invention of the internet.<sup>101</sup> It has enabled both private and public parties to collect massive amounts of data about their users.<sup>102</sup> The internet began to expand beyond traditional computers when other electronic devices, such as phones, TVs, watches, and even homes, suddenly became "smart." Using these smart devices quickly became the norm for many individuals in today's digital society.<sup>103</sup> Not long thereafter, other physical items—or simply

---

98. See Patricia L. Bellia, *Federalization in Information Privacy Law*, 118 YALE L.J. 868, 874–85 (2009).

99. See Paul Ohm, *The Rise and Fall of Invasive ISP Surveillance*, 2009 U. ILL. L. REV. 1417, 1417 (2009) (discussing the threats to privacy that arise from ordinary internet usage).

100. See Ohm, *supra* note 1, at 1143–44. It should be noted that geolocation information is generally not protected under the sectoral approach, rather, only under COPPA or regarding governmental access to information. See 16 C.F.R. § 312.2 (2018).

101. For more on the history of the public internet, see generally Jonathan Zittrain, *A History of Online Gatekeeping*, 19 HARV. J.L. & TECH. 253 (2006).

102. See Ben Popken, *Google Sells the Future, Powered by your Personal Data*, NBC NEWS (May 10, 2018), <https://www.nbcnews.com/tech/tech-news/google-sells-future-powered-your-personal-data-n870501> [<https://perma.cc/2SMD-NUVM>].

103. See, e.g., Larry Downes, *Why you may have good reason to worry about all those smart devices*, WASH. POST (Dec. 6, 2016), [https://www.washingtonpost.com/news/innovations/wp/2016/12/06/why-you-may-have-good-reason-to-worry-about-all-those-smart-devices/?noredirect=on&utm\\_term=.f6d8fcb1e7c5](https://www.washingtonpost.com/news/innovations/wp/2016/12/06/why-you-may-have-good-reason-to-worry-about-all-those-smart-devices/?noredirect=on&utm_term=.f6d8fcb1e7c5) [<https://perma.cc/8N7Z-UWG6>].

“things”—also emerged as interconnected. Not surprisingly, this technology is hence termed the Internet of Things, or IoT.

IoT undoubtedly increases the possibility of data gathering, in both the types of data and their potential volume, and it could signal a step up in a new generation of data mining.<sup>104</sup> These “things” are capable of gathering massive amounts of data about their users; for example, smart TVs, refrigerators, and even smart washing machines can collect, analyze, and retain data on their users’ habits.<sup>105</sup> Smart TVs can also listen to, record, and send to a third party whatever their microphones catch<sup>106</sup> and can even acquire data from a built-in camera.<sup>107</sup> Smartphones in particular enable information gathering of various types by various service providers.<sup>108</sup>

In the development of IoT, an emerging generation of technology could further elevate data collection: devices that operate in an always-on mode, meaning they can constantly collect data even without being active.<sup>109</sup> The definition of always-on devices is largely self-explanatory: such devices either

104. See Rostow, *supra* note 40, at 686.

105. See Chris Hoffman, *How to Stop Your Smart TV From Spying on You*, HOW-TO GEEK (Nov. 16, 2015), <https://www.howtogeek.com/233742/how-to-stop-your-smart-tv-from-spying-on-you> [<https://perma.cc/KD2N-5LWA>]. Some Smart TVs could also transmit the names of files on USB drives connected to the television and capture data from networks to which they are attached. See Joseph Steinberg, *These Devices May Be Spying on You (Even in Your Own Home)*, FORBES (Jan. 27, 2014), <http://www.forbes.com/sites/josephsteinberg/2014/01/27/these-devices-may-be-spying-on-you-even-in-your-own-home/#15407ce56376> [<https://perma.cc/Q5WK-9RN3>].

106. See April Glaser, *Philip K. Dick Warned Us About the Internet of Things in 1969*, SLATE (Feb. 10, 2015), [http://www.slate.com/blogs/future\\_tense/2015/02/10/philip\\_k\\_dick\\_s\\_1969\\_novel\\_ubik\\_on\\_the\\_internet\\_of\\_things.html](http://www.slate.com/blogs/future_tense/2015/02/10/philip_k_dick_s_1969_novel_ubik_on_the_internet_of_things.html) [<https://perma.cc/A6DR-FE98>].

107. See Steinberg, *supra* note 105.

108. For instance, cellular providers could track information about their users, such as with whom the user communicates and where the user goes; manufacturers and providers of software for smartphones, such as Google (Android phones) and Apple (iPhones), could track the actions their users are taking on their phone; and app developers often use their installed apps to extract information from the phone’s contact list, microphone, and camera. See *id.* In fact, many flashlight apps gained a reputation of data exfiltration. See Robert McMillan, *The Hidden Privacy Threat of...Flashlight Apps?*, WIRED (Oct. 20, 2014), <https://www.wired.com/2014/10/iphone-apps> [<https://perma.cc/J8PS-AC94>].

109. There is, however, a difference between always-ready and always-on statuses. Always-ready devices usually process locally to detect a “wake phrase,” which triggers the device to begin transmitting data. But always-on devices transmit data all the time while the processing occurs only externally. For the purposes of this Article, always-ready devices count as always-on, since always-ready devices are constantly awaiting the trigger phrase, they must always be “on” and thus could potentially transfer and collect data constantly. For more on this categorization, see *Microphones & the Internet of Things*, FUTURE PRIVACY F. (Aug. 2017), <https://fpf.org/wp-content/uploads/2017/08/Microphones-Infographic-Final.pdf> [<https://perma.cc/EUM2-D6QG>].

always await a trigger phrase to begin operating at any moment (“always-ready”) or operate constantly without a moment of idleness (“always-on”).<sup>110</sup>

Examples of always-on devices can be found in many areas, from young kids using smart, connected toys and devices to individuals (young and old) using computerized personal assistants or operating a smart home.<sup>111</sup> Their functionality and operation can be exemplified through computerized personal assistants like Amazon Echo.<sup>112</sup> For example, Amazon Echo is an “always-ready” device, meaning that it only becomes active upon a voice command like “Alexa” or “Amazon,” depending on users’ preferences. But for the device to know when the user has operated the activation command, it must constantly await commands by “listening” to its users. Therefore, the device is labeled as always on, even if it presumably deactivates without the command.<sup>113</sup>

These innovative technologies mark the beginning of the always-on era. Due to the devices’ mode of operation and their data collection abilities, the

---

110. For many IoT devices, users simply need to say the voice command to activate them, which Stacey Gray suggested terming as microphone-enabled devices. *See* Stacey Gray, *Always On: Privacy Implications of Microphone-Enabled Devices*, FUTURE PRIVACY F. 3 (Apr. 2016), [https://fpf.org/wp-content/uploads/2016/04/FPF\\_Always\\_On\\_WP.pdf](https://fpf.org/wp-content/uploads/2016/04/FPF_Always_On_WP.pdf) [<https://perma.cc/QS4Q-KJ8C>].

111. Computerized personal assistants are software agents that can perform tasks or services for an individual, usually based on user input, location awareness, and the ability to access information from a variety of online sources. There are various types of computerized personal assistants (e.g., Apple’s Siri and Microsoft’s Cortana). In 2014, Google even embedded such technology under a pre-installed ability in Google’s Chrome browser, which passively listened for the phrase “OK, Google” to launch a voice-activated search function. *See* Tony Bradley, ‘OK Google’ Feature Removed from Chrome Browser, FORBES (Oct. 17, 2015), <http://www.forbes.com/sites/tonybradley/2015/10/17/ok-googlefeature-removed-from-chrome-browser/#16d299a44e27> [<https://perma.cc/8SL7-XFFM>]; *see also* Top 22 Intelligent Personal Assistants or Automated Personal Assistants, PREDICTIVE ANALYTICS TODAY, <https://www.predictiveanalyticstoday.com/top-intelligent-personal-assistants-automated-personal-assistants/> [<https://perma.cc/K2W4-RGPQ>] (last visited Jan. 9, 2020).

112. Amazon Echo is “a hands-free speaker you control with your voice.” *Amazon Echo*, AMAZON, <https://www.amazon.com/Amazon-Echo-Bluetooth-Speaker-with-WiFi-Alexa/dp/B00X4WHP5E> [<https://perma.cc/9LUQ-FRT3>] (last visited Jan. 9, 2020). It “connects to the Alexa Voice Service to play music, make calls, send and receive messages, provide information, news, sports scores, weather, and more—instantly. . . . When you want to use Echo, just say the wake word ‘Alexa’ and Echo responds instantly.” *Id.*

113. Notably, it is difficult to estimate if these devices constantly collect data. Amazon, for instance, claimed that their Echo device only starts recording upon the trigger phrase. Google argues that Google Home (as another example of a computerized personal assistant) only “listens in short (a few seconds) snippets for the hotword. Those snippets are deleted if the hotword is not detected, and none of that information leaves your device until the hotword is heard.” *See* Scott Carey, *Does Amazon Alexa or Google Home Listen to My Conversations?*, TECHWORLD (May 25, 2018), <https://www.techworld.com/security/does-amazon-alexa-listen-to-my-conversations-3661967> [<https://perma.cc/Y4W9-HPHX>].

always-on era could lead to the collection and storage of massive quantities of user data of various types. Almost anything could be transmitted and stored,<sup>114</sup> depending mostly on the plausibility of obtaining authorized or unauthorized access to data. It might lead to a constant—and almost endless—collection and retention of data, even when the device seems to have been deactivated.

In the context of sectoral privacy, always-on devices may very well collect and retain information deemed sensitive by Congress and hence should become a subject for sectoral protection under federal laws. Operators of always-on devices might easily cover all the categories of sectoral privacy.<sup>115</sup> Consider, for example, Amazon Echo. If one is present in your household, it can capture any conversation in its vicinity and thus collect data regarding your finance, health, school performance, and any other information that you might consider sensitive. If children are present, it might capture their voices, conversations, questions, and even their musical preferences if they ask the device to play songs. Also, consider an always-on smart TV, or another IoT device equipped with a camera. Besides potentially acquiring watching habits, it could be used or misused to collect sound and imagery from its surroundings—Echo Show and Echo Look are just two examples of devices with a microphone and a camera.<sup>116</sup> Thus, many smart devices can gather almost any information that is already deemed sensitive by Congress. Worse yet, collecting some sensitive data is not just possible; it is highly probable.

To date, this so-called always-on era has had little influence on reforming sectoral privacy, nor modifying it even slightly.<sup>117</sup> When it comes to IoT, sectoral privacy borders on irrelevance. Other than COPPA, which marginally applies to some IoT devices—namely connected smart toys, such as Hello Barbie and My Friend Cayla<sup>118</sup>—most sectoral privacy laws do not apply to

---

114. See, e.g., Nick Ismail, *Storage Predictions: Will the Explosion of Data in 2017 be Repeated in 2018?*, INFORMATION-AGE (Dec. 6, 2017), <http://www.information-age.com/explosion-data-2017-repeated-2018-123469890> [<https://perma.cc/9YGX-G4YZ>].

115. As previously mentioned, sectoral privacy mainly protects financial privacy, educational privacy, health privacy, children's privacy, and consumer data privacy. See *supra* Section II.A.

116. See *Echo Look*, AMAZON, <https://www.amazon.com/Amazon-Echo-Look-Camera-Style-Assistant/dp/B0186JAEWK> [<https://perma.cc/6ESZ-AY7H>] (last visited Jan. 9, 2020); *Echo Show*, AMAZON, [https://www.amazon.com/Amazon-Echo-Show-Alexa-Enabled-Black/dp/B01J24C0TI/ref=sr\\_1\\_1?s=amazon-devices&ie=UTF8&qid=1528381083&sr=1-1&keywords=echo+show&dpID=51syqGPcCmL&preST=\\_SY300\\_QL70\\_&dpSrc=srch](https://www.amazon.com/Amazon-Echo-Show-Alexa-Enabled-Black/dp/B01J24C0TI/ref=sr_1_1?s=amazon-devices&ie=UTF8&qid=1528381083&sr=1-1&keywords=echo+show&dpID=51syqGPcCmL&preST=_SY300_QL70_&dpSrc=srch) [<https://perma.cc/49UY-DJAR>] (last visited Jan. 9, 2020).

117. While IoT has been debated in Congress, no significant legislation has passed thus far to protect information privacy.

118. Hello Barbie and My Friend Cayla are examples of connected smart toys, or IoT Toys, which are connected to the internet and can communicate with their users through voice

IoT because they have been crafted very narrowly to address a specific problem. Financial privacy, for instance, will be protected only if the IoT operator is an institution engaging in financial activities, or a certain entity that receives non-public personal information from non-affiliated financial institutions.<sup>119</sup> FCRA will probably not apply unless IoT operators are treated as consumer-reporting agencies.<sup>120</sup> Educational privacy will be generally excluded, as it applies to educational institutes or agencies that receive federal funds from the DoE.<sup>121</sup> Health privacy and consumer data privacy will likewise not be easily protected by these Acts when it comes to IoT devices, as they will not be considered covered entities under federal regulations.<sup>122</sup> All in all, sectoral privacy will not greatly concern IoT.

At root, policymakers must realize that many IoT devices are likely to acquire sensitive data, so privacy protection of relevant data should not be restricted merely to covered entities. Even the context of specific industries (e.g., educational facilities) or a specific population (e.g., young children) is inadequate to protect sensitive information today. When almost everything around us becomes a computer that is connected to the internet,<sup>123</sup> and data become a substantive part of the business model for many companies, the notion of how better to protect users' privacy has to be reconsidered. With few exceptions, most of the core values protected by the sectoral approach could become meaningless with the advance of new technologies, especially IoT. Therefore, protecting sectoral privacy in the always-on era calls for some form of intervention, legal or technological.

### C. ALWAYS-ON REGULATIONS

That privacy had met its demise became a popular opinion toward the end of the twentieth century.<sup>124</sup> Some regulators, however, like those of the

---

commands. For more on the regulation of IoT in the United States, see generally Haber, *supra* note 95.

119. See 12 U.S.C. §§ 3401–3422.

120. See 15 U.S.C. § 1681(b) (2012).

121. See 20 U.S.C. § 1232(g).

122. See 45 C.F.R. §§ 160.102–103.

123. See generally BRUCE SCHNEIER, *CLICK HERE TO KILL EVERYBODY: SECURITY AND SURVIVAL IN A HYPER-CONNECTED WORLD* 5–12 (2018) (describing how IoT turns almost any item into a computer).

124. Many argued that privacy is dead or that it deserves at most minimal protection in the digital age. Others argued that privacy should be treated as a tradeable currency. Scott McNealy, chief executive officer of Sun Microsystems, is quoted as saying, “You have zero privacy anyway . . . . Get over it.” Polly Sprenger, *Sun on Privacy: ‘Get Over It’*, WIRED (Jan. 26, 1999), <https://www.wired.com/1999/01/sun-on-privacy-get-over-it> [<https://perma.cc/FLA2-7EW5>]; see also A. Michael Froomkin, *The Death of Privacy?*, 52 STAN. L. REV. 1461, 1462 (2000). For more on the currency argument, see James P. Nehf, *Shopping for Privacy Online*:



European Union, recently made a clear statement regarding privacy under its General Data Protection Regulation (GDPR): protecting privacy still matters, perhaps even more today than ever before.<sup>125</sup> Therefore, private companies should grant their users more effective means of control and protection. While American policymakers still do not protect information privacy as robustly as the European Union does, and perhaps granting such protection might not be achieved easily, they evidently do not disregard this right and instead still seek proper ways to better protect it under their regulatory approach.<sup>126</sup>

Notably, privacy protection does not necessarily require legal intervention. As Lawrence Lessig famously argued, other potential modalities, like the market, social norms, and architecture, could also regulate behavior, with or without the law.<sup>127</sup> The problem with some of these potential modalities in the privacy-protection field and the always-on era lies in their failure to optimally regulate privacy protection on their own. For instance, as history shows, the market as a modality—while arguably an important component of any solution—might be insufficient to regulate privacy due to existing market failures.<sup>128</sup> Likewise, social norms will not effortlessly change the data-mining practices of commercial entities.<sup>129</sup> While this Article considers the potential of both the market and social norms to regulate privacy, it focuses mainly on the modalities of law and technology, probing mainly the law in this Section.

The use of the law as a modality to better protect privacy, by embedding the values protected by the sectoral approach, can take many forms. One might

---

*Consumer Decision Making Strategies and the Emerging Market for Information Privacy*, 2005 U. ILL. J.L. TECH. & POL'Y 1, 14–17 (2005).

125. See generally Commission Regulation 2016/679, 2016 O.J. (L 119) (EU) 1 (repealing Directive 95/46/EC) (General Data Protection Regulation).

126. See, e.g., FED. TRADE COMM'N, INTERNET OF THINGS: PRIVACY AND SECURITY IN A CONNECTED WORLD (Jan. 2015), <https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf> [<https://perma.cc/62U9-6Y6J>].

127. See LAWRENCE LESSIG, CODE: VERSION 2.0 120–37 (2006); LAWRENCE LESSIG, FREE CULTURE 116–73 (2004) (suggesting four modalities that regulate behavior).

128. To exemplify, many OSPs rely on data as a business model; this can be a market failure if they are a monopoly or operate in an oligopolistic market and thus lack incentives to provide proper privacy protections. Users will generally lack the opportunity to indicate their discontent with such practices. For more on privacy and market failures, see Victoria L. Schwartz, *Corporate Privacy Failures Start at the Top*, 56 B.C. L. REV. 1693 (2016).

129. To name a few examples, social norms generally fail to solve this conundrum, as consumers are generally unaware of data-mining practices; some view privacy as a currency; many fail to understand the implications of data storage; and even those who use these services might not be able to use them in the IoT context. For a general discussion on privacy and social norms, see Randall P. Bezanson, *Privacy, Personality, and Social Norms*, 41 CASE W. RES. L. REV. 681 (1991).

argue that to protect privacy properly in the digital age, Congress should choose a different framework (i.e., abandon the sectoral approach entirely), as this does little to advance the rationales behind the current regulatory approach to privacy protection.

Many scholars have long warned how poorly suited the sectoral approach is to protecting privacy in this era.<sup>130</sup> Indeed, it is difficult to grasp why personal information should be treated differently simply because of the identity of the private party that holds it, as the sectoral approach suggests. But it did seem to make sense that your doctor, rather than most individuals you encounter, should obtain your personal information or have full access to your entire medical history. The present nature of data mining could challenge these assumptions. Google most likely has far more intimate and personal information about you than your doctor. It might even know more about you than anyone else in the world, including your family and perhaps even yourself. That is why the sectoral privacy approach might be inadequate and ill-suited to protecting privacy in today's always-on era.

These changes might eventually lead to a comprehensive federal privacy law that could provide a one-size-fits-all approach or create a federal baseline for all industries when dealing with sensitive information.<sup>131</sup> While this remains to be seen, what should be evident in the always-on era is that regulating privacy in industries does little to achieve the goals of the sectoral approach. When the devices that surround us can collect protected forms of information very similar to those of the regulated industries, any sectoral regulation must also apply to OSPs of IoT. The current patchwork regime to protect privacy is thus too outdated to deal with current challenges. Therefore, policymakers are duty-bound to reevaluate the collection, storage, and transfer of information across the private sector, and to regulate it accordingly.

Developing a one-size-fits-all approach does not necessarily mean abandoning the sectoral approach entirely, as other forms of offline data collection might still exist. An extreme *ex ante* approach, for instance, might argue that the solution for protecting information privacy would be simply to ban data collection and retention in general, at least for some companies or sectors. This general approach, which focuses on how to prevent some forms of data from being retained from the outset, is unsuitable. This is because data serve as a business model for many companies. It is a multibillion-dollar

---

130. See, e.g., Ohm, *supra* note 61, at 1762; cf. Schwartz, *supra* note 74, at 922–31 (discussing the drawbacks of embracing an omnibus privacy regime in the United States).

131. See, e.g., Bellia, *supra* note 98, at 890–900 (advocating for the importance of federalization of information privacy law).

industry with many benefits for its users, such as the offer of free services.<sup>132</sup> Data could be highly valuable for companies, and for some users, as data processing could enable, inter alia, targeted—perhaps more accurate—advertising and suggest personalized services.<sup>133</sup>

Furthermore, the practice of data collection and retention serves many functions and values. It advances knowledge and innovation and is crucial for the development of machine learning, deep learning, and big data analysis, to name but a few examples, all of which rely heavily on large quantities of training data.<sup>134</sup> In addition, various parties, such as credit card companies, use data to reduce exposure to risks and costs of doing business, while they increase companies' effectiveness at raising revenues.<sup>135</sup> In sum, a general approach banning information or simply ignoring the opportunities in data mining altogether is neither practical nor desirable.<sup>136</sup>

But gray areas in such an ex ante approach exist. Depending on various factors regarding information privacy, companies might be allowed to process some forms of data, but not others. They could also be obliged or incentivized to incorporate privacy-enhancing principles into their practices, which could include, inter alia, limits on data collection and retention, data disposal, data accuracy, and various cybersecurity measures—at least for protection against the unauthorized use of these data.<sup>137</sup> In addition, technological developments could aid companies in understanding which data should be retained and which should not. Imagine that your smart device could actually deduce the speaker's identity, and therefore could change privacy settings in keeping with its preferences, or even more closely, protect the privacy of those that Congress sought to protect. So if your smart assistant could differentiate you from your under-thirteen-year-old child, it could potentially retain information

---

132. See Rostow, *supra* note 40, at 687; Cuaresma, *supra* note 40, at 506.

133. Notably, data could be highly valuable for non-profit companies as well, as they might, inter alia, use free open-sourced technology variants of for-profit company technologies in order to service users' device. Some users, however, might view targeted advertising and personalized services as a nuisance.

134. This is especially evident in the context of deep learning, which yields state-of-the-art results in fields such as speech, image, and text recognition, but based on billions of records that were collected from private users. For more on deep learning, see generally Liangpei Zhang et al., *Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art*, 4 IEEE GEOSCIENCE & REMOTE SENSING MAG. 22 (2016); Xue-Wen Chen & Xiaotong Lin, *Big Data Deep Learning: Challenges and Perspectives*, 2 IEEE ACCESS 514 (2014).

135. See Cuaresma, *supra* note 40, at 506.

136. See Jane Yakowitz, *Tragedy of the Data Commons*, 25 HARV. J.L. & TECH. 1, 4–5 (2011).

137. For similar recommendations in the United States, see, for example, *Protecting Consumer Privacy in an Era of Rapid Change*, *supra* note 24, at 15–71.

from you alone and exclude data mining from the child. Developments in voice or facial recognition could advance this rationale.<sup>138</sup>

The problem, however, is that other than children's privacy, protecting sensitive data goes far beyond the speaker's identity. It requires context and evaluation of the data to ascertain whether it is sensitive. It would be very difficult to achieve such a goal *ex ante*—requiring measures that could accurately predict which information it should not store beforehand.

Even with the invention of such technological measures, or if at least we accept *ex ante* privacy protection depending on the user's identity, these measures might raise further privacy issues. Embedding technologies like facial or voice recognition in IoT devices might exert a significant negative effect on the right to privacy, as these means rely on biometric features. Even if these features were stored only internally on the device, and even assuming that the biometric data were encrypted against the potential abuse of them, these measures would essentially rely on a form of anonymization, which could easily be de-anonymized, and in consequence users could be re-identified and suffer further damage.<sup>139</sup> For instance, if an Amazon Echo “knows” how to locate persons in a household and retains their preferences, revelation of the household's preferences could easily identify what data were linked to each person. In other words, using this method of protecting one cohort, such as children, might eventually lead to diminished protections for others.

A less extreme *ex ante* approach could rely on users' preferences (i.e., depending on whether they consent to such potential threats to their privacy in the always-on era). This so-called notice-and-consent mechanism already exists in the United States as part of the Fair Information Practice Principles (FIPPs), which are generally designed to empower data subjects by ensuring that they have sufficient knowledge of a data collector's activities in order to choose to consent to them or not.<sup>140</sup> However, this regulation-by-information approach, which relies on the concept of “informed consent,” has proven

---

138. Facial recognition is also developing at a rapid pace. Both Google Home and Amazon Echo have gained the ability to recognize individual voices by creating a voice profile. See Chris Welch, *Amazon's Alexa Can Now Recognize Different Voices and Give Personalized Responses*, VERGE (Oct. 11, 2017), <https://www.theverge.com/circuitbreaker/2017/10/11/16460120/amazon-echo-multi-user-voice-new-feature> [https://perma.cc/SRN9-JFBX].

139. Paul Ohm categorizes this phenomenon as the “accretion problem,” where “[o]nce an adversary has linked two anonymized databases together, he can add the newly linked data to his collection of outside information and use it to help unlock other anonymized databases.” Ohm, *supra* note 61, at 1746–48.

140. FIPPs could include, *inter alia*, notice, choice, access, accuracy, data minimization, security, and accountability. See Woodrow Hartzog, *The Inadequate, Invaluable Fair Information Practices*, 76 MD. L. REV. 952, 973–75 (2017).

ineffective in protecting privacy,<sup>141</sup> and likewise from a legal standpoint, individuals are unlikely to challenge potential violation of their privacy for a variety of reasons.<sup>142</sup>

Having determined that an ex ante legal approach is ineffective and insufficient to protect privacy, let us move on to examine ex post approaches. If we allow private companies to maintain their data-mining practices, we could limit them—or shape their practices according to core protected values—by imposing ex post liability. This could be civil, administrative, or even criminal. It could be promoted by fines, for instance, akin to what the GDPR imposes,<sup>143</sup> or it could be reputational and monetary like data breach notifications, which are currently legislated by states.<sup>144</sup> But these forms of regulation, which rest somewhat on deterrence theory, might also prove ineffective for the intended goals.<sup>145</sup>

---

141. As history shows from terms of service agreements, end-user license agreements (EULAs), and privacy policies, most consumers do not bother reading them for two main reasons: these documents are usually long and written in a legal language almost incomprehensible to most people, and consumers today already experience information flooding. See, e.g., Daniel B. Ravicher, *Facilitating Collaborative Software Development: The Enforceability of Mass-Market Public Software Licenses*, 5 VA. J.L. & TECH. 11, 13 (2000); Garry L. Founds, *Shrinkwrap and Clickwrap Agreements: 2B or Not 2B?*, 52 FED. COMM. L.J. 99, 100 (1999); George R. Milne & Mary J. Culnan, *Strategies for Reducing Online Privacy Risks: Why Consumers Read (or Don't Read) Online Privacy Notices*, 18 J. INTERACTIVE MARKETING 15, 20–21 (2004); Joel R. Reidenberg et al., *Privacy Harms and the Effectiveness of the Notice and Choice Framework*, 11 I/S: J.L. & POL'Y FOR INFO. SOC'Y 485, 491 (2015).

142. To name a few reasons, it is usually difficult for individuals to know when their rights were violated, to prove these violations, and to satisfy the injury-in-fact standing requirement under Article III of the Constitution without concrete harm. See U.S. CONST. art. III; *Lujan v. Defs. of Wildlife*, 504 U.S. 555 (1992); *Spokeo, Inc. v. Robins*, 136 S. Ct. 1540, 1545 (2016) (holding that a plaintiff does not satisfy Article III standing without identifying a concrete harm). It should be noted, however, that some courts ruled that violation of some Acts could constitute injury in fact sufficient to satisfy standing. See, e.g., *Matera v. Google Inc.*, No. 15-CV-04062-LHK, 2016 WL 5339806, at \*14 (N.D. Cal. Sept. 23, 2016) (holding that violations of the Wiretap Act and state law constitute injury in fact); cf. *Hancock v. Urban Outfitters*, 830 F.3d 511, 514 (D.C. Cir. 2016) (holding that injury in fact depends also on the type of information which would be sufficient for standing).

143. See Commission Regulation 2016/679, 2016 O.J. (L 119) (EU) 1 (repealing Directive 95/46/EC) (General Data Protection Regulation).

144. Data breach notifications usually require some private and government entities to notify individuals of security breaches of information involving personally identifiable information. See David Thaw, *The Efficacy of Cybersecurity Regulation*, 30 GA. ST. U. L. REV. 287, 297 (2014).

145. Deterrence theory had been criticized over the years by many scholars. See, e.g., Dan M. Kahan, *The Theory of Value Dilemma: A Critique of the Economic Analysis of Criminal Law*, 1 OHIO ST. J. CRIM. L. 643, 643–47 (2004).

An ex post approach could also take many other forms. Upon communication, companies could try to assess whether a data chunk should be deemed sensitive and retain only pieces of data deemed non-sensitive. In other words, policymakers could impose obligations on private companies to monitor their communications, analyze them, and opt for such a solution if the data are sensitive. Yet, this approach might also be problematic. For instance, it would be difficult to implement in practice because it would require an in-depth, sometimes subjective, analysis of data to determine its sensitivity. But who will decide whether a piece of data is sensitive or not? Should the state delegate this power to quasi-judicial or private entities?<sup>146</sup>

This measure would most likely be taken by computerized systems, as it would be impractical—and not necessarily desirable—to assign individuals to make these decisions, considering that it would be impossible for a human being to perform this task with the necessary accuracy and efficiency.<sup>147</sup> In addition, many companies might not have the capacity to conduct such analyses. Thus imposing strict obligations to review the sensitivity of data on companies might raise the barrier to entry in a market. Here, perhaps, it is even preferable that companies adhere to an “ignorance is bliss” approach, since obliging companies to obtain actual knowledge of the information that is conveyed and stored might defeat the very purpose of safeguarding users’ privacy.

Overall, perhaps technology will eventually make the sectoral approach obsolete. Much like technology-sparked discussions on information privacy for a specific cohort (i.e., children) or a specific context (i.e., video rental), IoT might challenge the current perception of what data should be protected, and Congress might eventually add more protections to other types of cohorts or contexts. This might make sense, as the technological changes of the always-on era could be perceived as much more comprehensive in the sense of privacy than those that Congress has regulated over time. But perhaps technology should be viewed as not only the problem, but also the solution for protecting privacy in the always-on era.

---

146. The practice of delegating quasi-judicial powers to intermediaries is, however, not unheard of. The Digital Millennium Copyright Act (DMCA), for example, created a notice-and-takedown regime against copyright infringement and de facto required search engines to “receive requests from copyright owners or their representatives to remove search results that link to allegedly infringing materials.” Eldar Haber, *Privatization of the Judiciary*, 40 SEATTLE U. L. REV. 115 (2016). Another example is the so-called right to be forgotten (or right to erasure) in the European Union, which also obliges some intermediaries to delist or even delete data that relate to the right of information privacy under some circumstances. For an overview and criticism of these and other privatization practices, see *id.*

147. *Id.* at 144 (discussing the costs of content reviewers).

The next Part shows how various technological measures could be embedded within IoT devices or services to better protect the values that the sectoral approach sought to protect, offering a toolkit for policymakers and OSPs to embrace a new approach, with or without legal intervention.

#### IV. REGULATING THE ALWAYS-ON ERA THROUGH TECHNOLOGY

In the always-on era, sensitive data are increasingly collected by unregulated entities under federal laws. Technology in the context of privacy protection, however, is not simply a problem, but may also be a viable solution. Accordingly, this Part discusses potential technological solutions to properly balance data utility with privacy interests and proposes the use of what will be defined as coresets for differential privacy and homomorphic encryption to be embedded in the operation of IoT devices. This Article proposes to use differential privacy *ex ante* based on the probability of sensitivity, as illustrated in the final Section.

##### A. TECHNOLOGY AS A SOLUTION

Using technology could enhance privacy protection for users even without abandoning the sectoral approach. In other words, technology might substantially help protect the very same values that it might help infringe.<sup>148</sup> But before discussing specific technological solutions to enhance privacy protection in the always-on era, it is essential to first acknowledge that such protection might depend greatly on the ways they are implemented. As a general framework, this Article advocates the use of an approach termed Privacy by Design (PbD): a “systematic approach to designing any technology that embeds privacy into the underlying specification or architecture.”<sup>149</sup> As a concept, PbD could be implemented to help manage various privacy challenges.<sup>150</sup> For example, this concept could be interpreted as calling for structural support for privacy protection and advocating privacy protection by an organization’s default mode of operation.<sup>151</sup> As explained later, PbD could be embedded in any technological solution.

---

148. See, e.g., Urs Gasser, *Recoding Privacy Law: Reflections on the Future Relationship Among Law, Technology, and Privacy*, 130 HARV. L. REV. F. 61 (2016).

149. See Ira S. Rubinstein, *Regulating Privacy by Design*, 26 BERKELEY TECH. L.J. 1409, 1411–12 (2011).

150. See Gasser, *supra* note 148, at 65–66.

151. See Ann Cavoukian, *Privacy by Design: The 7 Foundational Principles*, IPC (Jan. 2011), <https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf> [<https://perma.cc/Z94X-NTJK>].

This will not be the first instance where technology is suggested—and used—as a solution to protect privacy. This trend began in the 1970s with the development of Privacy-Enhancing Technologies (PETs), designed to be responsive to new information and communication technologies.<sup>152</sup> Many of these technological measures had been suggested or even implemented to protect privacy in the past, but with only modest success.<sup>153</sup> However, PETs could assume many forms, as this Section will show. Generally, privacy protection will adhere to various methods of de-identification of personal data. The ultimate goal will be to preserve the value of data and provide safeguards against identifying users at the same time. These methods could include, *inter alia*, anonymization and encryption. Here, “identifying users” means that one would not be able to learn from the resulting output (e.g., noisy database or statistics) regarding an individual record of the original database. The formal definition or lack of definition of privacy or the underlying assumptions/model is a crucial issue that is discussed in the next paragraphs.

We begin with one of the most common techniques in practice, one somewhat infamous in cryptography: data anonymization. Under this method, information in databases could be manipulated to make it intuitively difficult to identify data subjects.<sup>154</sup> Anonymization could be achieved by a variety of techniques (e.g., suppression of data,<sup>155</sup> generalizing identifiers,<sup>156</sup> or providing only aggregate statistics).<sup>157</sup> For example, Congress effectively chose anonymization to regulate healthcare data under HIPAA by specifying eighteen data identifiers whose removal from a dataset would, allegedly at least, protect privacy.<sup>158</sup>

---

152. See Gasser, *supra* note 148, at 65.

153. Such technological measures include adblockers, cryptography, and virtual private networks, to name a few. Another method focuses on the output of a query to a given database, meaning that the input itself is not in play, but rather through computation—the output of a query could aid in affording privacy protection. While these measures could play an important role in protecting privacy, they are generally insufficient to grant proper protection to all consumers in the IoT age. See Rostow, *supra* note 40, at 694–95; Kobbi Nissim et al., *Bridging the Gap Between Computer Science and Legal Approaches to Privacy*, 31 HARV. J.L. TECH 689, 695–96, 702 (2018).

154. See Ohm, *supra* note 61, at 1707–08.

155. Suppression means removing all identifying features from a dataset. *Id.* at 1707.

156. One technique would be suppressing or replacing users’ IDs that appear in each record. For example, during the 1990’s, America on-line (AOL) collected internet search queries of its users and published them for the research community. To preserve privacy, each user’s name was replaced by a random ID number. See Karim Z. Oussayef, *Selective Privacy: Facilitating Market-based Solutions to Data Breaches by Standardizing Internet Privacy Policies*, 14 B.U. J. SCI. & TECH. L. 104–05 (2008).

157. See Ohm, *supra* note 61, at 1714–16.

158. See 45 C.F.R. § 164.514(b)(2)(i)–(ii) (2018).



Data anonymization sounds like an almost-perfect solution to protect privacy, but it is not enough. While anonymization and aggregated data could help privacy protection,<sup>159</sup> it does not ensure privacy or protect it sufficiently.<sup>160</sup> Using reidentification or deanonymization methods, an adversary can link anonymized records to auxiliary information and discover the identity of data subjects.<sup>161</sup> This has proven possible by researchers in many instances.<sup>162</sup>

A famous example is Netflix, which publicly released one hundred million anonymized records that revealed how users rated movies. They did so by allowing teams to compete to improve their recommendation algorithm to win the “Netflix Prize.”<sup>163</sup> But it was not long before researchers proved how easy it was for an adversary to reidentify many data subjects using merely a smattering of outside knowledge about the subjects’ movie-watching preferences. Researchers combined Netflix’s published records of movie reviews with other public data, such as IMDB recommendations, that partially matched those records, thereby potentially revealing sensitive data about those

---

159. See Ira S. Rubinstein et al., *Data Mining and Internet Profiling: Emerging Regulatory and Technological Approaches*, 75 U. CHI. L. REV. 261, 266, 268 (2008).

160. See generally Ohm, *supra* note 61, Andreas Haeberlen et al., *Differential Privacy Under Fire*, PROC. 20TH USENIX SECURITY SYMP. 1 (Aug. 12, 2011), <http://www.cis.upenn.edu/~ahae/papers/fuzz-sec2011.pdf> [<https://perma.cc/EFD4-CGKS>]; Andrew Chin & Anne Klinefelter, *Differential Privacy as a Response to the Reidentification Threat: The Facebook Advertiser Case Study*, 90 N.C. L. REV. 1417, 1417–28 (2012).

161. See Ohm, *supra* note 61, at 1707–08; Michael Barbaro & Tom Zeller, *A Face is Exposed for AOL Searcher No. 4417749*, N.Y. TIMES (Aug. 9, 2006), <http://www.nytimes.com/2006/08/09/technology/09aol.html> [<https://perma.cc/7MW6-GUGR>].

162. See Ashwin Machanavajjhala et al., *L-diversity: Privacy Beyond K-anonymity*, 22ND INT’L CONF. ON DATA ENGINEERING IEEE (2006); Josep Domingo-Ferrer & Vicenç Torra, *A Critique of K-Anonymity and Some of Its Enhancements*, THIRD INT’L CONF. ON AVAILABILITY, RELIABILITY & SECURITY IEEE (2008). For further examples, see Latanya Sweeney, *K-Anonymity: A Model for Protecting Privacy*, 10 INT’L J. UNCERTAINTY, FUZZINESS & KNOWLEDGE-BASED SYSS. 557 (2002); Latanya Sweeney, *Weaving Technology and Policy Together to Maintain Confidentiality*, 25 J.L. MED. & ETHICS 98 (1997); Latanya Sweeney, *Simple Demographics Often Identify People Uniquely*, DATA PRIVACY LAB TECHNICAL REP. (2000); Cynthia Dwork, *Differential Privacy*, AUTOMATA, LANGUAGES & PROGRAMMING, 33RD INT’L COLLOQUIUM PROC. PART II 1 (2006); Yakowitz, *supra* note 136, at 3.

163. See *Netflix Prize*, NETFLIX, <https://www.netflixprize.com> [<https://perma.cc/9JAY-THEK>] (last visited Jan. 9, 2020) (“The Netflix Prize sought to substantially improve the accuracy of predictions about how much someone is going to enjoy a movie based on their movie preferences.”); see also James Bennett & Stan Lanning, *The Netflix Prize*, 2007 PROC. KDD CUP & WORKSHOP (2007).

Netflix users.<sup>164</sup> Eventually Netflix settled a class-action lawsuit regarding these potential privacy violations.<sup>165</sup>

Furthermore, de-identification methods like anonymization are notably advancing. One such common approach is a privacy model called k-anonymity.<sup>166</sup> It defines models that provide desiderata with provable guarantees but only under certain threat models (ways that the adversary may attack).

An example privacy mechanism that is used to satisfy k-anonymity is to remove features from each record so that there will be at least k duplications of each record in the data set. The idea is that, as a result, someone possessing the data set will not be able to distinguish among the records in such a cluster yet can still extract data utility from statistics. The main disadvantage of this approach is that we can learn about the user by learning from other records in a user's cluster. For example, each individual in the dataset can easily recognize her own record and thus her cluster, so if all but one person in the cluster band together, they can deduce the remaining person in the cluster. Over the years, heuristics have been suggested to cure this problem,<sup>167</sup> but k-anonymity was nevertheless criticized when researchers proved that k-anonymity and its variants could not preserve privacy in principle and for most basic definitions of the term.<sup>168</sup> Moreover, it might become even less effective for protecting privacy in the IoT context, since the data are usually signals (e.g., audio, video, and GPS) and not strings. For example, it makes sense to remove the last digits

---

164. More specifically, Netflix offered an award for those that will improve the rating prediction of their users by more than 10%. To do so, they published records of movie ratings from thousands of "anonymized" users. Researchers compared these records with published IMDB records that included the names of reviewers. Since it is very unlikely that a pair of users will give exactly the same rank, even for as little as four movies, it was fairly easy to identify users in Netflix database by comparing them to the IMDB records, often revealing users' sex or political preferences, among other sensitive attributes. See Arvind Narayanan & Vitaly Shmatikov, *How to Break Anonymity of the Netflix Prize Dataset*, ARXIV (Oct. 18, 2006), <https://arxiv.org/abs/cs/0610105> [<https://perma.cc/2SJQ-7MTL>]; Ryan Singel, *Netflix Cancels Recommendation Contest after Privacy Lawsuit*, WIRED (Mar. 12, 2010), <https://www.wired.com/2010/03/netflix-cancels-contest> [<https://perma.cc/688X-4ZSC>]; Ohm, *supra* note 61, at 1720–21; see also Arvind Narayanan & Vitaly Shmatikov, *Robust De-anonymization of Large Sparse Datasets*, PROC. 2008 IEEE SYMP. ON RES. IN SECURITY & PRIVACY 111 (2008).

165. Steve Lohr, *Netflix Cancels Contest Plans and Settles Suit*, N.Y. TIMES BITS BLOG (Mar. 12, 2010, 2:46 PM), <http://bits.blogs.nytimes.com/2010/03/12/netflix-cancels-contest-plans-and-settles-suit/> [<https://perma.cc/XW5B-7G4K>]; Ohm, *supra* note 61, at 1722.

166. See Sweeney, *supra* note 162, at 557.

167. See Josep Domingo-Ferrer & Vicenç Torra, *supra* note 162 (criticizing systematically others' suggested heuristics).

168. See, e.g., Rolando Trujillo-Rasua & Josep Domingo-Ferrer, *On the Privacy Offered by (K, Δ)-Anonymity*, 38 IEEE INFO. SYSS. 491 (2013).

of zip codes in order to try to preserve privacy as usually done in k-anonymity; however, it is less clear what to remove from a GPS or a speech signal.

In other words, protecting privacy and innovation with de-identification methods proves a double-edged sword. Only aggressive suppression of data could make reidentification or deanonymization almost impossible, but such suppression would also make the data almost useless.<sup>169</sup> While prohibiting reidentification could seemingly help resolve this puzzle, such a mechanism might be difficult to implement and enforce.<sup>170</sup> When adversaries can easily and legally learn from publicly available information, anonymization methods will not properly advance privacy protection. Thus, in the IoT context, even with data anonymization adversaries might still be able to reveal the identity of the data subject.<sup>171</sup>

We now turn to one of the most common ways to protect privacy: encryption—a field usually related to security.<sup>172</sup> In the context of security, encryption is now ubiquitous for transmitting sensitive data online, such as a credit card number. Even if an unauthorized party saw the communication, it would learn nothing about the transmitted data in cleartext. Encryption is also effectively used for some IoT communications, meaning that an adversary that might view a communication without a decryption key will not be able to extract any data from the content of the message.<sup>173</sup> In some instances, such as in Google Drive, even the stored data might be encrypted, in a way that only the user (not even Google) will possess the secret encryption key.<sup>174</sup>

---

169. See Ohm, *supra* note 61, at 1714.

170. See *id.* at 1758.

171. Notably, these de-identification methods are analogous to early cryptography, from simple algorithms such as Caesar cipher that naively add, say, the number three to each letter, to the complicated Enigma machine of World War II whose code was broken during the war by Alan Turing. Like the new de-identification methods, these cryptographic schemes intuitively looked good, but in fact were broken by researchers, sometimes with the help of additional external databases or prior knowledge. In contrast, modern cryptography is based on provable reductions to mathematical problems that are assumed to be too hard to solve in practice and reasonable time using state-of-the-art software and hardware.

172. See Hui Suo et al., *Security in the Internet of Things: a Review*, 3 COMPUTER SCI. & ELECTRONICS ENGINEERING (ICCSEE) 650 (2012).

173. Amazon, for instance, declares that it encrypts all communication between the Amazon Echo, the Alexa App, and Amazon servers. See Kate O’Flaherty, *How to Secure the Amazon Echo*, FORBES (May 25, 2018, 2:26 PM), <https://www.forbes.com/sites/kateoflahertyuk/2018/05/25/amazon-alexa-security-how-secure-are-voice-assistants-and-how-can-you-protect-yourself/#476433cb3734> [https://perma.cc/NB83-MSP4].

174. See Darren Quick & Kim-Kwang Raymond Choo, *Google Drive: Forensic Analysis of Data Remnants*, 40 J. NETWORK & COMP. APPLICATIONS 179, 179 (2014). It should be emphasized that even if attackers cannot read the encrypted message, they may still learn meta-data regarding the message (e.g., when it was sent? What is its length? Etc.). Simple possible solutions were suggested in Adi Akavia et al., *Secure search on encrypted data via multi-ring sketch*,

In many cases, however, encryption of IoT data poses problems for innovating and providing services. Many, if not most, IoT devices provide services that rely on data processing, and it is vital to learn from many users' data for machine learning, deep learning, and big data analysis.<sup>175</sup> In this sense, it is vital that the OSPs learn from a private dataset, rather than just store an encrypted version of it. Thus, classic encryption will generally be problematic for IoT data.<sup>176</sup>

To some extent, new forms of encryption methods can help preserve users' privacy while maintaining data utility. One example is homomorphic encryption—a research area in cryptography that aims to solve the problem of outsourcing the computational task without risking privacy.<sup>177</sup> Generally, homomorphic encryption is designed to enable the server or cloud to run computation services without learning anything about the transmitted data in cleartext, by running it on the encrypted data and returning an encrypted result.<sup>178</sup> Hence, unlike standard encryption techniques, homomorphic encryption ensures that only the user possesses the secret key, while computations can be performed on the encrypted IoT data.<sup>179</sup>

---

2018 ACM CONF. ON COMPUTER & COMM. SECURITY (where the client communicates with the server all the time, possibly using dummy message, in order to hide the time stamp of the real messages).

175. See, e.g., Mohammad Saeid Mahdavejad et al., *Machine Learning for Internet of Things Data Analysis: a Survey*, 4 DIGITAL COMM. & NETWORKS 161 (2018).

176. For example, in order for Amazon's Alexa to answer a user's question, Amazon must not only obtain the user's voice records, but also process them and return the answer to the user. The processing further requires using Alexa's powerful computation service and accessing all of its databases. See Hyunji Chung et al., *Digital Forensic Approaches for Amazon Alexa Ecosystem*, 22 DIGITAL INVESTIGATION 15 (2017).

177. See CRAIG GENTRY & DAN BONEH, A FULLY HOMOMORPHIC ENCRYPTION SCHEME 20 (2009).

178. The question whether it is even possible to run any algorithm on encrypted data without knowing the secret key was raised in 1978, within one year of the development of RSA—the first and most common message encryption algorithm. See Ronald L. Rivest et al., *On Data Banks and Privacy Homomorphisms*, in FOUNDATIONS OF SECURE COMPUTATION 169 (1978). For over thirty years, it was unclear whether a solution, called a fully homomorphic scheme, existed. The first construction was suggested only in 2009 and was considered a major theoretical breakthrough. See Craig Gentry, *Fully Homomorphic Encryption Using Ideal Lattices*, 41 ACM SYMP. ON THEORY OF COMPUTING (STOC) 2 (2009).

179. To exemplify, suppose that the IoT device (client) wants to solve a problem or compute  $f(D)$  on its data  $D$ , where  $f$  is the desired function, task, or algorithm. The client encrypts the data  $D$  to get its encrypted version  $[D]$  and sends it to the cloud. Homomorphic encryption allows the cloud to compute  $[f(D)]$ , the encrypted version of  $f(D)$ , using only  $[D]$ . It then sends  $[f(D)]$  to the IoT device that decrypts  $[f(D)]$  using its internal secret key and obtain the result  $f(D)$  to perform the client's command.

Today, however, homomorphic encryption is used relatively rarely, as it runs into practical barriers.<sup>180</sup> In the context of privacy for IoT, using this method entails two main disadvantages. First, while homomorphic encryption solves the computational outsourcing problem on the cloud, it does not enable the OSP to learn the transmitted data in cleartext from users' statistics to improve its model, since we cannot learn from encrypted data without having its key. Second, while homomorphic encryption might sound like a good solution in theory, in practice it is known to be unwieldy and unworkable, except for very simple tasks of adding encrypted numbers.<sup>181</sup> In particular, while in theory any algorithm can be applied to the encrypted data, hardly any machine-learning algorithms that can run in this model exist in practice. This makes homomorphic encryption currently unsuitable for many, if not most, IoT services that run machine-learning algorithms.<sup>182</sup>

The potential technological solutions presented in this Section alone are therefore currently ineffective for preserving users' privacy and data utility. Still, technological solutions can be viable if they combine new techniques with at least some of the existing techniques presented in this Section. As the next Section argues, a relatively new approach in computer science could effectively preserve users' confidentiality (to some extent) while keeping data utility at a proper level for the IoT context.

#### B. DIFFERENTIAL PRIVACY USING CORESETS

Instead of focusing on protecting merely personally identifiable information as in the method of anonymization, this Article suggests focusing on the data subjects themselves.<sup>183</sup> By doing so, this Section proposes a model that can manage the practical and utility issues arising from handling IoT data and preserve provable guarantees regarding users' privacy at the same time. We intend to expand the use of mathematical tools in the context of privacy,<sup>184</sup> while further addressing the core values of the sectoral approach in the always-

---

180. Wei Wang et al., *Accelerating Fully Homomorphic Encryption Using GPU*, 2012 IEEE CONF. ON HIGH PERFORMANCE EXTREME COMPUTING IEEE 1 (2012).

181. *See id.*

182. Exact running times and performance measures can be found in Miran Kim et al., *Secure Logistic Regression Based on Homomorphic Encryption: Design and Evaluation*, 6 JMIR MED. INFORM. 1, 1–3 (2018).

183. *See* Nissim et al., *supra* note 153, at 687–88.

184. For other suggestions to combine mathematical tools within the notion of privacy protection, see Omar Chowdhury et al., *Privacy Promises That Can Be Kept: A Policy Analysis Method with Application to the HIPAA Privacy Rule*, PROC. 18TH ACM SYMP. ON ACCESS CONTROL MODELS & TECH. 3 (2013); Henry DeYoung et al., *Experiences in the Logical Specification of the HIPAA and GLBA Privacy Laws*, PROC. OF 9TH ACM WORKSHOP ON PRIVACY IN ELECTRONIC SOC'Y (2010).

on era in the following Section. Our model is based on a combination of a few recent techniques in the theory of differential privacy,<sup>185</sup> computational geometry,<sup>186</sup> and homomorphic encryption.<sup>187</sup> The link among these techniques is a modern data summarization technique named coresets (or core-sets), as will be further explained.

Introduced in 2006, differential privacy is a standard that strives to assure that the presence or absence of an individual in a dataset does not make any significant difference to the outcome of any given database query.<sup>188</sup> It mathematically ensures that breaking confidentiality will be limited in probability,<sup>189</sup> and that individuals' data could remain in the database without anyone knowing that it exists.<sup>190</sup> It does so by sanitization of the data (i.e., by adding noise ("blur") to the data) in order to hide information about individual users, while keeping the global statistics, or the ability to construct efficient classifiers from the sanitized data.<sup>191</sup>

Before exemplifying the use of differential privacy in the context of IoT, first it is important to clarify how noise could be introduced. Deciding the level of sanitization or noise to be added to the data requires discussion of two computation models and communication protocols: centralized or local. Under a centralized model, the OSP collects the data from its users and is also responsible for adding the noise.<sup>192</sup> The original data must be deleted or at least not be used by the OSP prior to adding the noise. The data will then be sanitized with noise, and the learning algorithms will be fed the "sanitized

---

185. For more on differential privacy, see Cynthia Dwork, *Differential Privacy: A Survey of Results*, in THEORY AND APPLICATIONS OF MODELS OF COMPUTATION 1 (Manindra Agrawal et al. eds., 2008).

186. See Dan Feldman & Michael Langberg, *A Unified Framework for Approximating and Clustering Data*, PROC. 43D ANN. ACM SYMP. ON THEORY COMPUTING 569, 569–71 (2011).

187. See Adi Akavia et al., *Secure Search on the Cloud via Coresets and Sketches*, ARXIV (Aug. 19, 2017), <https://arxiv.org/abs/1708.05811> [<https://perma.cc/Z4C5-HAF7>].

188. See Cynthia Dwork et al., *Calibrating Noise to Sensitivity*, PRIVATE DATA ANALYSIS, PROC. 3RD CONF. THEORY OF CRYPTOGRAPHY 265 (2006); Chin & Klinefelter, *supra* note 160, at 1427 (citing Cynthia Dwork, *A Firm Foundation for Private Data Analysis*, COMM. ASS'N FOR COMPUTING MACHINERY 86, 91 (2011)). For more on differential privacy, see Felix T. Wu, *Defining Privacy and Utility in Data Sets*, 84 U. COLO. L. REV. 1117, 1139–40 (2013); Ohm, *supra* note 61, at 1756; Chin & Klinefelter, *supra* note 160, at 1452–54; Jane Bambauer et al., *Fool's Gold: An Illustrated Critique of Differential Privacy*, 16 VAND. J. ENT. & TECH. L. 701, 712–17 (2014).

189. See Ohm, *supra* note 61, at 1756.

190. See Chin & Klinefelter, *supra* note 160, at 1430.

191. See Shuchi Chawla et al., *Toward Privacy in Public Databases*, 2 THEORY OF CRYPTOGRAPHY CONF. (2005).

192. See, e.g., Xi Xiao et al., *CenLocShare: a Centralized Privacy-preserving Location-sharing System for Mobile Online Social Networks*, 86 FUTURE GENERATION COMPUTING SYSS. 863 (2018).

data.” The centralized model is often used as a method for allowing data use by third parties or for spreading data in different OSP departments, thus lowering the risk of information leakage to competitors by employees.<sup>193</sup> But this model is weakened by its reliance on trust and efficiency. The method relies on trusting the OSP to actively sanitize data and on sanitization by OSPs, proving too late for some users. Data from IoT devices might be hacked, stolen, lost, or just poorly sanitized.<sup>194</sup>

The alternative model is local, wherein the IoT client does not share its raw data with the OSP.<sup>195</sup> Instead, the sanitization is done *ex ante* on the client’s side. Only the sanitized dataset is sent to the OSP. As a result, trusting the OSP and preventing data leakage prior to sanitization are not a challenge as long as the sanitization is performed properly. The disadvantage of the local model is that the per-user raised noise level is significantly greater than in the centralized approach, where small added noise suffices to blur the original statistics. In addition, in many instances locally added noise is still too low to preserve users’ privacy.<sup>196</sup>

Unlike previous techniques, such as k-anonymity, which also rely on mechanisms such as adding noise or hiding data, differential privacy suggests a very strong definition of privacy that is resistant to external databases that the adversary may have. An algorithm is differentially private only if it provably meets this definition. Such an algorithm is unlike de-identification techniques, with which a sanitized database has noise added per record (and the adversary

---

193. For example, some argue that Facebook uses this technique to publish click rates to its ad publishers. See Yehuda Lindell & Eran Omri, *A Practical Application of Differential Privacy to Personalized Online Advertising*, 2011 IACR CRYPTOLOGY EPRINT ARCHIVE 152 (2011). This method is also common in governments’ Bureau of Statistics in order to publicly share their collected data. See Boaz Barak et al., *Privacy, Accuracy, and Consistency Too: a Holistic Solution to Contingency Table Release*, PROC. 26TH ACM SIGMOD-SIGACT-SIGART SYMP. ON PRINCIPLES DATABASE SYSS. (2007).

194. See, e.g., Ryan Singel, *Netflix Spilled Your Brokeback Mountain Secret, Lawsuit Claims*, WIRED (Dec. 17, 2009), <https://www.wired.com/2009/12/netflix-privacy-lawsuit> [<https://perma.cc/BP47-FGS3>]; Narayanan & Shmatikov, *supra* note 164, at 6–10.

195. For more on the local model, see generally Peter Kairouz et al., *Extremal Mechanisms for Local Differential Privacy*, 17 J. MACHINE LEARNING RES. 1 (2016).

196. For example, instead of collecting GPS data from its users and adding noise to it (i.e., centralized privacy), Apple added sanitization mechanisms on its smartphones, so some type of noise is added to the GPS samples before transmitting them to Apple from the user’s smartphone. However, researchers that reverse engineered this protocol claimed that the amount of noise added is far too small to preserve users’ privacy. See Andy Greenberg, *How One of Apple’s Key Privacy Safeguards Falls Short*, WIRED (Sept. 15, 2017, 09:28 AM), <https://www.wired.com/story/apple-differential-privacy-shortcomings> [<https://perma.cc/LAN5-K7LC>]; Jun Tang et al., *Privacy Loss in Apple’s Implementation of Differential Privacy on macOS 10.12.*, ARXIV (Sept. 11, 2017), <https://arxiv.org/pdf/1709.02753.pdf> [<https://perma.cc/4R58-KH6V>]. See generally Chin & Klinefelter, *supra* note 160.

can tell, for example, if another user has been added since the previous version of the database). Instead, a differentially private algorithm completely replaces the database with a new database containing “global” noise.

The problem, however, is that most of the literature in modern computer science that discusses privacy, including k-anonymity and differential privacy, does not generally fit the IoT model. In particular, as was the case with the Netflix Prize, the literature assumes that the OSP holds the complete original (non-sanitized) database of its users. When that is the case, privacy issues are assumed to arise only once the OSP reveals its user data to a third party (e.g., when Google sells or discloses parts of its database to advertisers) or at least reveals a derivation of the data, such as classifiers or statistics, that may leak information about individuals.

By contrast, in the foregoing Sections we assume that there are *many* users that send their private data to the OSP and wish to preserve their privacy. While in principle the OSP might itself add noise to the collected records—that is, after collecting and before using them—this is too late if the users do not trust the OSP. The hidden and natural implication is that the noise should be added *locally* on the user’s side *before* even reaching the OSP. This communication model for privacy was suggested relatively recently and requires much more noise to be added than the centralized model requires.

To understand how differential privacy could help in the always-on era, we begin with an example: baby diapers. Suppose a company that sells diapers asks how many Amazon Echo users have a baby at home. The motivation may be to decide where to place their ads or perhaps to use the device itself for marketing purposes. This can be done, for example, by listening for a baby crying at some point in time or analyzing the voice after a conversion. If  $x_i = 1$  when the *i*th client has a baby at home and  $x_i = 0$  when not, the answer when there are *n* clients is the sum:  $S = x_1 + x_2 + \dots + x_n$ . Suppose that the diapers company already knows some of these values (e.g., the sum of the first *n*-1 numbers in this equation), through either relying on another database or computing this number on the day before the last *n*th client joined.

Within the diapers example, if Amazon publishes this number, and the diapers company knows the sum of the first *n*-1 clients, they will be able to compute the value  $x_n = S - x_1 - \dots - x_{n-1}$  of the last client. That is, they will be able to compute whether the *n*th client has a baby at home. To avoid this, Amazon could compute the sum *S* and add a little noise before presenting the noisy value of *S* to the diapers company. Even if the diapers company knows all the values of  $x_i$  except one, as long as sufficient noise was added, the company will not be able to extract  $x_i$  from the noisy value of *S*. Thus, they will not be able to compute whether the *n*th client has a baby at home.



Some scholars claim Facebook already uses this technique for sharing information about the number of its users' clicks for third-party sponsors.<sup>197</sup>

The main challenge in using differential privacy is knowing how to add noise to the data that is sufficiently large to preserve the individual's privacy, but sufficiently small to allow a good approximation of useful statistics. For example, if a random number is added to a given sum of numbers from a Laplacian<sup>198</sup> distribution with zero mean and scale of roughly  $1/\epsilon$ , where  $\epsilon$  is a number usually between 0 and 1, then the adversary that receives  $S$  will (depending on the value of  $\epsilon$ ) be unable to determine whether any specific  $x_i$  is 0 or 1, even if the adversary knows that this particular algorithm outputted  $S$ .

More precisely, the probabilities that  $x_i = 0$  and  $x_i = 1$  given  $S$  are approximately the same, up to an additive error of  $\epsilon$ . In other words, whatever the adversaries know or wish to know regarding a specific value  $x_i$ , and whatever external database or knowledge they already have, they will not be able to learn about  $x_i$  merely because it was part of the original input. Formally, the output of the randomized algorithm that computes the approximation of  $S$  has the same distribution (up to  $\epsilon$  additive factor) if we change a single value. True, a dummy algorithm that outputs a random number will also satisfy this privacy guarantee. However, the above private algorithm is more desirable because it is efficient: with high probability, depending on  $\epsilon$ , it gives a good, provable approximation of  $S$ . This property of allowing approximation is called the utility of the algorithm.

In the diapers example, we assumed that Amazon computes  $S$  in a centralized private fashion. But a locally private version for the above solution is also possible. If we do not want the actual values of  $x_i$  to get to Amazon via its Echo device in the first place, each  $i$ th device should add its own noise to its value  $x_i$ . A common approach that has provable guarantees is to send the "wrong" value with some fixed probability. For example, an algorithm might send as its vote the real binary value  $x_i$  with probability 0.75, and otherwise send the "wrong" value  $1 - x_i$ .<sup>199</sup> The privacy of each user is preserved in the sense that with 0.25 probability (a 25% chance),  $x_i$  is not the real value, while

---

197. See generally Chin & Klinefelter, *supra* note 160.

198. The Laplacian distribution is used since it is proportional to  $\exp(-|x|)$ , which yields the desired property  $\exp(-|x + \epsilon|)/\exp(-|x|) = \exp(-\epsilon)$ . This property does not hold for, for example, the Gaussian distribution that is proportional to  $\exp(-x^2)$ . For further detail, see sources cited *supra* note 188; Dan Feldman et al., *Private Coresets*, PROC. 41ST ANN. ACM SYMP. ON THEORY OF COMPUTING (2009).

199. See generally Stanley L. Warner, *Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias*, 60 J. AM. STAT. ASSOC. 63 (1965).

the utility is preserved for a sufficiently large number of users: for example, if the approximated value of  $S$  is  $n/4$ , then probably most of the users sent 0, and  $n/4$  is due to the noise. On the other hand, if  $S = \frac{3n}{4}$ , the real value of  $S$  is close to  $n$ . In both of these cases, there is a “doubt” of probability 0.25 whether each user’s vote was her real value or the opposite. This is regardless of her real value, or the real values of the other people in the group. Note that the expected error of 0.25 in this local privacy model is by order of magnitude larger than the expected error in the previous centralized model of  $\frac{1}{n}$ , which also decreases with the number of users.

Notably, the use of differential privacy as a solution for privacy has also been subjected to criticism in various respects. Scholars argue that introducing noise is limited in value for a few reasons: the use of noisy data might yield inaccurate outcomes; it might require complex and costly calculations; and it could cause chaos in database systems.<sup>200</sup> It will also not apply to data already collected. However, while acknowledging potential shortcomings, some scholars have recently considered differential privacy as a potential solution to protect educational privacy from a legal perspective.<sup>201</sup>

Indeed, one of the main challenges in differential privacy is to explain the intuition behind its guarantee of privacy. A common explanation is that a differential privacy algorithm is private in the sense that an adversary can only obtain information that can be learned *whether you participate in the database or not*. In other words, the algorithm and its output are insensitive to any single user. For instance, given a sanitized database of patients, we may conclude that heavy smoking may cause cancer. So, if we see someone smoking, we can infer that she may have a higher probability of getting cancer. While we learned *something* about this individual from the database, it was not because she participated in the database; indeed, the database may not have contained any information specific to her at all. If her specific information were added to the database, we would learn nothing new about her. Thus, in that sense, her privacy is preserved. In contrast, this property does not hold in alternatives with no provable guarantees, such as in the Netflix case. In those cases, we can often learn information about an individual that we would never have known

---

200. See Arvind Narayanan & Vitaly Shmatikov, *Privacy and Security Myths and Fallacies of “Personally Identifiable Information”*, COMM. ASS’N FOR COMPUTING MACHINERY 24, 26 (2010), [http://www.cs.utexas.edu/~shmat/shmat\\_cacm10.pdf](http://www.cs.utexas.edu/~shmat/shmat_cacm10.pdf) [https://perma.cc/YUG4-RTXH]; Bambauer et al., *supra* note 188, at 704; Ohm, *supra* note 61, at 1757.

201. See generally Nissim et al., *supra* note 153 (“[D]ifferential privacy satisfies a large class of reasonable interpretations of the FERPA privacy standard.”). Notably, the scholars argue that FERPA and differential privacy were used to illustrate an application of their approach and that it “may be developed over time and applied, with potential modifications, to bridge between technologies other than differential privacy and privacy laws other than FERPA.” *Id.*

but for her specific information being included in the database. Admittedly, this intuition about differential privacy can be difficult to grasp.

Another challenge of differential privacy is applying it to more complex calculations. The diapers example involved only simple calculations. However, in applications such as machine learning, the input is not just a binary number, but a long database record of each user. Moreover, the data emitted is not from mere summation, but from more complicated functions such as neural networks or logistic regression. More generally, differentially private algorithms traditionally use impractical models, such as Curator,<sup>202</sup> and solve very specific, and arguably artificial, theoretical problems with guarantees that engineers and lawyers may understand significantly less than the other methods.<sup>203</sup>

One final challenge is practical and arises from utility challenges. In many existing protocols, such as Curator, the user is limited to asking a small number of very specific questions. However, while data scientists and learning algorithms may handle added noise, they usually need to learn a single sanitized database for all their queries. Even when sanitized databases are used today, they are useful for a specific family of problems.

To overcome the challenges of differential privacy, we propose combining it with other techniques, most importantly with the notion of coresets: a small representation of the data, such that querying the coreset will yield a provably small approximation to the original data. In particular, solving an optimization problem or running a learning algorithm on the coreset will yield a near optimal

---

202. Curator is a model of computation. In this model, there is a specific service (e.g., a web interface) with access to the original (non-noisy) data of users that gives noisy answers to given questions. Usually, these questions are restricted to a specific type. This service—the curator—may operate either on the end user’s side and answer queries of the company, or between the company that stores the original data and its third-party clients. The curator can usually answer specific types of statistical questions regarding the data with some additional noise. Since each answer admits some small privacy leakage, once a specific number of questions have been asked, the curator refuses to answer further questions in order to avoid too much leakage. The curator model is more common in academic papers than in practice: while it can provide provable privacy guarantees, data scientists are accustomed to working with databases (noisy or otherwise) rather than such answering services. Moreover, machine-learning algorithms are usually applied and trained on data sets, and it is unclear how to use curators with these methods. For more on the curator model, see CYNTHIA DWORK & AARON ROTH, *THE ALGORITHMIC FOUNDATIONS OF DIFFERENTIAL PRIVACY* 211–407 (2014).

203. This relates to an academic debate that has serious implications for the industry and IoT privacy. In particular, it implicates what it means to preserve the privacy of a user and how to achieve it. See Bambauer et al., *supra* note 188, at 712–17; Nancy Victor et al., *Privacy Models for Big Data: A Survey*, 3 INT’L J. BIG DATA INTELLIGENCE 61, 61–65 (2016); George Danezis & Seda Gürses, *A Critical Review of 10 Years of Privacy Technology*, 2010 PROC. SURVEILLANCE CULTURES: A GLOBAL SURVEILLANCE SOC’Y 1 (2010).

solution on the original data.<sup>204</sup> Unlike other forms of lossy compression, like MP4 or JPEG compression, coresets are considered lossy compression for specific optimization problems or statistics to be applied on the data, rather than generic compression of the data itself. Coresets have been suggested in recent years to resolve key problems in machine learning, and a single coreset may be the union of multiple coresets to solve the numerous corresponding problems.<sup>205</sup> By means of a technique called sup-sampling (or non-uniform sampling), a single coreset may handle many problems by uniting coresets of the database for these problems.<sup>206</sup>

Coresets are especially useful in the context of IoT and Big Data in general, since they usually possess an important property: the union of a pair of coresets yields a coreset for the union of the underlying data.<sup>207</sup> This implies that one can compute the coreset on streaming IoT data by compressing small batches of data that arrive on the fly and recompress them. So, at any given moment, we can have a small coreset for all the seen streamed data and thus can apply existing (possibly inefficient) algorithms to the small data. Similarly, the IoT service provider can easily compute coresets on the cloud by computing a coreset in each machine on its own streaming data. Then, a main server can collect the coresets for all the machines and solve the optimization problem on it, possibly after an additional final compression.<sup>208</sup>

---

204. See Feldman & Langberg, *supra* note 186.

205. See Artem Barger & Dan Feldman, *k-Means for Streaming and Distributed Big Sparse Data*, PROC. 2016 SIAM INT'L CONF. ON DATA MINING, SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS 1–2 (2016); Dan Feldman et al., *Coresets for Vector Summarization with Applications to Network Graphs*, INT'L CONF. ON MACHINE LEARNING (2017).

206. See Michael Langberg & Leonard J. Schulman, *Universal  $\epsilon$ -approximators for Integrals*, PROC. 21ST ANN. ACM-SIAM SYMP. ON DISCRETE ALGORITHMS SOC'Y INDUS. & APPLIED MATHEMATICS (2010). Even without knowing how to compute a coreset for a given problem or classifier, a coreset for a related problem may suffice. In practice, a coreset generally yields a good approximation for a problem it was not designed to solve, since intuitively, a representative point for one problem is also a good representation for the other problem. A coreset for an optimization problem is usually more general in the sense that it usually can approximate queries of a certain type and not just solve the relevant optimization problem. See Rohan Paul et al., *Visual Precis Generation Using Coresets*, 2014 IEEE INT'L CONF. ON ROBOTICS & AUTOMATION (2014).

207. See Piotr Indyk et al., *Composable Core-sets for Diversity and Coverage Maximization*, in PROC. 33RD ACM SIGMOD-SIGACT-SIGART SYMP. ON PRINCIPLES DATABASE SYSS. (2014).

208. Indeed, coresets for many fundamental problems in machine learning, including experimental results, appeared during the recent decade in machine learning conferences. See, e.g., Dan Feldman & Tamir Tassa, *More Constraints, Smaller Coresets: Constrained Matrix Approximation of Sparse Big Data*, PROC. 21TH ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING (2015); Mario Lucic et al., *Training Gaussian Mixture Models at Scale via Coresets*, 18 J. MACHINE LEARNING RES. 5885 (2017).

The main challenge in the research of coresets construction is thus to prove that, for any possible input set, we can compute a small coreset whose approximation error is small, with a good tradeoff between the approximation error (say  $\epsilon$ ) and size of the coreset (say  $1/\epsilon$ ). But how can coresets, which enable efficient compression of the data, help us solve the utility issues of differential privacy? In principle, there is no linkage between the ability of data compression, as in the coresets above, and the ability to add a small amount of noise that will preserve the desired approximation error while maintaining privacy—we can always reduce a sanitized database, using coresets to reduce its size, without losing more privacy. However, a perhaps surprising theorem forges a link between the two: if we have a (non-private) small coreset for a problem, we also can have a sanitized database (called a private coreset) where the size of roughly  $1/\epsilon$  of the small coreset turns into the additive error (noise) of a similar order.<sup>209</sup> That is, a (non-private) coreset of small size implies a (not necessarily small) private coreset that is computed via a differential  $\epsilon$ -private coreset.

Different from the Curator model, such a sanitized database can be queried unlimited times without further information leakage once the private coreset is computed from the raw data. That is, the existence of a small coreset implies a sanitized database that preserves the desired statistics (in terms of utility) and also preserves privacy. This theorem is very promising, (e.g., for machine learning in IoT), since dozens of coresets for main problems are already known. Unfortunately, the proof of the above theorem is not constructive in the sense that the computation time of this generic coreset reduction is impractical. How to implement it efficiently is still an open question. Instead, specific private coresets have been suggested in recent years for specific problems.<sup>210</sup>

Hence, private coresets may be used to obtain a single sanitized database, unlimitedly applicable to many machine-learning algorithms with no additional noise. Now the problem remains of computing a private coreset with little added noise (as in the centralized model) while using the localized model for IoT applications. Recall that the main advantage of centralized models compared to local ones is that less added noise is required, and the main disadvantage is that the company has the users' original (non-noisy) data, whereas in local privacy the user sends only noisy data. Indeed, while private

---

209. See Dan Feldman et al., *Private Coresets*, PROC. 41ST ANN. ACM SYMP. ON THEORY COMPUTING (2009).

210. See Dan Feldman et al., *Coresets for Differentially Private K-Means Clustering and Applications to Privacy in Mobile Sensor Networks*, 2017 16TH ACM/IEEE INT'L CONF. ON IEEE 3 (2017).

coresets may be computed on the client's side (as in local privacy) or on the server's side (as in centralized privacy), the following technique may allow us to get the small error of centralized privacy, while preserving the client's privacy as in local privacy. To that end, we suggest computing private coresets using homomorphic encryption.

More precisely, the Homomorphic Encryption Coreset (CHE)<sup>211</sup> is a modern tool that may resolve this conflict, namely, to get a small error without letting the company access the original raw IoT data. We denote by  $D$  the database of users' data, and by  $\text{sanitized}(D)$  its sanitized version (private coreset). In the traditional centralized model, users send their records of raw data to the OSP, which maintains the database  $D$  and then computes its sanitized version  $\text{sanitized}(D)$  that can be used for publishing or learning without sacrificing privacy. The main challenge is learning how the OSP can compute the private coreset  $\text{sanitized}(D)$  without having access to the original database  $D$ .

We suggest using homomorphic encryption to compute differential private coresets as follows: instead of sending their original records, or noise records, the client or clients send an encrypted version of their records (but without noise)  $[D]$  of, for example, GPS or an Echo's data  $D$  to the OSP. The OSP adds noise to the data, as in the centralized model. However, this is done on the encrypted version of the data so there is no privacy loss at all. The result is an encrypted private coreset  $[\text{sanitized}(D)]$ . Now, this is a private coreset that can be exposed to the OSP, but it is still encrypted. At this point the OSP sends the data back to the clients' IoT device that uses its secret key to decrypt  $[\text{sanitized}(D)]$  and obtain the private (non-encrypted) coreset. This sanitized dataset  $\text{sanitized}(D)$  is sent back to the server, which can use it (e.g., to improve the machine-learning results for other users as well). Under this proposition, only a small amount of noise has been added by the server as in centralized privacy, and still the server has never seen the complete original data. More generally, this problem can be applied to data from multiple users where each user has its own key and the sanitized database is computed for all of them.<sup>212</sup> This is how private coresets via homomorphic encryption can make differential privacy more practical, without losing its theoretical guarantees.

Gaps and many handling problems in IoT still await private solutions, on both the theoretical and practical sides of computations, as well as through laws and regulations. For example, the algorithms for computing sanitized

---

211. See generally Adi Akavia et al., *Secure Search via Multi-Ring Fully Homomorphic Encryption*, 25TH ACM CONF. ON COMPUTER & COMM. SECURITY (2018).

212. See Adriana López-Alt et al., *On-the-Fly Multiparty Computation on the Cloud via Multikey Fully Homomorphic Encryption*, PROC. 44TH ANN. ACM SYMP. ON THEORY COMPUTING (2012).

databases are usually applied to static database records. However, IoT is made of streaming data that grow larger with time. There are very few results for handling such streaming data privately, especially without introducing too large an additive noise. Similarly, for most problems it is not clear how to compute the sanitized database in parallel on distributed data (e.g., cloud or smartphones), unless we use local privacy. Private coresets may help in handling these issues because there are simple reductions that show how, given a coreset for a set of models, we can compute it on the streaming and distributed model. The main idea is that two coresets can be computed independently on distributed machines such as the cloud, or different subsets of streaming data. Then they can be merged and reduced again on each machine or device.

Our proposed model thus requires further discussion on when to add noise to IoT devices, and what level of noise will ultimately preserve privacy and remain useful for OSPs. Too much noise will make the data unusable, while too little will defeat the purpose of preserving privacy. Accordingly, the next Section proposes a theoretical *ex ante* approach which strives to protect privacy by the probability of gathering sensitive data, which will depend on various factors.

#### C. MEASURING NOISE VIA THE PROBABILITY OF SENSITIVITY

To date, scholars have only limitedly applied differential privacy in the context of sectoral privacy protection by, for example, offering differential privacy to satisfy the requirements of a particular legal standard of privacy (FERPA in this instance).<sup>213</sup> Our intention is to broaden this innovative argument. Using the concept of differential privacy, combined with other mathematical models, we offer an analytic framework for any policymaker—taking the U.S. approach to privacy—to protect privacy while still acknowledging the value of data. Further, the level of noise added to the model could be measured—at least to some extent—on the potential sensitivity of the data, depending on various factors related to the IoT in question.

It is generally difficult to define when data become sensitive, although a few scholars have attempted to do so.<sup>214</sup> Evaluating the probability that data will be sensitive *ex ante*—under the sensitive categories that Congress has set—is even more ambitious. It is generally an almost impossible task to accomplish. Sectoral privacy, however, is implemented almost precisely by this ambitious method. It is regulated through the notion that with some entities,

---

213. *See generally* Nissim et al., *supra* note 153.

214. *See, e.g.*, ÉLOÏSE GRATTON, UNDERSTANDING PERSONAL INFORMATION: MANAGING PRIVACY RISKS (2013); Ohm, *supra* note 1, at 1733–34.

and in some contexts, there is increased probability that sensitive data will be shared digitally, as is the case in medical, educational, or financial institutions. This measure could be also implemented—perhaps even more accurately—in the always-on era. Thus, without belittling the important scholarly debate on data sensitivity, this Article focuses on the current categories of sensitive information, as reflected in the federal statutes of sectoral privacy: financial data, health information, education records, children’s data, and consumer data. As shown below, considering various factors that relate to IoT, and the nature of the use of these devices, could aid in assessing such probability.

The factors to consider when assessing the probability of IoT devices gathering sensitive data depend on various potential characteristics. As explained below, the factors we suggest include the device’s architecture, its sensors, its physical location, the nature of gathered data, and the age of its potential users. The probability of infringing on information privacy should be evaluated through the aggregation of these factors: that is, prior to any query from the database. But this does not mean that these factors should not be reevaluated continuously. On the contrary, we encourage such reevaluation, followed by proper modifications and adaptations. Additionally, these factors might greatly vary depending on their users’ input, implemented with the use of the device.

We begin with the architecture. As previously noted, not all IoT devices operate alike. Some might have to be turned on manually to begin their data collection (e.g., the smart connected toy Hello Barbie). Others operate in an “always-ready” mode like Amazon Echo or Google Home, meaning that they await their trigger phrase prior to any data collection or retention. Finally, we have the devices that are “always on” (i.e., that constantly collect and transmit data, like Fitbit).<sup>215</sup>

The architecture of the device could greatly influence the probability of collecting sensitive data. Clearly, and without considering other factors like the types of data collected, devices that constantly collect data will have greater probability of collecting sensitive data than always-ready devices. Consequently, in many instances always-ready devices could have higher probability of collecting sensitive data than those operated manually, simply due to their architecture (i.e., it is much more convenient for many individuals to use them so they could be more frequently used). We suggest that OSPs

---

215. Fitbit is a fitness tracker that monitors steps and could provide insights on, inter alia, an individual’s heart rate or quality of sleep. See Andrew Hilts et al., *Every Step You Fake: A Comparative Analysis of Fitness Tracker Privacy and Security*, OPEN EFFECT REP. 3–6 (2016), [https://openeffect.ca/reports/Every\\_Step\\_You\\_Fake.pdf](https://openeffect.ca/reports/Every_Step_You_Fake.pdf) [<https://perma.cc/89PV-Z5TF>].



might be required to differentiate three probability levels, depending on architecture: manual, always-ready, and always-on.

The second factor to be considered is the sensors of the device. They may vary greatly, but we can still further divide them into five categories: sensors that measure the environment (e.g., temperature or air quality), that measure human activity (e.g., movements, location, and heartrate), that capture written communication, that capture oral communication, and that capture visual communication (cameras). The types of sensor could greatly affect the probability of gathering sensitive data. In this regard, every type of sensor could be given a numerical representation related to the value of  $\epsilon$ .

The third factor is the device's physical location. Some devices are more portable than others. Some might be placed in locations that may have a greater probability of gathering sensitive data than others. As for the first argument, devices like smart refrigerators or smart washing machines will most likely not change their placement much, while mobile health ("mHealth") wearables, like Fitbit, are more likely to be on the move. Even smart personal assistants can be moved more easily than smart refrigerators or washing machines, hence could be placed anywhere that has connectivity to both external power and the internet.

The placement of such devices could affect the sensitivity of data gathered as some locations could impact the type of data conveyed. For instance, a house's bedroom could convey data on sexual activity more than the kitchen area. Placing an Amazon Echo in your living room might not be the same as placing a similar device in your office. However, this concern must be evaluated with respect to sectoral privacy. It is extremely difficult to determine the link between placement and sensitivity in general, so instead we merely suggest considering two broad categories of devices: wearable and not wearable. A wearable device, especially one constantly worn, is more likely to gather sensitive data simply due to its ability to collect sensitive data directly from an individual more easily.

The fourth factor to be considered is the nature of the gathered data. Some IoT do not gather any sensitive data at all, or at least have a low probability of gathering such data. For instance, if a smart refrigerator knows which food or drinks it contains, and perhaps even consumption habits, it has low probability of collecting sensitive data, if any at all. Other IoT devices have a higher probability of collecting such sensitive data. While Amazon Echo is not generally likely to gather health information in the course of its communications, as its business model does not depend on health data per se, it might do so upon communicating with it, and it could gather consumer-sensitive data if used for purchasing. A smart TV could also fit this category,

as it might be aware of your watching habits and might also capture sensitive communication. Finally, we have the devices that by default collect sensitive data. Health wearables, for example, depend on health data; hence the gathering of such data is almost certain. Thus, the nature of gathered data will depend on the three-fold categorization of their potential nature: low, high, and certain.

The fifth and final factor to be considered is the user's age. As federal law generally protects children younger than thirteen in some circumstances,<sup>216</sup> we must evaluate the probability of having an IoT device gather data from this cohort. Thus, this factor divides devices into two main categories: devices that are, and that are not, targeted at children younger than thirteen. For the devices that do target children, like smart connected toys or kids' wearables, the probability of gathering sensitive data is absolute. This is the easy case, as COPPA applies to the OSPs of these devices which must ensure that they comply with its regulations.<sup>217</sup> This category is thus excluded from this factor, as it is already implemented and labeled "certain" through the fourth factor, namely the nature of gathered data. The second sub-category is more challenging and will depend, inter alia, on users' inputs. When configuring the device, users will be obliged to answer various questions that will help determine the probability that children's data will be obtained. So, if a user operates an Amazon Echo device in his or her living room, having children aged under thirteen in the household will increase the probability of gathering such data. This probability will change depending on the number of children in the household and their cognitive abilities, among other potential factors.

The probability of sensitivity can be calculated through each of these five factors, and perhaps mostly through their correct combination. This calculation will involve an ex ante evaluation of the IoT device in question, along with input from users (e.g., information about whether there are children present in a household) that will allow fine-tuning of such an evaluation. Upon evaluation of these factors, and perhaps others, OSPs could translate them into a relative ordering of whether the risk is high, medium, or low.<sup>218</sup> Such probability could be implemented through the mathematical models we have suggested, thereby adding noise only to the IoT devices that present higher probability of sensitivity without an ex post evaluation of the data or the data

---

216. See 16 C.F.R. § 312.2 (2018); 15 U.S.C. §§ 6501(1), 6502, 6501(8) (2012).

217. See *Children's Online Privacy Protection Rule: A Six-Step Compliance Plan for Your Business*, FED. TRADE COMM'N (June 2017), <https://www.ftc.gov/tips-advice/business-center/guidance/childrens-online-privacy-protection-rule-six-step-compliance> [<https://perma.cc/276M-9AH2>] (last visited Feb. 10, 2019). For more on smart connected toys and COPPA, see generally Haber, *supra* note 95.

218. See Ohm, *supra* note 61, at 1765.

subject. In other words, IoT devices could measure, to some extent, the probability of sensitivity, give it a numerical representation, and add noise to the IoT device according to such privacy risk assessment.

We turn to briefly explain how the value of  $\epsilon$  can be adjusted to properly balance data utility and privacy.<sup>219</sup> This Article suggests defining the value of  $\epsilon$  based on the probability of sensitivity. IoT devices with a high probability of gathering sensitive data—like Fitbit, which acquires sensitive health data—should add more noise to any gathered data. Meanwhile, IoT devices with a low probability of gathering sensitive data, like smart refrigerators, should add less noise—or none at all—depending on the privacy risk assessment. For example, an Amazon Echo in a household with children under the age of thirteen should add more noise to its data than an Amazon Echo in a household without children, all other things being equal. Essentially, any interaction with IoT technology will require an *ex ante* evaluation of the probability of sensitivity, followed by adding sufficient noise to the device's operation.

At this point it is essential to underline some caveats. First, our model is built on the current values embedded in the federal sectoral approach. It thus excludes, *inter alia*, state legislation that might also be relevant for privacy protection. It is also only natural that the perception of sensitivity of data change with technology and potential social changes. Thus, as previously mentioned, this model must be constantly challenged and recalibrated when necessary.

Second, our quantification of the level of data sensitivity could be viewed as somewhat arbitrary. In that regard, our intention is rather modest. We strive to show mostly how data probability could be assessed and used by differential privacy, but we make no binding statements regarding the actual values linked to data or context. These values could be challenged and changed by scholars or policymakers when necessary. Moreover, as previously mentioned, this model must also be adaptive, and obviously also include other types of sensitive data that must always be reevaluated in light of potential technological or social changes. Accordingly, any mechanism of probability is flexible and

---

219. Calibrating the privacy level or the added noise to the data is related to the value of the privacy parameter  $\epsilon$  defined in the context of differential privacy. Calibrating  $\epsilon$  depends on the specific application and type of data. When the data can be visualized, like GPS data for example, it may be computed interactively in a graphic way. Such a solution has been suggested, where the user visually sees how the data change in real-time while changing  $\epsilon$  via a slide-bar on a graphic user interface. When the data look sufficiently noisy, and the user sees that the secret data cannot be extracted intuitively, the current value of  $\epsilon$  is chosen. *See generally* Adi Akavia et al., *supra* note 211.

could be fine-tuned with time, especially if a specific form of technology is found able to collect more sensitive data than the model anticipated.

Third, if policymakers impose obligations on industries because of such an evaluation, the obligations might be detrimental if they are broader than intended or negatively affect the data's utility. This could lead to what is known as the principle of parsimony, meaning that taking broader action under uncertainty might have negative consequences.<sup>220</sup> The negative consequences of using differential privacy, however, are not worrisome. Concededly, some data might become less valuable for industries subject to overbroad regulation. But this potential drawback should be balanced against the benefits of using such a model. Thus, while sometimes differential privacy might be used under conditions of uncertainty, its impact on the quality of the data is not substantial. Essentially, it retraces the trade-off between privacy and utility.

Finally, our technological solution is likely to be combined with the modality of the law. Lacking external incentives for information protection, market actors' self-regulation is bound to fail.<sup>221</sup> There must be some form of incentive for companies to adhere to these requirements. This could be achieved, for example, by obliging private companies to implement these technological measures *ex ante* in order to begin operating (e.g., by requiring licenses) or *ex post* (by imposing high fines for noncompliance or data breaches). It could also be achieved by granting a safe harbor from liability lawsuits on the fulfillment of these standards, which will be treated as evidence of compliance *vis-à-vis* liability or even combining the modalities of social norms and the market to drive consumers to demand that these companies protect their privacy better.

With these caveats in mind, the main purpose of this Article is not to provide a definitive formula that will apply perfectly in every context, not to mention that is well near impossible to achieve. Its intention is to introduce a new mechanism that combines the notions of privacy perception with differential privacy, thus providing a relative form of privacy. This form of privacy is not only a practical means for privacy protection, but it also broadens the discussion on the use of technology to meet new challenges better. Without adhering to such methods, regulating privacy in the always-on era by the sectoral approach will defeat many of the purposes behind such forms of legislation, and will ultimately fail to properly protect individuals' privacy.

---

220. See Schwartz, *supra* note 74, at 923 (explaining principle of parsimony in context).

221. Kenneth A. Bamberger & Deirdre K. Mulligan, *Privacy on the Books and on the Ground*, 63 STAN. L. REV. 247, 258–59 (2011).

## V. CONCLUSION

Protecting privacy in an always-on era is very challenging. When individuals are constantly surrounded by devices that might capture their daily routine, conversations, location, imagery, and vital signs, they must have safeguards against misuse of these data. The sectoral approach does little to advance the rationales of protecting privacy in this age, so policymakers must further examine it. And policymakers should strive to embrace other regulatory mechanisms that would better protect sensitive data as sensors become more embedded in our lives. But as illustrated above, technological solutions must also be considered, as they might enhance privacy protection for individuals while preserving the value of data to a greater extent than the current regulatory approaches can. To accomplish this, OSPs might be obliged, or incentivized, to deploy mathematical solutions that will depend on an ex ante evaluation of the probability of data sensitivity.

Data sensitivity will also change with technology. As individuals make more use of IoT technology, including its potential embedment in the public infrastructure, we might divulge more data to both private companies and government agencies than ever before. Thus, any privacy model, including the one proposed in this Article, must be further examined and recalibrated to embed the values that society wishes to protect. For the time being, policymakers must consider requiring OSPs to implement innovative technological and mathematical solutions, such as the proposed framework, to address the profound privacy concerns that emerge from the always-on era.



# WHAT IS IT ABOUT LOCATION?

Kirsten Martin<sup>†</sup> & Helen Nissenbaum<sup>††</sup>

## TABLE OF CONTENTS

<b>I.</b>	<b>INTRODUCTION .....</b>	<b>253</b>
A.	BACKGROUND AND MOTIVATION .....	254
B.	OUTLINE.....	257
<b>II.</b>	<b>BACKGROUND AND RELATED WORK .....</b>	<b>258</b>
A.	TECHNOLOGY .....	259
B.	REGULATION .....	263
C.	COURTS .....	266
1.	<i>Who Collects Location Data Is Important .....</i>	<i>267</i>
2.	<i>How Location Data Is Collected Is Important.....</i>	<i>269</i>
3.	<i>What May Be Inferred on the Basis of the Location Data in Question ..</i>	<i>271</i>
D.	RELATED EMPIRICAL WORK .....	272
<b>III.</b>	<b>STUDY DESIGN .....</b>	<b>275</b>
A.	CONTEXTUAL INTEGRITY .....	275
B.	METHODOLOGY.....	277
1.	<i>Factorial Vignette Survey.....</i>	<i>278</i>
2.	<i>Respondent Controls.....</i>	<i>279</i>
a)	Privacy and Trust .....	279
b)	Authoritarianism .....	280
3.	<i>Analyzing Respondent-Level Variables .....</i>	<i>280</i>

---

DOI: <https://doi.org/10.15779/Z382F7JR6F>

© 2020 Kirsten Martin & Helen Nissenbaum.

<sup>†</sup> William P. and Hazel B. White Professor, University of Notre Dame's Mendoza College of Business.

<sup>††</sup> Professor of Information Science, Cornell Tech. The authors would like to thank the participants of the 2017 Privacy Law Scholars Conference for their helpful comments on an early study in this series. We are supremely grateful to colleagues who read earlier drafts and offered invaluable suggestions: Paul Ohm, Joel Reidenberg, Frederik Zuiderveen Borgesius, Deborah Estrin, Mainack Mondal, and Eran Toch provided critical insights, particularly into the technical correlates, in turn spurring ideas on the empirical and policy issues. We are grateful for support from the National Science Foundation under Grants No. 1311823, No. 1649415, CNS-1801501, and NSA H98230-18-D-006. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or NSA.

<b>IV. PILOT STUDY .....</b>	<b>281</b>
A. PILOT DESIGN .....	281
B. PILOT RESULTS.....	281
C. DISCUSSION OF PILOT STUDY .....	282
<b>V. MAIN STUDY.....</b>	<b>283</b>
A. MAIN STUDY DESIGN .....	283
1. <i>Vignette Factors</i> .....	283
2. <i>Vignette Template and Example for Main Study</i> .....	285
3. <i>Vignette Rating Task</i> .....	287
4. <i>Sample</i> .....	287
B. MAIN STUDY RESULTS .....	290
1. <i>Significance of Vignette Factors</i> .....	290
a) Actors.....	290
b) Duration .....	291
c) Source .....	292
d) Inferred Information .....	293
2. <i>Interactions</i> .....	294
a) Appropriateness of Source by Actor.....	294
b) Appropriate Duration by Actor.....	295
C. DISCUSSION OF MAIN STUDY .....	296
<b>VI. FOLLOW-UP STUDY .....</b>	<b>297</b>
A. FOLLOW-UP STUDY DESIGN .....	298
B. FOLLOW-UP STUDY RESULTS .....	298
1. <i>Average Rating Vignette Is “Okay”</i> .....	298
2. <i>Actors</i> .....	299
3. <i>Source</i> .....	300
C. FOLLOW-UP STUDY DISCUSSION.....	301
<b>VII. SIGNIFICANCE FOR TECHNOLOGY, REGULATION, AND LAW</b> .....	<b>301</b>
A. TECHNOLOGY .....	303
B. SIGNIFICANCE FOR REGULATION .....	304
C. SIGNIFICANCE FOR LEGAL DECISIONS .....	306
D. SIGNIFICANCE FOR HOW LOCATION IS LABELED IN SURVEYS AND LAW .....	307
<b>VIII. CONCLUSION.....</b>	<b>308</b>
<b>APPENDIX A – PILOT STUDY FOR SURVEY DESIGN .....</b>	<b>309</b>



A.	PILOT STUDY SURVEY DESIGN .....	310
1.	<i>Features Tested</i> .....	310
2.	<i>Vignette Factors in Pilot Study</i> .....	311
3.	<i>Vignette Template for Pilot Study</i> .....	312
4.	<i>Vignette Rating Task</i> .....	312
B.	PILOT RESULTS .....	312
1.	<i>Ordering of Controls and Vignettes</i> .....	312
2.	<i>Vignette Voice (“You” Versus “A Person”)</i> .....	313
3.	<i>Location</i> .....	313
4.	<i>Storage Versus Frequency of Data Collection</i> .....	314
5.	<i>Discussion of Pilot Survey</i> .....	315
	<b>APPENDIX B – FOLLOW-ON STUDY</b> .....	<b>316</b>
A.	FOLLOW-ON STUDY #1: ADDING PLACE TO A SURVEY ABOUT LOCATION .....	316
1.	<i>Average Rating Vignette Is “Okay”</i> .....	318
2.	<i>Actors</i> .....	318
3.	<i>Source</i> .....	319
B.	FOLLOW-ON STUDY #2: ADDING PLACE TO SURVEY WITH DURATION INCLUDED .....	320
1.	<i>Average Rating Vignette Is “Okay”</i> .....	320
2.	<i>Actor</i> .....	321
3.	<i>Source</i> .....	321
4.	<i>Duration</i> .....	322
	<b>APPENDIX C – QUALITY OF SAMPLES</b> .....	<b>323</b>

## I. INTRODUCTION

This Article reports on a set of empirical studies that reveal how people think about location data, how these conceptions relate to expectations of privacy, and consequently, what this might mean for law, regulation, and technological design. Despite the great debates, published commentary, court action, regulatory activity, and scholarly literature, not enough is known about how people understand location data, and what specifically about it affects people’s judgments about others’ access to their whereabouts.<sup>1</sup> Further, despite

---

1. See Matthew B. Kugler & Lior Jacob Strahilevitz, *Actual Expectations of Privacy, Fourth Amendment Doctrine, and the Mosaic Theory*, 2015 SUP. CT. REV. 205, 261–63 (2015) (calling for more empirical research on people’s perception of location data and the psychological basis of privacy expectations).

efforts to stem location tracking, it remains rampant. Stern rules<sup>2</sup> aimed at curtailing location tracking are a poor match for the ingenuity of seekers of this information who, among other tactics, exploit enormous ambiguity in how location is interpreted and operationalized to make end runs around these rules.<sup>3</sup>

Filling this gap is critical to a frontier of privacy regulation that has been sorely neglected. This neglect exists in part because the significance of location data was not fully appreciated until the recent ubiquity of technology-enabled location tracking, and in part because its murkiness has suited the beneficiaries of location surveillance. Although our findings alone do not support specific lines of legal regulation, they leave little doubt of a damaging rift between how these beneficiaries of location surveillance communicate their practices and how we, its subjects, understand these practices. Only when this rift is repaired will it be possible to adequately regulate location surveillance—through policy, law, and technology—to meet privacy expectations and promote privacy’s societal value.

#### A. BACKGROUND AND MOTIVATION

The set of empirical studies on which this Article reports is the third in a series, initiated in 2015, which challenges the role of the public-private dichotomy in privacy law and regulation by scrutinizing the extent to which

---

2. See, e.g., *Privacy, Security, and Deception*, GOOGLE PLAY DEVELOPER POL’Y CTR., <https://play.google.com/about/privacy-security-deception/> [https://perma.cc/DG49-XBPG] (last visited Dec. 30, 2019); *App Store Review Guidelines*, APPLE DEVELOPER, <https://developer.apple.com/app-store/review/guidelines/> [https://perma.cc/MZ47-J7P3] (last visited Dec. 30, 2019).

3. Several companies collect and monetize location data, including precise GPS coordinates, the name of Wi-Fi routers, and whether users have Bluetooth on or off. See, e.g., Michael Grothaus, *Google Tracks Your Movements Even if You’ve Turned Location History Off*, FAST COMPANY (Aug. 13, 2018), <https://www.fastcompany.com/90217689/google-tracks-your-movements-even-if-youve-turned-location-history-off> [https://perma.cc/4CNY-M2KZ]; Adrienne Jeffries, *Why Is This Company Tracking Where You Are on Thanksgiving?*, OUTLINE (Nov. 15, 2017, 9:50 AM), <https://theoutline.com/post/2490/why-is-this-company-tracking-where-you-are-on-thanksgiving> [https://perma.cc/4WN5-D3T4] (last visited Nov. 16, 2017); Taylor Hatmaker, *Users Dump AccuWeather iPhone App After Learning It Sends Location Data to a Third Party*, TECHCRUNCH (Aug. 22, 2017, 1:19 PM), <http://social.techcrunch.com/2017/08/22/accuweather-revealmobile-ios/> [https://perma.cc/8NLX-3RPA]; Robbie Gonzalez, *The “Thanksgiving Effect” and the Creepy Power of Phone Data*, WIRED (May 31, 2018, 2:29 PM), <https://www.wired.com/story/the-thanksgiving-effect-and-the-power-of-phone-data> [https://perma.cc/CMV3-JUQ8]; Frank Bajak, *Mobile Carriers Cut Off Flow of Location Data to Brokers*, AP NEWS (Jun. 19, 2018), <https://apnews.com/8582857aff8146f8ac81d247533b2177/APNewsBreak-Verizon-to-end-location-data-sales-to-brokers> [https://perma.cc/9Q5E-7SDV].

peoples' privacy expectations align with the dichotomy.<sup>4</sup> Contrary to received views,<sup>5</sup> we found that they do not align very well at all. Utilizing concepts from the theory of contextual integrity,<sup>6</sup> the first two sets of studies revealed that in the right circumstances (defined by social domains, recipients, and purposes), people are quite ready to share information deemed private with others. However, for information deemed public (so defined by its placement in public records), people maintain highly modulated privacy expectations.<sup>7</sup>

These studies extended over diverse categories of information types, but, quite early in their design, we set aside location, realizing that this category deserved special and separate attention. For one, location has had strong historical associations with both the private (e.g., one's home) and the public (e.g., the proverbial public square). For another, it has become a target of great interest and value as a raft of existing and emerging technologies have rendered location information accessible to an unprecedented degree. In so doing, these technologies and associated practices have muddied historical lines between public and private spaces, both by giving public exposure to that which was considered private, and also by revealing legitimate privacy interests in erstwhile public locations.

The focus of our studies here is the latter; that is, privacy interests in location data gleaned from spaces deemed public and historically not warranting legal or other forms of protection. Although novel capabilities eroding the sanctity of historically private spaces are deeply worrying,<sup>8</sup> the

---

4. For a fuller discussion of this point, see generally HELEN NISSENBAUM, *PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE* (2010), especially Chapters 4, 5, and 6.

5. The clearest articulation of the private-public dichotomy is in the plain view and third-party doctrines; or, as summarized by Monu Bedi, the Fourth Amendment Disclosure Doctrines, which equate making something available to be seen as, therefore, relinquishing privacy expectations. Monu Bedi, *The Fourth Amendment Disclosure Doctrines*, 26 WM. & MARY BILL RTS. J. 461, 461–63 (2017); see also Ian Kerr & Jena McGill, *Emanations, Snoop Dogs and Reasonable Expectations of Privacy*, 52 CRIM. L.Q. 392, 407–11 (2007); Orin S. Kerr, *The Fourth Amendment and New Technologies: Constitutional Myths and the Case for Caution*, 102 MICH. L. REV. 801 (2004); Orin S. Kerr, *The Case for the Third-Party Doctrine*, MICH. L. REV. 561, 566 (2009).

6. NISSENBAUM, *supra* note 4. For a definition of privacy as contextual integrity, see *infra* Section III.A. According to the theory of CI, whether privacy has been preserved or violated depends on whether a given flow of information (or data) is *appropriate*, which in turn depends on whether this flow conforms with entrenched and contextual informational norms (sometimes abbreviated as “privacy norms”).

7. Kirsten Martin & Helen Nissenbaum, *Privacy Interests in Public Records: An Empirical Investigation*, 31 HARV. J.L. & TECH. 111 (2017); Kirsten Martin & Helen Nissenbaum, *Measuring Privacy: An Empirical Test Using Context to Expose Confounding Variables*, 18 COLUM. SCI. & TECH. L. REV. 176 (2017).

8. See Kerr & McGill, *supra* note 5, at 393–94 (describing how bodily emanations like sweat and scents can be harnessed by new technologies for surveillance purposes); Kerr, *The*

erosion of freedom in spaces deemed public seems to defy standard labels. The way we see it, regulation (or absence of regulation), guided by a principle of laissez-faire or “up for grabs,” reflects intuitions based on the material capabilities of prior eras. Details aside, the so-called plain view or public disclosure doctrine is one such—a comfortable fit for traditionally-defined public spaces viewed through human eyes and recorded by notes on paper.<sup>9</sup> We should not be surprised, therefore, to discover that these ideas are desperately inadequate for public spaces of the present day—monitored by sophisticated systems of fixed and mobile networked sensors and recorded into computerized databases. Regulation that embodies intuitions and norms of past eras is bereft of concepts for handling present day privacy threats in historically public spaces, in turn handicapping courts and other regulatory efforts to identify, grasp, acknowledge, and protect against them. While people struggle to convey the nature of these wrongs, stakeholders continue to exploit this convenient lacuna.

Our studies offer insights into how people think about location data and the factors affecting how we evaluate common location-tracking practices. In so doing, these studies may serve the needs of courts, regulators, and system designers seeking to address diverse challenges without compromising the normative standing of privacy interests in location data. One important instance is the need to flesh out the meaning of “reasonable expectation of privacy” in the myriad of privacy cases that reach courts. Studies such as ours serve decision makers, including judges and regulators, who could benefit from robust empirical findings rather than intuition, hearsay, or anecdote as grounds for deciding whether practices in question either meet or do not meet societal expectations.<sup>10</sup> Likewise, social actors using and offering digital devices and

---

*Fourth Amendment and New Technologies*, *supra* note 5, at 865–66 (offering examples of how technological developments allow for increasing intrusion by law enforcement into private spaces).

9. See Bedi, *supra* note 5, at 470 (“[T]he public disclosure doctrine, which says that there is no privacy protection for a person’s movements in public.”); Kerr, *The Fourth Amendment and New Technologies*, *supra* note 5, at 827–28.

10. Professors Kugler and Strahilevitz nicely summarize why actual beliefs (as measured in surveys) are relevant to court opinions. Kugler & Strahilevitz, *supra* note 1, at 220 (“[W]e show how scientific polling can alleviate concerns that, in undertaking such an inquiry, judges will place undue weight on their own beliefs or on the beliefs of people in their social orbits.”). Around location data specifically, Kugler and Strahilevitz quote Justice Alito, who argued that reasonable expectations of privacy are “the average person’s expectations” or “popular expectations.” *Id.* at 207 (quoting *United States v. Jones*, 132 S. Ct. 945, 964 (Alito, J., concurring)); see also Christopher Slobogin & Joseph E. Schumacher, *Reasonable Expectations of Privacy and Autonomy in Fourth Amendment Cases: An Empirical Look at Understandings Recognized and Permitted by Society*, 42 DUKE L.J. 727 (1992).

services would do well to heed these findings in order to comply with them and avoid scandals of noncompliance when discovered.<sup>11</sup>

Despite the great debates, published commentary, court action, regulatory activity, and scholarly literature, not enough is known about how people understand location data, what specifically about location tracking affects their judgments of it, and what their expectations are regarding others' access to their whereabouts.<sup>12</sup> Given breakneck development of location tracking systems and the fundamental importance of a reasonable expectation standard in deciding legal and regulatory questions about privacy, answers to these questions are urgently needed. Our studies seek to fill some of the gaps in knowledge by focusing on location data and location tracking in public places. One of the most dramatic findings is that people's expectations of privacy are not correlated with the traditional dichotomy of private versus public. Moreover, privacy expectations in public spaces are far from haphazard but are tied systematically to factors that our studies reveal.

## B. OUTLINE

Part II of this Article provides a backdrop for our studies showcasing related work on privacy and location data. We have highlighted work on location privacy in technology design, regulation, and the courts that has particularly informed and influenced our own. We also explain how our studies extend past and contemporaneous empirical work on location and privacy.

Part III describes the design of our studies, including the factorial vignette methodology. It also outlines the theory of contextual integrity, which provides the framework structuring our survey instrument.

In Part IV, we describe a series of pilot studies that guided the design of the main survey and were critical in informing its structure, such as the study's 'voice' and the ordering of the questions. Results, some of which were quite surprising, shaped our main studies.

Part V describes our main study. This study presented a series of scenarios involving the capture and flow of location data to a nationally representative sample of 1,500 respondents. Respondents were asked to rank these scenarios in terms of how appropriate they judged the practices to be.<sup>13</sup>

---

11. See generally Ira S. Rubinstein & Nathaniel Good, *Privacy by Design: A Counterfactual Analysis of Google and Facebook Privacy Incidents*, 28 BERKELEY TECH. L.J. 1333 (2013) (analyzing the ways in which digital services are being designed to violate privacy); WOODROW HARTZOG, *PRIVACY'S BLUEPRINT: THE BATTLE TO CONTROL THE DESIGN OF NEW TECHNOLOGIES* (2018).

12. See Kugler & Strahilevitz, *supra* note 1.

13. These results are reported in the full Article to explain the study design.

Part VI takes up a question that emerged from findings in both Pilot and Main studies. It was clear that we needed to learn more about how respondents conceptualized location and how this affected their expectations of location privacy. To this end, we investigated different ways of describing location tracking, from merely numeric representations to semantically meaningful descriptions of *place*. To isolate the importance of adding place to vignettes describing a generic location, we ran two factorial vignette surveys: the first merely referenced *location*, and the second referenced a meaningful *place* (e.g., school, hospital).

In Part VII, we discuss the significance of the findings of all three studies for technology, regulation, and the courts. Our results immediately debunk the idea that people have no expectations of privacy in public.<sup>14</sup> The findings call common practices of amassing location data by government and commercial entities into question by showing that these practices flout expressed privacy expectations in systematic and specific ways.

The studies further reveal that *how* we ask about location in surveys makes a difference to how people react. Details such as duration of collection, place, and inferences drawn significantly affect respondent ratings. Strikingly, the respondents were far more attuned to location tracking when it revealed place (e.g., home, work, shopping) than GPS coordinates. By implication, regulating standard technical markers (e.g., GPS) representing location in technical systems may not assuage location privacy worries. Another surprising result is that the duration of location-tracking loses significance when inferences are drawn, which suggests that inference trumps duration and that concerns over duration may be proxies for more fundamental concerns over what can be inferred from longer-term location surveillance.

Finally, in line with our earlier studies, respondents consistently found most repugnant data capture and flow practices involving data aggregators or data brokers.

## II. BACKGROUND AND RELATED WORK

Our work has been prompted and shaped by much that has come before, including the developmental trajectories of technology, regulation, and court decisions. It has also drawn from related empirical work, which like ours has sought to understand the influence of diverse factors over privacy expectations concerning location. A caveat (for which we hope to be forgiven) is that in

---

14. See Helen Nissenbaum, *Protecting Privacy in an Information Age: The Problem of Privacy in Public*, 17 LAW & PHIL. 559 (1998) (examining why theories of privacy neglect or dismiss questions of privacy in public).

acknowledging influences from all four domains—technology, courts, regulation, and empirical studies—we have had to be selective in reviewing each of them.

#### A. TECHNOLOGY

This Section provides a selective survey of technologies that enable and facilitate the monitoring and tracking of individuals through space, with a focus on mobile devices or “smartphones.” A person’s whereabouts may be noted, tracked, and recorded by a variety of means, ranging from the plain sight of other people to technology-enabled image capture. The class of digital technologies that generate and record location data is broad and diverse, including fixed sensors that locate individuals within their ranges to mobile location sensors that people increasingly carry around with them. Such technologies span traditional CCTV systems to newer forms of networked cameras (still and video), license plate readers, RFID tags associated with traditional forms of identification (e.g., credit cards or passports), mobile phones, Internet-of-Things (IoT) devices, location-specific social media, and more. The emerging arena of urban tech—so-called “smart cities”—which, by definition, involves a myriad of system-integrated sensors interacting with physical bodies in motion as well as signals from mobile devices, introduces acute privacy challenges.<sup>15</sup> Few are more urgent than those associated with the capture of location data generated by individuals via innumerable transceivers “communicating” with an equally diverse range of transmitters from familiar mobile phones to novel, smart (driverless) vehicles.

This expanding array of location-generating and location-capture technologies requires a full reckoning outside the scope of this Article; however, a closer examination of one case, namely smartphones, helps to showcase at least one reason why location privacy has fallen into a mire of confusion. We further confine the examination under this heading to devices powered by Apple’s iOS and Google’s Android OS.<sup>16</sup> Without doing justice to all relevant developments, it is fair to say that since we began our studies of the determinants of privacy expectation roughly four years ago, advances in the scope and sophistication of consumer mobile technologies have been staggering.

For the two major competing mobile operating system (OS) platforms, numbers, one might say, are the tail that wags the dog. The more apps and app

---

15. *See, e.g.*, BEN GREEN, *THE SMART ENOUGH CITY: PUTTING TECHNOLOGY IN ITS PLACE TO RECLAIM OUR URBAN FUTURE* (2019).

16. The discussion of mobile privacy owes a huge debt to Mainack Mondal and Eran Toch, who should not, however, be blamed for any inaccuracies.

developers are attracted to respective operating systems, the greater the value to users and, so the argument goes, the greater the likelihood they will choose respective operating systems. To take one slim measure, the number of offerings in Apple's app store jumped from 800 in 2008 to 1 million in 2013.<sup>17</sup> And within the four-year timespan of our studies, the number jumped from 1.3 million in 2014 to 2.1 million by 2017.<sup>18</sup> With respect to the Android operating system, while slower to introduce third-party apps, Play Store offerings grew from 1.38 million in 2014 to 2.7 million by 2017.<sup>19</sup>

It is not surprising that, in reverse symbiosis, Apple and Google extend capacity and power to developer communities through Application Programming Toolkits (APIs) and Software Developer Kits (SDKs)<sup>20</sup> to capitalize on data naturally generated by their respective systems. For location, the OS provides not only GPS, but other markers such as position in relation to nearby Wi-Fi routers<sup>21</sup> and the closest cellular service towers. In addition to location markers, iOS and Android OSs are constantly updating, refining, and augmenting their offerings with a myriad of others (gyroscope, compass, identity verification, time, etc.) in service of the nearly 5 million total apps in the App Store and Play Store. Various uses of these developer kits and interfaces have stoked public outcry. For example, the popular Brightest Flashlight app was discovered to be tracking users' location and selling it to

---

17. Caroline McCarthy, *Apple: One Million iPhones Sold, 10 Million App Store Downloads in First Weekend*, CNET (July 15, 2008), <https://www.cnet.com/news/apple-one-million-iphones-sold-10-million-app-store-downloads-in-first-weekend/> [<https://perma.cc/AL6Z-6JAB>]; *App Store Sales Top \$10 Billion in 2013*, APPLE (Jan. 7, 2014), <https://www.apple.com/newsroom/2014/01/07App-Store-Sales-Top-10-Billion-in-2013/> [<https://perma.cc/B6T7Z-DQC2>].

18. Nick Summers, *The App Store Now Boasts 1.3 Million iOS Apps*, NEXT WEB (Sept. 9, 2014), <https://thenextweb.com/apple/2014/09/09/now-13million-apps-app-store/> [<https://perma.cc/U2A4-BX9R>]; Shannon Liao, *Apple's Total Number of Apps in the App Store Declined for the First Time Last Year*, VERGE (Apr. 5, 2018, 6:07 PM), <https://www.theverge.com/2018/4/5/17204074/apple-number-app-store-record-low-2017-developers-ios> [<https://perma.cc/ZJ6R-ZBLV>].

19. *Number of Android Applications*, APP BRAIN STATS (Oct. 5, 2014), <https://web.archive.org/web/20141006142446/https://www.appbrain.com/stats/number-of-android-apps> [<https://perma.cc/L5EA-GJ7G>]; *Number of Android Applications*, APP BRAIN STATS (Feb. 9, 2017), <https://web.archive.org/web/20170210051327/https://www.appbrain.com/stats/number-of-android-apps> [<https://perma.cc/ZB8D-SLCD>].

20. APIs and SDKs provide convenient programming interfaces that aid application developers in making their systems function within operating systems, such as mobile operating systems, or platforms, such as Facebook.

21. See, e.g., WIGLE.NET, <https://wigle.net/> [<https://perma.cc/9LZF-GLF6>] (last visited Dec. 30, 2019) (offering geolocated Wi-Fi network services).



third parties,<sup>22</sup> the Weather Channel was sued by the city attorney of Los Angeles for passing its users' location data to other IBM-owned services as well as outside entities,<sup>23</sup> and Accuweather stirred ire when investigators discovered that it was recording and selling location data even after users had said no.<sup>24</sup>

To rein in practices where app developers extract ostensibly unnecessary data, government regulators and OS providers have tightened policies for accessing various classes of information. Because of growing public distaste over stealth capture of device-generated data, regulators and OS providers are suggesting, and in some cases requiring, just-in-time, explicit requests for access to specific categories of data, with location data an important category among those singled out for special treatment.<sup>25</sup> Should we be satisfied that, with these explicit requests, websites, services, and mobile apps are finally doing right by their users? Can users be confident that their expressed preferences will, in fact, determine how location data is handled “in the machine” and beyond? Will their expectations be met?

In our view, the only correct answer to these questions is “we don’t know,” because the internal practices of OS providers, as well as the data flowing back and forth between the OS and app providers, remain opaque to the vast majority of users and to regulators. Only with considerable ingenuity have experts developed tools, such as Serge Egelman’s AppCensus, to ferret out some level of insight, far from complete.<sup>26</sup> But another reason, not previously recognized, why these questions are impossible to answer directly, is the conceptual ambiguity of location. In turn, this conceptual ambiguity poses challenges even to good faith efforts to regulate location tracking and to

---

22. Robert McMillan, *The Hidden Privacy Threat of...Flashlight Apps?*, WIRED (Oct. 20, 2014, 6:30 AM), <https://www.wired.com/2014/10/iphone-apps/> [<https://perma.cc/HQC4-GY8A>].

23. See Complaint for Injunctive Relief and Civil Penalties for Violations of the Unfair Competition Law, *People v. TWC Prod. and Tech., L.L.C.* (2019), <https://int.nyt.com/data/documenthelper/554-l-a-weather-app-location/8980fd9af72915412e31/optimized/full.pdf> [<https://perma.cc/57CT-RU5X>].

24. See Hatmaker, *supra* note 3.

25. Currently, there are twenty-eight such categories requiring special permissions, out of a total of ninety-one possible. *Permissions Overview*, ANDROID DEVELOPERS, INTERNET ARCHIVE, <https://web.archive.org/web/20190303040327/https://developer.android.com/guide/topics/permissions/overview> [<https://perma.cc/K7MU-PCFM>] (last visited Dec. 30, 2019).

26. See Irwin Reyes et al., “Won’t Somebody Think of the Children?” *Examining COPPA Compliance at Scale*, PROC. ON PRIVACY ENHANCING TECH., June 2018, at 63–83 (analyzing the privacy behavior of the mobile apps by “dynamic test,” which contains App Corpus, Analysis Environment, Event Extraction, etc.). The AppCensus search is available at <https://search.appcensus.io> [<https://perma.cc/6TSR-9VGY>] (last visited Dec. 30, 2019).

represent and enforce it in systems in concert with the ways people conceive, interpret, and value it. In other words, technical efforts to protect location privacy may stumble because of a failure to map the meaning that people assign to location with its representations in technical systems.

To illustrate the discrepancy between the meaning that people assign location with its representation in technical systems, let us return to the Accuweather scandal and consider a hypothetical explanation that gives Accuweather the benefit of the doubt. To begin, let's assume that Accuweather represented location in the system as coordinates derived from GPS. When users answered "no" to location tracking, Accuweather respected this expressed preference by ceasing to attach GPS coordinates to their respective records. Still wanting information about users' whereabouts, it sought alternative markers; in this instance, lookup tables from closest Wi-Fi routers. While users might be outraged by the workaround, Accuweather could counter that by ceasing to collect GPS signal, they were dropping location as it is normally represented in its system. Although we have not seen evidence of this precise dialog, the indignation registered in reports of this incident suggests that people are neither attuned to nor impressed by such distinctions. Our hypothetical account could continue. Even assuming that Accuweather has taken this criticism to heart and now eschews location markers drawn from GPS, Wi-Fi, and cellular towers, they have not exhausted all sources: in particular, semantic sources. Consider, for example, a user paying with anything but cash at a CVS branch on Bleecker Street, New York City. In this case, location is rendered semantically as "a CVS drugstore on Bleecker." Such data also may have been shared in the text message to a friend, "I am just finishing up at the CVS on Bleecker!" or tagged in a selfie posted on Instagram. Accuweather could hypothetically purchase such information from CVS, or one of the many location data brokers.

The point is that location can be characterized in many different ways, from GPS coordinates to semantically rich labels. Viewed in this light, one could conceive of the constellation of location-tracking mobile apps as a massive and distributed system for producing layer upon layer of meaning to numeric location coordinates. This system is akin to Geographic Information Systems, which attach meaningful labels to numerical geographical coordinates, but far more varied and potentially threatening. Similarly, meanings that apps attach to particular locations may be rich and complex, and potentially uncomfortably revealing. For example, in a familiar case, an app may identify a given set of coordinates as a person's "home" or "work." In more complicated instances, it may connect locations with app-labeled activities (e.g., "exercise" or "having sex") or even through co-presence with

other people (e.g., in social apps).<sup>27</sup> With greater sophistication, these systems may infer even more.

To put the conundrum plainly, when people respond “no” to location tracking, what is it that they believe, expect, and want to be happening? And whatever this is, does it map onto how systems developers represent and enforce this? For anyone committed to privacy-by-design or, more concretely, committed to ensuring that people’s location privacy expectations can be represented and enforced in technical systems, a sound mapping between those expectations and those systems is a necessary condition. The goal of such a mapping between technical representations and people’s privacy expectations is a key motivator of our work. As such, we have also sought to demonstrate where revealed expectations currently are asynchronous with efforts on the technical side.

## B. REGULATION

In this Section, we discuss the regulation of the location-tracking practices of commercial entities. The sources of this regulation are far less clear than the constitutional principles that apply to governmental actors, as discussed below in Section II.C.

As we know, the information technology and service industry functions under a model of “self-regulation,”<sup>28</sup> particularly in relation to privacy. Following concerns over the information practices of apps, the major mobile operating systems have issued sets of policies and guidelines for app developers.<sup>29</sup> Acknowledging deep anxiety over location, as noted above in Section II.A, they have become more demanding in requiring mobile app developers to provide finer grained notices about the data fields they seek to

---

27. See Jennifer Valentino-De Vries et al., *Your Apps Know Where You Were Last Night, and They’re Not Keeping It Secret*, N.Y. TIMES (Dec. 10, 2018), <https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html> [<https://perma.cc/R53-BSZT>].

28. Mary J. Culnan, *Protecting Privacy Online: Is Self-Regulation Working?*, 19 J. PUB. POL’Y & MARKETING 20, 20–26 (2000); FED. TRADE COMM’N, *PRIVACY ONLINE: FAIR INFORMATION PRACTICES IN THE ELECTRONIC MARKETPLACE* 1–7 (2000), <https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-fair-information-practices-electronic-marketplace-federal-trade-commission-report/privacy2000.pdf> [<https://perma.cc/SE88-33SK>]; Siona Listokin, *Industry Self-Regulation of Consumer Data Privacy and Security*, 32 J. MARSHALL J. INFO. TECH. & PRIVACY L. 15, 15 (2015).

29. See, e.g., *Developer Policy Center*, GOOGLE PLAY, <https://play.google.com/about/developer-content-policy/> [<https://perma.cc/S2B8-AKPU>] (last visited Dec. 30, 2019) (describing Android policies); *App Review*, APPLE, <https://developer.apple.com/app-store/review/> [<https://perma.cc/EXP8-VL3M>] (last visited Dec. 30, 2019) (describing iOS policies).

collect as well as finer grained choices for users, particularly as applied to location data.<sup>30</sup>

Although these policies and guidelines have somewhat constrained app developer access to user data generated by mobile devices,<sup>31</sup> by no means do they address the full scope of vulnerability to location tracking. First, quite obviously, location tracking is not limited to mobile apps; for example, fitness trackers may provide users with information about their runs by mapping and measuring their routes.<sup>32</sup> Second, the guidelines have still not stopped controversial practices that have raised eyebrows, if not vocal protest.<sup>33</sup> For example, having secured users' permission to monitor location data, companies may then provide this data to brokers.

One might argue that the status quo is not surprising, given the general backdrop of weak privacy regulation in the United States. Over the past decade, however, due to increasing pressure from advocacy organizations<sup>34</sup> and the public exposure of high-profile industry missteps,<sup>35</sup> the appetite for

30. See *Permissions Overview*, ANDROID DEVELOPER, *supra* note 25.

31. For example, from Google Play Developer Policy Center:

Limit your collection and use of this data to purposes directly related to providing and improving the features of the app (e.g. user anticipated functionality that is documented and promoted in the app's description).

Post a privacy policy in both the designated field in the Play Console and within the app itself. The privacy policy must, together with any in-app disclosures, comprehensively disclose how your app collects, uses, and shares user data. Your privacy policy must disclose the type of parties to which any personal or sensitive user data is shared.

Handle all personal or sensitive user data securely, including transmitting it using modern cryptography (for example, over HTTPS).

*Personal and Sensitive Information*, GOOGLE PLAY, [https://play.google.com/about/privacy-security-deception/#!?zippy\\_activeEl=personal-sensitive#personal-sensitive](https://play.google.com/about/privacy-security-deception/#!?zippy_activeEl=personal-sensitive#personal-sensitive) [<https://perma.cc/MXM8-JDXW>] (last visited Dec. 30, 2019).

32. See, e.g., Alex Hern, *Fitness Tracking App Strava Gives Away Location of Secret US Army Bases*, GUARDIAN (Jan. 28, 2018), <https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases> [<https://perma.cc/29CQ-SPZPJ>]; Liz Sly, *U.S. Soldiers Are Revealing Sensitive and Dangerous Information by Jogging*, WASH. POST (Jan. 29, 2018, 2:22 AM), [https://www.washingtonpost.com/world/a-map-showing-the-users-of-fitness-devices-lets-the-world-see-where-us-soldiers-are-and-what-they-are-doing/2018/01/28/86915662-0441-11e8-aa61-f3391373867e\\_story.html](https://www.washingtonpost.com/world/a-map-showing-the-users-of-fitness-devices-lets-the-world-see-where-us-soldiers-are-and-what-they-are-doing/2018/01/28/86915662-0441-11e8-aa61-f3391373867e_story.html) [<https://perma.cc/V3DH-ZNEQJ>].

33. See Valentino-DeVries et al., *supra* note 27.

34. See, e.g., ELECTRONIC PRIVACY INFORMATION CENTER, <https://epic.org> [<https://perma.cc/NT5T-E2WP>] (last visited Dec. 30, 2019); ELECTRONIC FRONTIER FOUNDATION, <https://www.eff.org> [<https://perma.cc/TD8A-WGDH>] (last visited Dec. 30, 2019).

35. See, e.g., Matt Warman, *Google: We Failed to Delete All Streetview Data*, TELEGRAPH (July 27, 2012), <https://www.telegraph.co.uk/technology/google/9432518/Google-we-failed-to-delete-all-Streetview-data.html> [<https://perma.cc/5NQA-42CU>]; Ritchie S. King & Mika

privacy regulation is slowly growing, with location privacy at the leading edge. A 2013 FTC Staff Report defined geolocation as “critical information” in need of greater regulation,<sup>36</sup> and location data was the focus of the Future of Privacy Forum’s “Mobile Location Analytics Code of Conduct.”<sup>37</sup> Yet even while warning that location data as generated by and garnered from mobile devices may be deeply revealing, these documents did not disrupt the reigning notice-and-choice model and merely offered “suggestions” and “recommendations” for how to communicate location data practices with greater salience, such as with “just-in-time” notices. Although this model allowed the FTC to issue a complaint against Goldenshores Technologies, LLC, maker of the “Brightest Flashlight” Android app, for misrepresenting its privacy practices,<sup>38</sup> it is impotent against accurate representations that are nevertheless incomplete and difficult to follow.

There is sufficient alarm over the insidious practices surrounding location data that it has gained the attention of lawmakers. Notably, in the European Union, the General Data Protection Regulation (GDPR), implemented in May 2018, singled out location data for special attention along with other types of data in the tightly regulated category of personally identifying information.<sup>39</sup>

---

Gröndahl, *How Google Collected Data from Wi-Fi Networks*, N.Y. TIMES (May 23, 2012), <https://archive.nytimes.com/www.nytimes.com/interactive/2012/05/23/business/How-Google-Collected-Data-From-Wi-Fi-Networks.html> [https://perma.cc/KP2P-NFYB]; Matthew Rosenberg et al., *How Trump Consultants Exploited the Facebook Data of Millions*, N.Y. TIMES (Mar. 17, 2018), <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html> [https://perma.cc/U67M-TN3K]; Matthew Rosenberg & Sheera Frenkel, *Facebook’s Role in Data Misuse Sets Off Storms on Two Continents*, N.Y. TIMES (Mar. 18, 2018), <https://www.nytimes.com/2018/03/18/us/cambridge-analytica-facebook-privacy-data.html> [https://perma.cc/4A7L-YSJM].

36. FED. TRADE COMM’N, MOBILE PRIVACY DISCLOSURES: BUILDING TRUST THROUGH TRANSPARENCY 3 (2013), <https://www.ftc.gov/sites/default/files/documents/reports/mobile-privacy-disclosures-building-trust-through-transparency-federal-trade-commission-staff-report/130201mobileprivacyreport.pdf> [https://perma.cc/JT6B-QB2A].

37. *Mobile Location Analytics Code of Conduct*, FUTURE PRIVACY F. (Oct. 22, 2013), <https://fpf.org/wp-content/uploads/10.22.13-FINAL-MLA-Code.pdf> [https://perma.cc/3CZE-JUNT].

38. *Android Flashlight App Developer Settles FTC Charges It Deceived Consumers*, FED. TRADE COMMISSION (Dec. 5, 2013), <https://www.ftc.gov/news-events/press-releases/2013/12/android-flashlight-app-developer-settles-ftc-charges-it-deceived> [https://perma.cc/ZAY9-6JYS].

39. In the GDPR, personal information includes “name, identification number, location data or online identifier . . . .” *Frequently Asked Questions about the GDPR*, EU GDPR PORTAL, <http://eugdpr.org/gdpr-faqs.html> [https://perma.cc/F27U-H2K6] (last visited Sept. 5, 2018); *Overview of the General Data Protection Regulation (GDPR)*, INFO. COMMISSIONER’S OFF. (2016), <https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/> [https://perma.cc/853C-GQTA] (last visited Dec. 30, 2016); ICO, *What is personal data?*, ICO’S GUIDE GDPR (2019), <https://ico.org.uk/for-organisations/guide-to-data-protection>

Specifically, processing identifiable information is regulated. To do so, data processors must meet one of a few criteria: the processing of the data must be necessary (1) to complete a contractual obligation, (2) to protect vital interests of the data subject or other person, (3) to perform a task in the public interest, (4) to comply with a law or regulation, or (5) “for the legitimate interests” pursued by the data controller or a third party.<sup>40</sup> These requirements would, for example, clearly and immediately rule out Brightest Flashlight.

Our assessment is that as the hardware, software, and political economy of data advance, the practices of location tracking are diverging from people’s expectations of appropriate behaviors. These discrepancies between expectations and common practices, despite efforts to regulate, suggest at least two possibilities, not necessarily mutually exclusive. First, the crafters of regulation, government and industry, are knowingly trading off privacy expectations and interests of data subjects in favor of location data collectors. Or, second, they do not properly grasp how people understand and value location data. Although our studies mainly shed light on the latter possibility, in so doing, they raise the stakes by revealing the nature and extent of the tradeoff.

### C. COURTS

In this Section, we consider how the courts have dealt with privacy and location data. Historically, the “third-party doctrine” has reflected the idea that individuals have no reasonable expectation of privacy in information they willingly give to others. But two landmark court cases have suggested that the doctrine is stretched thin in the face of location tracking technologies. First, in *United States v. Jones*, the Supreme Court held that police could not attach a GPS device to a defendant’s vehicle and track its movement for a period of twenty-eight days. While the majority focused on the trespass to property, Justice Sotomayor wrote in a concurring opinion that the third-party doctrine is “ill suited to the digital age, in which people reveal a great deal of information about themselves to third parties in the course of carrying out mundane tasks.”<sup>41</sup> In a second case, *Carpenter v. United States*, the Court held that a

---

/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/what-is-personal-data/ [https://perma.cc/5MAY-ASQJ] (last visited Dec. 30, 2019); *GDPR FAQs*, EUGDPR.ORG, <http://eugdpr.org/gdpr-faqs.html> [https://perma.cc/CML4-77DL] (last visited Sept. 5, 2018).

40. *Lawful Basis for Processing*, INFO. COMMISSIONER’S OFF., <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/> [https://perma.cc/N95S-GWNB] (last visited Dec. 30, 2018).

41. *United States v. Jones*, 565 U.S. 400, 417 (2012) (Sotomayor, J., concurring) (“[I]t may be necessary to reconsider the premise that an individual has no reasonable expectation

defendant had a reasonable expectation of privacy in his cell phone's location data, even though it was in the hands of his service provider, a third party.<sup>42</sup>

In the legal literature, much has been written about these two important cases and others involving location tracking.<sup>43</sup> Insofar as they relate to and influence our work, we have focused on factors that have systematically affected how courts have resolved questions about reasonable expectations of privacy in location, and how these factors have evolved over time. Guided by the theory of contextual integrity and characterizing location tracking practices as special cases of information flow, we considered how courts took the following features into consideration in determining whether practices involving the collection and uses of location data were legally acceptable: (1) who collects the data, (2) how it is gathered, and (3) the meaning that can be extracted from it.

### 1. *Who Collects Location Data Is Important*

An initial factor critical to determining whether reasonable expectations of privacy have been respected is *who* collects the data (or in contextual integrity terms, who receives the data). The courts have often differentiated between law enforcement versus private actors, with the former subject to rigorous constitutional constraints and the latter to far fewer.<sup>44</sup> The rise of commercial information intermediaries such as data brokers and credit agencies drove an active discussion in the courts and among legal scholars about the third-party doctrine.<sup>45</sup> This discussion focused specifically on the legal issues when intermediaries, with whom one has no reasonable expectations of privacy, provide information to government agencies, with whom one has a

---

of privacy in information voluntarily disclosed to third parties . . . . This approach is ill-suited to the digital age . . . .").

42. *Carpenter v. United States*, 138 S. Ct. 2206, 2222 (2017) (holding that individuals, in "rare case[s]," may have "a legitimate privacy interest in records held by a third party"); *see also* Bedi, *supra* note 5, at 486–88.

43. *See, e.g.*, Christopher Slobogin, *Making the Most of United States v. Jones in a Surveillance Society: A Statutory Implementation of Mosaic Theory*, 8 DUKE J. CONST. L. & PUB. POL'Y 1 (2012); Kevin S. Bankston & Ashkan Soltani, *Tiny Constables and the Cost of Surveillance: Making Cents Out of United States v. Jones*, 123 YALE L.J. 335 (2013); Paul Ohm, *The Many Revolutions of Carpenter* 32 HARV. J.L. & TECH. 357 (2019); Orin S. Kerr, *Initial Reactions to Carpenter v. United States* (USC Law Legal Studies Paper No. 18-14, 2018).

44. Kiel Brennan-Marquez, *Outsourced Law Enforcement*, 18 U. PA. J. CONST. L. 797, 797–99 (2016) (explaining that Fourth Amendment protections extend only to law enforcement seeking to gain information about citizens; commercial entities are able to surveil citizens at any time).

45. *See, e.g.*, Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 528 (2006); Susan Freiwald, *Cell Phone Location Data and the Fourth Amendment: A Question of Law, Not Fact*, 70 MD. L. REV. 681, 733 (2010).

constitutionally-based reasonable expectation of privacy.<sup>46</sup> One analogy supporting the third-party doctrine was likening private firms providing information to government actors to confidential informants, thus putting the onus on individuals such as clients, customers, and consumers for their misplaced confidences in untrustworthy actors or firms with whom they interact.<sup>47</sup> One problem with a focus on the actor as determinative of the norms of collecting and using location data is that law enforcement can then simply get the information from private parties.<sup>48</sup>

No matter what one's view on past cases, it would take willful avoidance to ignore epic transformations in the informational landscape. Writing about the burgeoning data broker industry, ranging from general brokers (such as Acxiom) to specialized providers (including some that focus on location data),<sup>49</sup> Chris Hoofnagle and others warn against private actors serving as government surrogates, calling them "Big Brother's Little Helpers."<sup>50</sup>

Another aspect of this transformation is the gradual elimination of choice in the transfer of data from individuals such as subscribers, consumers, and customers, to third parties, which are increasingly online, as a condition of a diverse array of services and transactions. This has led to a literature debating the idea of information intermediaries as fiduciaries.<sup>51</sup> Without pursuing this debate further, to us, significant progress will not be made that makes the actors in question—government or private—determinative of appropriate

---

46. See Kerr, *The Case for the Third-Party Doctrine*, *supra* note 5; Bedi, *supra* note 5.

47. In other words, the individual is at fault for sharing information with informants who, in turn, share that information with the government, whether a confidential informant in a criminal conspiracy or an untrustworthy firm with whom data is shared. See David C. Gray & Danielle Keats Citron, *The Right to Quantitative Privacy*, 98 MINN. L. REV. 62, 86 n.458 (2013) (describing the "misplaced trust rationale"); Kerr, *The Case for the Third-Party Doctrine*, *supra* note 5, at 568 (explaining that the Fourth Amendment does not protect defendants' misplaced confidence) (citing *Lewis v. United States*, 385 U.S. 206 (1966)).

48. Gray & Citron, *supra* note 47, at 140 ("If the government lacks legal authority to install and monitor a GPS-enabled tracking device, then it can get the same information by securing locational data from OnStar, Lojac, a cellular phone provider, or any number of 'apps' that gather and use locational information as part of their services.").

49. See Valentino-DeVries et al., *supra* note 27.

50. Chris Jay Hoofnagle, *Big Brother's Little Helpers: How Choice Point and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement*, 29 N.C. J. INT'L L. & COM. REG. 595 (2003); Michael D. Birnhack & Niva Elkin-Koren, *The Invisible Handshake: The Reemergence of the State in the Digital Environment*, 8 VA. J.L. & TECH. 1 (2003).

51. See, e.g., Kiel Robert Brennan-Marquez, *Fourth Amendment Fiduciaries*, 84 FORDHAM L. REV. 611 (2015); Neil M. Richards & Woodrow Hartzog, *Taking Trust Seriously in Privacy Law*, 19 STAN. TECH. L. REV. 431 (2016); Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C. DAVIS L. REV. 1183 (2016); Lina Khan & David Pozen, *A Skeptical View of Information Fiduciaries*, 133 HARV. L. REV. 497 (2019).



action; that is, action that conforms with reasonable privacy expectations. This point is elaborated in Part III below.

## 2. *How Location Data Is Collected Is Important*

Some legal scholars have focused on *how* location data is collected as determinative of the norms of data collection. David Gray and Danielle Citron focus on the investigative technique used to surveil the individual.<sup>52</sup> They argue that the technological advancements around indiscriminate data collection, aggregation, and storage remove the practical limitations on surveillance and, by this capability, run afoul of the traditional Fourth Amendment prohibition on dragnets.<sup>53</sup> Similarly, Margaret Hu shifts to a non-intrusion test to justify surveillance.<sup>54</sup> Hu focuses on big data technologies that facilitate horizontal cybersurveillance as a new technique.<sup>55</sup> Katherine J. Strandburg also argues that courts should apply a principle of technosocial continuity to respect privacy expectations of individuals.<sup>56</sup> The principle of technosocial continuity “requires that courts consider both the ways in which technology facilitates intrusive surveillance and the ways in which technology spurs social change that may make citizens more vulnerable to existing surveillance technologies.”<sup>57</sup>

Arguments to tie privacy expectations of location data to how the data is collected—if the technique is too invasive or pervasive, then privacy expectations are violated—closely align with Harry Surden’s theory of

---

52. Gray & Citron, *supra* note 47, at 102 (“Among the important factors that a court would need to consider are: (1) the inherent scope of a technology’s surveillance capabilities, be they narrow or broad; (2) the technology’s scale and scalability; and (3) the costs associated with deploying and using the technology.”).

53. *Id.* at 102.

54. Margaret Hu, *Cybersurveillance Intrusions and an Evolving Katz Privacy Test*, 55 AM. CRIM. L. REV. 127, 131 (2018) (“During oral argument in *Jones*, and in concurrences by Justices Alito and Sotomayor, the Court suggested that a nonintrusion test may be more appropriate given the scope of developing technology. A nonintrusion test is grounded in customary law, replacing an interpretation of the Fourth Amendment that is currently grounded in property and tort law, and presents a way to untether concepts of privacy from nondisclosure.”).

55. *Id.* at 361 (“Horizontal cybersurveillance makes possible what has been termed as ‘sentiment analysis.’ Sentiment analysis can be described as opinion mining and social movement forecasting. Through sentiment analysis, mass cybersurveillance technologies can be deployed to detect potential terrorism and state conflict, predict protest and civil unrest, and gauge the mood of populations and subpopulations. Horizontal cybersurveillance through sentiment analysis has the likely result of chilling expressive and associational freedoms, while at the same time risking mass data seizures and searches.”).

56. See Katherine J. Strandburg, *Home, Home on the Web and Other Fourth Amendment Implications of Technosocial Change*, 70 MD. L. REV. 614, 619 (2010).

57. See *id.*

structural privacy rights.<sup>58</sup> According to Surden, physical, societal, and technological constraints combine to make certain activities, including surveillance, difficult to complete without heavy costs; when one of these constraints is penetrated, we see our privacy as violated.<sup>59</sup> For Gray and Citron, technological advances serve to remove the structural constraints previously curtailing mass surveillance;<sup>60</sup> whereas obscurity, as defined by Professors Frederic Stutzman and Woodrow Hartzog, can be seen as adding to structural constraints.<sup>61</sup>

Along the line of location data collection and duration, Matthew Kugler and Lior Strahilevitz have examined if the duration of GPS data collection impacts people's reasonable expectations of privacy. Specifically, Kugler and Strahilevitz test the importance of duration in how the public regards the appropriateness of law enforcement needing a warrant to gather GPS data; they find it has no significant effect.<sup>62</sup> Importantly, these scholars frame the technology used to collect the data as critical to understanding whether privacy expectations are violated in the collection of location data.<sup>63</sup>

---

58. Harry Surden, *Structural Rights in Privacy*, 60 SMU L. REV. 1605 (2007).

59. *Id.*

60. Gray & Citron, *supra* note 47, at 63–67.

61. Woodrow Hartzog & Frederic Stutzman, *The Case for Online Obscurity*, 101 CALIF. L. REV. 1, 35–36 (2013) (“[W]e have identified four of these key factors: (1) search visibility, (2) unprotected access, (3) identification, and (4) clarity. The presence of these factors diminishes obscurity, and their absence enhances it. Thus, in determining whether information is obscure online, courts should consider whether any of these factors were present. Information that is entirely unobscure is completely obvious, and vice versa.”).

62. The authors ask a single question: Would it “violate people’s reasonable expectations of privacy if law enforcement” (1) used a car’s onboard GPS system to locate it on public streets without the owner’s permission? (2) used a car’s onboard GPS system to track its movements on public streets for one day without the owner’s permission? (3) same, but for one week? (4) same, but for one month? Kugler & Strahilevitz, *supra* note 1, at 246.

63. Rachel Levinson-Waldman argues that the following are important factors to consider in examining surveillance technologies:

(1) the duration of the surveillance; (2) the lowering of structural barriers to pervasive surveillance, reflected in the greatly reduced cost of tracking; (3) the recording of an individual’s or group’s movements; (4) the elicitation of information from within a protected space such as a home; and, as appropriate, (5) whether the technology undermines core constitutional rights and (6) whether surveillance technologies are piggy-backed on each other. Pulling out and articulating these factors, and analyzing how and why they should be considered, seeks to add rigor to the improvisatory method that has defined the judiciary’s consideration of these questions.

Rachel Levinson-Waldman, *Hiding in Plain Sight: A Fourth Amendment Framework for Analyzing Government Surveillance in Public*, 66 EMORY L.J. 527 (2016). The article goes on to examine various types of surveillance technologies (e.g. GPS, cellular phones, video cameras, drones, license plate readers, and body-worn cameras). *Id.*; see also Christopher Slobogin, *Making the*

### 3. *What May Be Inferred on the Basis of the Location Data in Question*

Pertinent to our work is the way courts have increasingly acknowledged the power of information technologies to transform information about one thing into another. Thus, in addition to how intrusive or pervasive are the *modes* of information collection, an important question is *what more can be inferred* from the information collected.

In other words, the methods for gathering information and the duration of the collection have historically been seen as a technological Peeping Tom peering into previously practically obscure spaces.<sup>64</sup> More recently, however, in both the Jones and Carpenter cases, location data over a period of time has been flagged for its capacity to generate new knowledge. The duration of the surveillance tells a new story about the individual, and individuals have a reasonable expectation of privacy in the whole of their movements.<sup>65</sup> Until now, “the Supreme Court has tended to pay more attention to the nature of the police intrusion required to obtain information than to the nature of the information obtained.”<sup>66</sup>

Paul Ohm takes up this shift from the duration of surveillance being a problematic technique to the duration of surveillance capturing new information and quotes the lower court in *Jones*: “[p]rolonged surveillance reveals types of information not revealed by short-term surveillance, such as what a person does repeatedly, what he does not do, and what he does ensemble. These types of information can each reveal more about a person than does any individual trip viewed in isolation.”<sup>67</sup>

For Ohm, the recent rulings validate the mosaic theory where the “accumulation of so many individual bits about a person’s life” results in a “personality picture that is worthy of conditional protection.”<sup>68</sup> Importantly,

---

*Most of United States v. Jones in a Surveillance Society: A Statutory Implementation of Mosaic Theory*, 8 DUKE J. CONST. L. & PUB. POL’Y 1 (2012) (concerning the duration of collection as important to understand privacy expectations around location data).

64. See, e.g., Levinson-Waldman, *supra* note 63, at 561–62 (arguing that duration could work as “a substantial intrusion on individuals’ privacy and diminish[] the obscurity that many people take for granted in their day-to-day movements . . . . The addition of technology has thereby both raised the stakes and lowered the barriers to intensive, intrusive surveillance”).

65. See *Carpenter v. United States*, 138 S. Ct. 2206, 2217 (2018) (“A majority of this Court has already recognized that individuals have a reasonable expectation of privacy in the whole of their physical movements.”) (quoting *United States v. Jones*, 400, 430 (Alito, J., concurring)).

66. Paul Ohm, *supra* note 43, at 362.

67. *Id.* at 373 (citing *United States v. Maynard*, 615 F.3d 544, 562 (D.C. Cir. 2010), *aff’d in part sub nom. United States v. Jones*, 565 U.S. 400 (2012)).

68. Slobogin, *supra* note 43, at 3–4; see also *United States v. Jones*, 565 U.S. 400, 416 (2012) (Sotomayor, J., concurring) (“I would ask whether people reasonably expect that their

this line of argument focuses on a new type of information that is revealed through the collection of location data as animating privacy concerns.<sup>69</sup>

#### D. RELATED EMPIRICAL WORK

Finally, we connect our studies with important instances of prior empirical work around privacy expectations and location data that has inspired and influenced it. We include work from the survey research literature examining privacy expectations primarily for purposes of influencing social science, law, and regulation. Further, we include empirical work in the user experience literature, primarily informing and targeting technology developers and designers, while aware that regulators are paying attention.

Previous work on location data has focused on the degree to which the method of collection (GPS tracker versus cell phone tower data) or duration of collection matters to reasonable expectations of privacy. The collecting agent is usually explicitly law enforcement. The closest attempt to measure privacy expectations surrounding the collection of location data centers on GPS location, law enforcement, and the duration of the collection.<sup>70</sup> In this study, Matthew Kugler and Lior Strahilevitz conducted a nationally representative survey to test the duration of location data collection that individuals judge as within their privacy expectations.<sup>71</sup> Their specific focus was on law enforcement. They tested whether duration (one day, one week, one month) impacted the degree to which use of “a car’s onboard GPS system to locate it on public streets without the owner’s permission” met privacy expectations.<sup>72</sup> The authors found that duration “barely affects” the degree to which the public regards geolocation tracking as invading their reasonable expectations of privacy.<sup>73</sup>

Alisa Smith, Sean Madden, and Robert Barton empirically examined how the method of government data collection impacted privacy, and found that respondents disapproved of government intrusion with aerial surveillance, a

---

movements will be recorded and aggregated in a manner that enables the Government to ascertain, more or less at will, their political and religious beliefs, sexual habits, and so on.”).

69. There exists a line of regulations focusing on types of information as requiring ‘special’ consideration including content versus metadata; medical; sensitive; financial, or intimate information. JULIE C. INNESS, *PRIVACY, INTIMACY, AND ISOLATION* (1996); Paul Ohm, *Sensitive Information*, 88 S. CAL. L. REV. (2015); Christopher Slobogin, *Cause to Believe What? The Importance of Defining a Search’s Object—Or, How the ABA Would Analyze the NSA Metadata Surveillance Program*, 66 OKLA. L. REV. 725 (2014).

70. Kugler & Strahilevitz, *supra* note 1, at 245–46.

71. *Id.*

72. *Id.* at 246.

73. *Id.* at 212.

GPS tracking device, or through cell phone towers.<sup>74</sup> Bernard Chao also compared reasonable expectations of privacy in different scenarios and observed that the highest proportion of respondents found the placement of a GPS device on a car for a duration of eighteen days to be a violation of reasonable expectations of privacy, as compared to seventeen other scenarios.<sup>75</sup> The question centered on the degree of intrusion of a government actor.<sup>76</sup> Similarly, Marc McAllister surveyed respondents with a series of questions involving location tracking through GPS devices versus cell phone tracking to compare the appropriateness of law enforcement surveillance as dependent on the seriousness of the crime.<sup>77</sup>

Outside law enforcement as the collecting agent, Jennifer Urban, Chris Hoofnagle, and Su Li found that “Americans overwhelmingly consider information stored on their phones to be private, and strongly reject systems that would rely on collecting and using contact data from their phones or tracking their locations.”<sup>78</sup> They found that 92% of respondents do not think their location data should be used for ads, and 46% say location should not be kept at all, even by cell phone companies.<sup>79</sup> Finally, Kirsten Martin and Katie Shilton compared location data to other data used for advertising, and found that the collection and use of location data for advertising negatively impacts privacy expectations in the mobile context, especially for high-use users.<sup>80</sup>

A series of studies has measured consumer behavior directly around location data to inform the tech industry. Eran Toch et al. employed a location

---

74. Alisa Smith et al., *An Empirical Examination of Societal Expectations of Privacy in the Digital Age of GPS, Cell Phone Towers, & Drones*, 26 ALB. L.J. SCI. & TECH. 26 111, 133–35 (2016). While Chao et al. dismiss these findings as not representative enough (Smith et al. have 54% women and 25% African American respondents), their own re-weighting in Chao et al. did not impact their results. Bernard Chao et al., *Why Courts Fail to Protect Privacy: Race, Age, Bias, and Technology*, 106 CALIF. L. REV. 263, 294, 297 (2018).

75. Chao et al., *supra* note 74, at 308–09. Chao’s examination of other forms of surveillance did not include duration. Among the seventeen other scenarios, accessing data stored in the cloud was second-highest, email was fifth, and roadblock was lowest. *Id.*

76. *Id.* at 303.

77. Marc McAllister, *GPS and Cell Phone Tracking: A Constitutional and Empirical Analysis*, 82 U. CINCINNATI L. REV. 207, 212 (2013). Kugler and Strahilevitz rightly identify the methodological issues and open questions of McAllister’s work, including no explanation of the sample. Kugler & Strahilevitz, *supra* note 1, at 223 n.113.

78. Jennifer M. Urban et al., *Mobile Phones and Privacy*, BERKELEY CTR. L. & TECH. 6 (2012), [https://www.ftc.gov/system/files/documents/public\\_comments/2013/12/00007-89101.pdf](https://www.ftc.gov/system/files/documents/public_comments/2013/12/00007-89101.pdf) [<https://perma.cc/A5HL-E7DC>].

79. *Id.* at 19, 20.

80. Kirsten Martin & Katie Shilton, *Why Experience Matters to Privacy: How Context-Based Experience Moderates Consumer Privacy Expectations for Mobile Applications*, 67 J. ASS’N INFO. SCI. & TECH. 1871, 1877–80 (2016).

sharing system to examine the actual behavior of study participants. The authors found that users were more willing to share location data when their location was frequented by a large and diverse set of people, thus suggesting a preference for areas where their identity would be obscured by others.<sup>81</sup> Michael Benisch et al. conducted a user study to measure when and where users would be willing to share their location data.<sup>82</sup> The authors found that day, time, and exact location are the significant factors driving users' willingness to share information rather than user activity, identity, or general concern as found in previous studies.<sup>83</sup> These findings suggest that users are quite nuanced about when and where they are willing to share their location data.<sup>84</sup>

There are three important gaps in the existing literature. First, location data has been operationalized in empirical studies as GPS without any explanation as to the types of inferences drawn about the user or the meaning of location data. Second, the majority of surveys have focused on law enforcement as the collecting actor, though the majority of location data collectors are actually private actors. Finally, the user studies have suggested that individuals have specific privacy expectations about how, when, and where location data should be gathered. Our study seeks to extend this important work on privacy by focusing on a diverse set of collecting actors and measuring the normative judgment of the respondents when the inferences drawn from location data are clear.

---

81. Eran Toch et al., *Empirical Models of Privacy in Location Sharing*, UBIComp '10 Proc. 12th ACM Int'l Conf. on Ubiquitous Computing, 129–138 (2010).

82. Michael Benisch et al., *Capturing Location-Privacy Preferences: Quantifying Accuracy and User-Burden Tradeoffs*, 15 PERS. & UBIQUITOUS COMPUTING 679, 679 (2011).

83. *Id.*

84. *Id.*; see also Adrienne Porter Felt et al., *I've Got 99 Problems, but Vibration Ain't One: A Survey of Smartphone Users' Concerns*, Proc. Second ACM Workshop on Security & Privacy in Smartphones & Mobile Devices 33 (2012) (finding that respondents differentiated their privacy expectations around location data based on who was receiving it; they were more concerned when friends, advertisers, or the public received it than when the server received it); Irwin Reyes et al., *supra* note 26, at 69–70 (discussing how apps targeted at children collected location data without consent); Primal Wijesekera et al., *Android Permissions Remystified: A Field Study on Contextual Integrity*, Proc. 24th USENIX Security Symp. 499, 508 (Aug. 12–14, 2015) (discussing situations where respondents did not find requests for location data from apps to be appropriate).

### III. STUDY DESIGN

#### A. CONTEXTUAL INTEGRITY

Our earlier work challenging the role of the private-public dichotomy revealed previously ignored factors that systematically affect people's privacy expectations. We called these confounding variables because they explained some of the inconsistencies between what people say and what they do, which commentators commonly—mistakenly in our view—call a “paradox.”<sup>85</sup> This work was guided by the theory of contextual integrity (CI), which pointed to variables that both refined and confounded the blunt categories of public and private. We have taken a similar approach in the present set of studies, in which we demonstrate that people's judgments about appropriate flows (in other words, their expectations) of location data are far more nuanced, in systematic ways, than the dichotomy would predict. Focusing solely on locations traditionally conceived as public, our study is able to hone in on what location means to respondents and the contextual parameters systematically affecting their judgments about location tracking and location data capture. Before proceeding, we offer a brief overview of CI, how it has guided our studies, and how, for pragmatic reasons, we have simplified it.

According to the theory of CI, whether privacy has been preserved or violated depends on whether a given flow of information (or data) is *appropriate*, which in turn depends on whether this flow conforms with entrenched and contextual informational norms (sometimes abbreviated as “privacy norms”).<sup>86</sup> When flows conform with entrenched norms, we say CI, *prima facie*, is respected. Otherwise, a further analysis is required in order to establish whether norms that have been contravened should override a practice under consideration or vice versa.

To establish conformance, a CI analysis needs to map actual flows against privacy norms (or expectations). Fully specifying a privacy norm requires specifying five key parameters: information type (about what), subject (about whom), sender (by whom), recipient (to whom), and transmission principle (flow under what conditions). Thus, when describing a given flow for purposes of evaluating its appropriateness, one needs to provide values for all five

---

85. Martin & Nissenbaum, *Measuring Privacy*, *supra* note 7, at 218; Martin & Nissenbaum, *Privacy Interests in Public Records*, *supra* note 7.

86. *Id.*; see also HELEN NISSENBAUM, *PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE* (2010). The part of CI theory that defines a series of steps to establish whether norms should prevail over conflicting practices, or vice versa, is concerned with moral legitimacy of norms or practices, respectively. Although answering questions about legitimacy is a defining component of CI theory itself, we set them aside for purposes of the current study.

parameters, or risk ambiguity resulting from missing variables.<sup>87</sup> An analysis that takes the public-private dichotomy as determinative would assert that reactions to flows of location data could be predicted solely on the basis of whether the location in question is public or private. By contrast, a CI analysis predicts a complex dependency between privacy expectations on the one hand, and the values for all five parameters on the other.

This thesis fundamentally informs the design of our studies. It also contrasts CI with some of the work discussed in Part II, notably efforts to decide cases or regulate data practices with reference to one factor alone (for example, the actor collecting information, the type of information, or the mode of collection) without recognizing that these factors interact. Although technical innovation has posed persistent challenges to institutional norms and structures, we ascribe painfully slow progress in coping with technology-induced privacy threats to an equally persistent failure to grapple with the interdependencies among key contextual factors. The studies reported in this Article (and the two previous articles), attempt to bring these interdependencies to light in the intersecting domains of law, policy, and technology.

Before describing our methodology and the studies themselves, two further points. First, we have not yet addressed the “context” in contextual integrity. The most we can say here, avoiding a long digression, is that context is roughly equivalent to social domain or sphere as theorized in social and political theory and reflected in the organization of societies (e.g., healthcare, family, commerce, finance, politics, etc.). Such domains are also frequently reflected in areas of law such as commercial law, family law, and constitutional law. Contexts in this sense are constituted by respective roles, activities, purposes, values, and norms. Among the norms, those governing information flows are associated with respective contexts in their characteristic ontologies, such as those defining contextual roles or capacities of actors (e.g., student, physician, senator, rabbi, etc.), and types or categories of information (e.g., diagnosis, blood type, vote, grades, marital status, criminal record, etc.). Accordingly, the scenarios we present to study respondents include values for parameters that are clearly associated with particular, familiar contexts (e.g., government, healthcare, etc.).

The second point is a caveat. Ideally, CI would require that the scenarios presented to respondents include the five parameters, with simultaneous variation of values for them. The reality of limited resources, time, human subjects, and requirements of statistical analysis has necessarily required

---

87. Martin & Nissenbaum, *Privacy Interests in Public Records*, *supra* note 7, at 123; Martin & Nissenbaum, *Measuring Privacy*, *supra* note 7, at 198.



pragmatic simplifications. These decisions were made with careful forethought and a disciplined effort not to claim more than the results allow.

#### B. METHODOLOGY

Our study comprises three key parts: (i) a set of pilot surveys which informed the design of the main study; (ii) a main study with a nationally representative sample to shed light on the attributes of location tracking instances that are systematically related to assessments of appropriateness of information flow (how “okay”); and (iii) a follow-up survey to assess the significance of location semantics relating how respondents’ understanding (conceptions) of location affect their judgments of the appropriateness of location tracking.

**Table 1: Overview of Studies**

Study	Sample	Goal
Pilot Study 1	Amazon Turk N = 1,200	Measure the impact of (1) the ordering of the control questions, (2) the voice of the vignettes, and (3) two parameters of the factorial vignette: (a) the precision of the location data described, and (b) the significance of frequency of tracking.
Main Study	Knowledge Networks N = 1,500	Understand what attributes of information flow are important to respecting contextual integrity in a public space. Survey 1. Actor, Source, Place Survey 2. Actor, Source, Place, Duration Survey 3. Actor, Source, Place, Duration, Inference.
Follow-Up Study	Amazon Turk N = 300	Explore how giving meaning to location data (including the place as understood from the location data) impacts consumers’ judgment.

In what follows, we outline general methods and our selection of respondent control ratings. To settle further design issues, we ran a pilot study which informed the factors we chose to include in subsequent surveys, the voice of the vignettes (2nd versus 3rd person), and the question order.

### 1. Factorial Vignette Survey

The method we used for our studies is known as the factorial vignette methodology.<sup>88</sup> Factorial vignette surveys present respondents with a series of vignettes in which multiple factors are systematically varied in order to test their relevance to respondents' assessments. These factors thus constitute the independent variables of our study. The variables chosen for our study correspond to a subset of the contextual factors (or parameters) of CI. For each vignette, values for the parameters are systematically and simultaneously varied. After seeing each vignette, respondents are asked to complete a simple rating task—the degree to which a scenario is appropriate or “okay”—from which we extract the statistical relevance of each of the factors.

The factorial vignette methodology has proven effective for addressing normative research questions which are notoriously difficult to study.<sup>89</sup> Because of the need to respond to several simultaneous contextual factors in the vignette, respondents are less likely to fall victim to two types of respondent bias. First, respondents may adjust answers in order to appear ethical or concerned in a traditional survey and are less likely to do so when many factors are changing simultaneously. This is particularly useful for privacy, which, according to skeptics, people claim to value while their behaviors communicate otherwise.<sup>90</sup> Second, respondents may have difficulty identifying and articulating the reasons behind their judgments, and the factorial vignette survey methodology supports the researcher in analyzing which factors moved the respondent's rating of the vignette without directly asking the respondent for a prioritized list of what is important to them in judging the vignette.<sup>91</sup>

---

88. Guillermina Jasso, *Factorial Survey Methods for Studying Beliefs and Judgments*, 34 SOC. METHODS & RES. 334, 342 (2006); Steven Nock & Thomas Guterbock, *Survey Experiments*, in HANDBOOK OF SURVEY RESEARCH (Peter V. Marsden & James D. Wright eds., 2010).

89. See, e.g., Jasso, *supra* note 88.

90. This is sometimes (mistakenly) referred to as the privacy paradox, where individuals are criticized for stating in surveys that they care about privacy while also sharing their data with companies. However, individuals are shown to not realize how their data is being tracked, shared, and used after disclosure, thereby rendering their behavior more closely aligned with their stated preferences. Individuals believe their privacy expectations are respected online and are shown to penalize companies when privacy expectations are violated. See Kirsten Martin, *Breaking the Privacy Paradox*, 32 BUS. ETHICS Q. 1 (forthcoming 2019); Kirsten Martin, *Privacy Notices as Tabula Rasa: An Empirical Investigation into How Complying with a Privacy Notice Is Related to Meeting Privacy Expectations Online*, 34 J. PUB. POL'Y & MARKETING 210, 220 (2015); Kirsten Martin, *The Penalty for Privacy Violations: How Privacy Violations Impact Trust Online*, 82 J. BUS. RES. 103, 110 (2018).

91. Martin & Nissenbaum, *Measuring Privacy*, *supra* note 7, at 195.

For our studies, vignettes described a scenario involving the collection, flow, or use of location data in public spaces, which respondents were asked to evaluate. Each respondent was presented with twenty to thirty vignettes, depending on the study. The survey instrument generates vignettes in real time by varying values randomly for each factor.

We asked respondents to rate the degree to which the vignette was “okay.” Choosing this language is part of our ongoing effort to elicit a sense of what is expected and what is normative. Although other studies of privacy might reasonably want to learn what people prefer, in taking guidance from CI, we strive to learn about people’s perception of norms. Nevertheless, more work is needed in defining an approach that encourages respondents to cast an objective eye.

## 2. Respondent Controls

Outside the vignettes, we also captured respondent-level controls based on previous privacy studies.<sup>92</sup> As before, we were interested in controlling for individual-level differences when the respondents answered a series of vignettes. Respondent-level beliefs and attributes that we selected (and discuss below) have all been shown to correlate with judgments about privacy and trust.

### a) Privacy and Trust

Privacy has been examined as impacting trust in prior studies and respondents’ general trust disposition has been found to impact their privacy concerns.<sup>93</sup> We captured the respondents’ disposition to trust by asking them to rate, on a scale from “strongly disagree” to “strongly agree,” their agreement with the statement: “In general, I trust people until proven otherwise.” We also captured the respondents’ institutional trust in government and business with the degree they agreed with, “In general, I trust the federal government,” and, “In general, I trust business.” Finally, we asked respondents to evaluate the statement, “In general, I believe privacy is important.”

---

92. See *id.*; Martin & Nissenbaum, *Privacy Interests in Public Records*, *supra* note 7.

93. Kirsten Martin, *The Penalty for Privacy Violations*, *supra* note 90, at 104. For a comparison of Westin’s privacy concern measurement to actual privacy expectations as well as individual’s trust disposition, see Martin & Nissenbaum, *Measuring Privacy*, *supra* note 7 (finding that respondents rated as low on Westin’s privacy concern measurement believed privacy to be important but trusted the firms and, therefore, had low concerns; and finding that Westin’s privacy concern measurement was not significant in driving specific judgments about privacy expectations).

b) Authoritarianism

In previous scholarship examining the privacy interests in public space and the privacy expectations around being tracked in public, Kugler and Strahilevitz found that respondents' affinity for authoritarianism impacted their expectations of privacy in regards to being tracked by the government in public.<sup>94</sup> To test the respondents' affinity for authoritarianism, an authoritarianism score was created from two questions based on existing scholarship: (a) "It's great that many young people today are prepared to defy authority" (reverse coded), and (b) "What our country needs most is discipline, with everyone following our leaders in unity."

3. *Analyzing Respondent-Level Variables*

Each control variable was captured using a slider with a scale of Strongly Disagree (-100) to Strongly Agree (+100). To standardize the responses, a new variable was created and assigned to each respondent as to what quartile their rating corresponded to (top 25%, bottom 25%, etc. of all ratings). This analysis was performed for each respondent control and used in the multi-level regressions as well as for splitting the sample when necessary.

Table 2: General Format of Surveys

Q #	Concept	Prompt
1	Trust in Business	In general, I trust business.
2–31	Vignettes (1 of 3 possible)	Please rate the degree to which this situation is okay, from Definitely Not Okay to Definitely Okay.
Respondent Controls: <i>How much do you agree or disagree with the following statement:</i>		
32	Privacy Important	In general, I find privacy important.
33	Trust in Government	In general, I trust the federal government.
34	RevAuthoritarianism 1	It's great that many young people today are prepared to defy authority.
35	Authoritarianism2	What our country needs most is discipline, with everyone following our leaders in unity.
36	Trust Disposition	In general, I give people the benefit of the doubt until shown otherwise.

94. Kugler & Strahilevitz, *supra* note 1, at 254–55.

#### IV. PILOT STUDY

##### A. PILOT DESIGN

In order to study what location data means to individuals, we needed to make decisions about terminology and study design. To this end, we ran a pilot study to test four facets of the survey design: (1) the ordering of control questions, (2) the voice of the factorial vignettes, and (3) two of the vignettes' parameters: (a) the importance of precision when presenting location data, and (b) significance of tracking frequency. Results of this pilot study, which were used to design the main surveys, are briefly described. A full description and analysis are provided in the Appendix.

##### B. PILOT RESULTS

1. **Ordering of Controls and Vignettes.** Did placing the controls before or after the vignettes matter to (i) the rating of the vignette or (ii) the respondents' ratings of the controls? To ensure the ordering did not impact the vignette ratings, we ran the pilot survey with the respondent controls both before and after asking the respondents to rate the vignettes. The average vignette rating did not change when the control questions were asked before versus after the vignettes. The average rating remained about -36 ("Not Okay"). Interestingly, the ratings for certain control variables did change when the controls were asked after the vignettes, as shown in Table A2 in the Appendix.

Specifically,

- The Authoritarian score decreased from -13.32 to -20.58 when the question is asked after the vignettes. In other words, the respondents are less authoritarian after rating scenarios about commercial and governmental tracking.
  - The average trust in business rating also decreases from -12.12 to -25.95 when the question is asked after the vignettes are rated. This is consistent with previous work on trust and privacy: respondents' institutional trust in business in general is diminished when the gathering and use of data is just explained in vignettes.<sup>95</sup>
2. **Vignette Voice** ("you" versus "a person"). We tested if the 'voice' of the vignette mattered to the judgment of whether the information flow was appropriate. The voice of a second person, third person, or third person plural impacted the privacy judgments of the respondents, as

---

95. Kirsten Martin, *Do Privacy Notices Matter? Comparing the Impact of Violating Formal Privacy Notices and Informal Privacy Norms on Consumer Trust Online*, 45 J. LEGAL STUD. 191, 206 (2016).

has been suggested before.<sup>96</sup> Voice did make a difference. When the vignettes included a reference to the respondent (“you”), the vignettes were rated less “okay” (-35.32) compared to a third-person voice (-27.05) or a third-person plural voice (-30.45). We decided to use third person voice for the live survey.

3. **Location Precision.** This may have been the most surprising finding. We were interested whether precision mattered, ranging from GPS (most precise) to location, street address, and city. The results suggest that the word “location” meant the same to respondents as “GPS” in judging the scenario as appropriate, with no significant difference between the two levels ( $p=0.95$ ). And even where the precision decreased, such as street address (+5.69) and city (+8.19), the degree of difference was only slightly over GPS and generic location. ( $p < 0.00$ ).
4. **Storage versus Frequency of Data Collection.** In order to test if the frequency of the data collection or its storage duration affected subjects’ responses, we included both factors in the vignette. The length of storage time was found to be inversely related to how “okay” the vignette was judged, as indicated by the steep negative slope in Figure 2 in the Appendix. Frequency, by contrast, was not significant; respondents did not rate the vignette any differently as the frequency levels varied.<sup>97</sup>

#### C. DISCUSSION OF PILOT STUDY

The results of the pilot study were surprising and essential in guiding aspects of the design of our main study. In sum:

1. We used the term “location” in the later studies, knowing that the term is equivalent to “GPS” for the respondent;
2. We dropped the use of frequency;
3. We shifted to the term “duration” for the duration of tracked location information;
4. We used the third-person plural in the later vignettes and asked the control questions after the vignettes in order to break up the control questions.

---

96. See Slobogin & Schumacher, *supra* note 10, at 736.

97. Because this result was somewhat surprising, we ran another vignette survey without storage included as a factor to allow the respondent to focus on frequency (from every five seconds to once per day). However, frequency was still not significant; the only difference was the average vignette rating decreased from -35.52 to -31.57 when storage was removed as a factor.

We discuss these decisions further in the Appendix.

## V. MAIN STUDY

Having settled some of these design issues, the purpose of the main vignette study is to identify what contextual factors are important to respondents' judgments of whether location data collected in public were appropriate. The study focused on the factors described below and shown in Table 3:

- Transmission Principles
  - Source: How the location data is gathered (phone signal, mapping app, license-plate reader, etc.)
  - Duration: How long the location data is gathered (from a few minutes to a year)
- Actors: Recipients of the location data (FBI, family, your employer, etc.)
- Attributes: What information can be inferred from the location data (who your friends are, how regularly you vote, etc.)<sup>98</sup>

### A. MAIN STUDY DESIGN

#### 1. *Vignette Factors*

- a. Source. How the location data is gathered and transmitted has been found to be important.<sup>99</sup> The source of collecting the location data varied across license plate readers, CCTV, phone tracking, social media, or a mapping application. Since sources affect the conditions or constraints of flow from subject to recipient, we took these to be operationalizations of Transmission Principles, as defined in CI.

---

98. Based on the design pilot, we used third-person voice in each scenario with the word location, which is equivalent to the term "GPS coordinates" for the respondents, given the pilot study described above. We had the respondents rate the vignettes before answering the control questions. We used duration rather than storage or frequency. *See infra* Appendix A.

99. *See generally* Surden, *supra* note 58; Luciano Floridi, *Network Ethics: Information and Business Ethics in a Networked Society*, 90 J. BUS. ETHICS 649 (2009); Kirsten Martin, *TMI (Too Much Information)*, 30 BUS. & PROF. ETHICS J. 1 (2011); Woodrow Hartzog & Evan Selinger, *Surveillance as Loss of Obscurity*, 72 WASH. & LEE L. REV. 1343 (2015); Lawrence Lessig, *The Architecture of Privacy*, 1 VAND. J. ENT. L. & PRAC. 56 (1999).

Table 3: Vignette Factors Included in National Study

Concept	Description	As operationalized in Vignette
<b>Duration</b>	How long you are tracked	A year, about six months, a month, a few days, a few minutes
<b>Actor</b>	Government	A city emergency service (ambulance, fire)
	Federal government	The FBI
	Employer	Their employer
	Commercial data aggregator	A commercial data broker
	Commercial	A commercial location-based service (e.g., Yelp)
<b>Source</b>	Family	A family member (e.g., parents, spouse, or sibling)
	License-plate reader	License-plate readers
	CCTV	CCTV cameras with facial recognition
	Phone	The signal from a mobile phone
	Fit Bit	A fitness app (e.g., FitBit or Strava)
	Social media	Geo-tagged posts on social media (e.g., Twitter, Facebook, Instagram)
<b>Additional Factor</b>	Mapping app	A mapping app (e.g., Google Maps)
	<b>Place</b>	<b><u>Inferences about individual (Survey 3 Only)</u></b>
	Association	A restaurant or cafe
	Protests/rallies	Who their friends are
		Whether they are active in their political beliefs in attending protests
	Sin Shopping	A liquor store
		Whether they have a drinking problem
	Shopping	A shoe store
		How susceptible they are to shoe ads
	Home	Home
		How often they spend the night away from home
	Work	Work
		Whether they are dedicated workers
	Medical	A medical clinic
		Whether they have a chronic illness
	Voting	A voting site
		How regularly they vote



- b. Duration of Collection. Previous work has found that the duration of data collection can affect privacy expectations.<sup>100</sup> We had the data collation range from a period of a few minutes to a year.
- c. Actors. In order to capture both government and commercial actors as well as different purposes of the data collection, the values of the actor (recipient) parameters ranged over FBI, a city planner, a commercial data broker, a location-based commercial service, and family members.
- d. Place and Inferences. We added this factor into Survey 2, which is explained below in order to understand the extent to which “bare” location was a stand-in or proxy for other qualitative locational information. Inferred information included a person’s associates, whether attending a protest, voting behavior, routine travel, whether frequenting a store, and whether frequenting a medical facility, in addition, simply, to where a person is. This tests whether the attribute of type of information inferred about a person drives expectations surrounding location information.

## 2. *Vignette Template and Example for Main Study*

The factors in Table 3 are used within a vignette template as described below. A specific level within each factor is randomly assigned as the vignette is generated for the respondent. Below the example vignettes for all three surveys are provided, as well as the general template for each.

### a. Survey 1 Template Baseline

{Actor} acquires location data from {Source} and uses this data to figure out if a person was at {Place}.

### b. Survey 1 Examples

A city emergency service (ambulance, fire) acquires location data from license-plate readers and uses the data to figure out if a person was at a shoe store.

---

100. *See, e.g.,* United States v. Jones, 565 U.S. 400 (2012) (finding that the data was collected for twenty-eight days); Shafer v. City of Boulder, 896 F. Supp. 2d 915 (D. Nev. 2012) (finding data was collected for two months); Carpenter v. United States, 138 S. Ct. 2206 (2018) (finding that data was collected for about 127 days); County of Riverside v. McLaughlin, 500 U.S. 44 (1991) (finding that data was collected for twenty-four hours).

A city emergency service (ambulance, fire) acquires location data from a mapping app (e.g., Google Maps) and uses the data to figure out if a person was at a restaurant.

A commercial location-based service (e.g., Yelp) acquires location data from license-plate readers and uses the data to figure out if a person was at a restaurant.

An employer acquires location data from geo-tagged posts on social media (e.g., Twitter, Facebook, or Instagram) and uses the data to figure out if a person was at home.

c. Survey 2 Template – Adding Duration

{Actor} acquires location data from {Source} for a period of {Duration} and uses this data to figure out if a person was at {Place}.

d. Survey 2 Examples

A family member (e.g., parents, spouse, or sibling) acquires location data from license-plate readers for a period of a year and uses the data to figure out if a person was at the National Mall.

An employer acquires location data from a mapping app (e.g., Google Maps) for a period of a few minutes and uses the data to figure out if a person was at a shoe store.

A commercial data aggregator acquires location data from license-plate readers for a period of a week or so and uses the data to figure out if a person was at a shoe store.

e. Survey 3 Template – Adding Inference

{Actor} acquires location data from {Source} for a period of {Duration} and uses this data to figure out if a person was at {Place} and {Inference}.

f. Survey 3 Examples

An employer acquires location data from a mapping app (e.g., Google Maps) and uses the data to figure out if a person was at a liquor store and whether they have a drinking problem.

A city emergency service (ambulance, fire) acquires location data from geo-tagged posts on social media (e.g., Twitter, Facebook,

Instagram) and uses the data to figure out if a person was at a liquor store and whether they have a drinking problem.

A commercial data aggregator acquires location data from a fitness app (e.g., FitBit or Stava) and uses the data to figure out if a person was at a shoe store and how susceptible they are to shoe ads.

### 3. *Vignette Rating Task*

For each vignette, respondents were instructed to indicate the degree to which they agreed with the question “Is this okay?” with a slider. The left side of the slider indicated “Definitely Not Okay” and the right of the slider indicated “Definitely Okay.” The slider was on a scale of -100 to +100 with the number suppressed so the respondents saw only the labels “Okay” and “Not Okay.”

### 4. *Sample*

In our previous studies, we utilized Amazon’s Mechanical Turk, which has become an accepted platform for empirical research such as this. Amazon Mechanical Turk offers a platform for researchers to post surveys (HITs) and respondents or workers to perform HITs they find worthwhile or interesting. Mindful of questions around this choice (not only aimed at our work), for our main survey, we deployed KnowledgeNetworks, an online research panel representative of the entire U.S. population. Approximately 1,500 respondents took one of three possible vignette surveys. KnowledgeNetworks panel members are randomly recruited through probability-based sampling. Households are provided with access to the internet and hardware if needed.<sup>101</sup> Importantly, Amazon Mechanical Turk provided higher quality sample with the same theoretical generalizability as KnowledgeNetworks.

---

101. For an overview of the KnowledgeNetworks sampling methodology and a comparison to the pilot tests on Turk, see *infra* Appendix C.

Table 4: Sample Statistics for Surveys 1–3

	Survey 1	Survey 2	Survey 3
	Base	+ Duration	+Inferred Info
Authoritarian Scale	5.57	6.83	2.76
Trust Scale	5.44	5.94	457
Female	49%	53%	50%
Age	50.1	49.5	49.4
Privacy Important	72.32	70.97	72.14
Trust Government	-23.08	-22.93	-27.67
Trust Business	2.85	2.53	4.10
DV Mean	-28.66	-35.96	-46.16
N (Respondents)	480	483	435

The sample was analyzed for unresponsive respondents. Since the respondents each rated thirty independently generated vignettes, the pattern of their rating on a sliding scale of -100 to +100 for each vignette could be analyzed as possibly unresponsive. We marked two types of surveys as nonresponsive: those that rated over twenty of the thirty vignettes as “0” (never moved the slider) and those that rated over twenty-five vignettes at one of the end points (moved the slider to the left or the right almost every time). For the KnowledgeNetworks sample, this resulted in 10% of Survey 1 respondents, 13% of Survey 2 respondents, and 16% of Survey 3 respondents being removed from the pool. The number of respondents listed in Table 4 above does not include those respondents removed from the analysis.<sup>102</sup>

102. Appendix C includes a comparison of the KnowledgeNetworks sample with the sample from running the same surveys on Amazon Mechanical Turk. The number of respondents discarded from non-responsive ratings was less for Turk than the national sample from KnowledgeNetworks. For the Turk sample, 2% of Survey 1 respondents, 5% of Survey 2 respondents, and 11% of Survey 3 respondents were found to be unresponsive. In comparison for the KnowledgeNetworks sample, 10% of Survey 1 respondents, 13% of Survey 2 respondents, and 16% of Survey 3 were unresponsive and removed from the sample. The Turk sample was higher quality than the KnowledgeNetworks sample with the same theoretically generalizable findings.

Table 5: Main Regression of Okay Rating on Vignette Factors and Respondent Controls

	Survey 1		Survey 2		Survey 3	
	BASE		DURATION		INFERENCE	
	<u>Coef</u>	<u>p</u>	<u>Coef</u>	<u>p</u>	<u>Coef</u>	<u>p</u>
FedGovtActor	45.55	0.00	33.22	0.00	10.08	0.00
DataAggregatorActor	0.77	0.61	0.03	0.98	-0.88	0.48
FamilyActor	33.70	0.00	23.04	0.00	19.08	0.00
EmployerActor	-3.00	0.05	-7.52	0.00	-1.28	0.31
CityServicesActor	42.39	0.00	16.74	0.00	4.42	0.00
<i>(null = Commercial Actor)</i>						
MappingAppSource	-1.53	0.31	-3.97	0.00	-1.01	0.41
PhoneSource	-0.19	0.90	-7.26	0.00	-0.73	0.56
LPRSource	-6.47	0.00	-7.40	0.00	-3.38	0.01
CCTVSource	-4.01	0.01	-5.11	0.00	-2.84	0.03
FitBitSource	-7.90	0.00	-10.35	0.00	-3.92	0.00
<i>(null = Social Media Source)</i>						
MedicalPlace	0.43	0.80	-1.01	0.52	0.58	0.68
RalliesPlace	3.34	0.06	1.45	0.36	-6.10	0.00
ShoppingPlace	-2.57	0.14	-4.39	0.01	-0.22	0.88
VotingPlace	-13.85	0.00	-12.74	0.00	-7.22	0.00
SinShoppingPlace	-3.09	0.08	-6.38	0.00	1.11	0.45
HomePlace	5.84	0.00	-0.82	0.61	-1.44	0.33
WorkPlace	4.78	0.01	0.84	0.60	2.07	0.16
<i>(null = Restaurant)</i>						
DurationScale	n/a	n/a	-1.64	0.00	-0.49	0.06
PrivacyImport	-0.27	0.00	-0.38	0.00	-0.26	0.00
HighAuthoritarianism	4.28	0.30	4.67	0.23	2.94	0.52
TrustScale	0.42	0.00	0.44	0.00	0.23	0.00
_cons	-32.04	0.00	-12.37	0.00	-31.14	0.00
N	480		483		435	
Vignettes	14,400		14,490		13,050	
DV Mean	-28.66		-35.96		-46.16	
ICC	32.7%		35.4%		45.2%	
ICC Null	33.6%		39.1%		48.0%	

## B. MAIN STUDY RESULTS

To analyze which vignette factors are significant to respondents' judgments about the appropriateness of the gathering and use of location data, we regressed the dependent variable—the rating that the collection of location data in the given vignette was “Okay”—on the vignette factors and the respondent controls. The results are in Table 5. The factor with the most impact on the rating task is the actor collecting the location data; changing who gathers the information had the largest impact on the rating that the gathering of location data was “Okay.” Below, each vignette factor—actor, source, and inference—is analyzed.

For the respondent controls, we found that authoritarianism was not significant to the rating task compared with the general trust scale, which was: the greater the respondents' trust in general (a composite of dispositional trust, trust in business, and trust in government), the more appropriate the respondent judged the collection of location data overall.

### 1. *Significance of Vignette Factors*

#### a) Actors

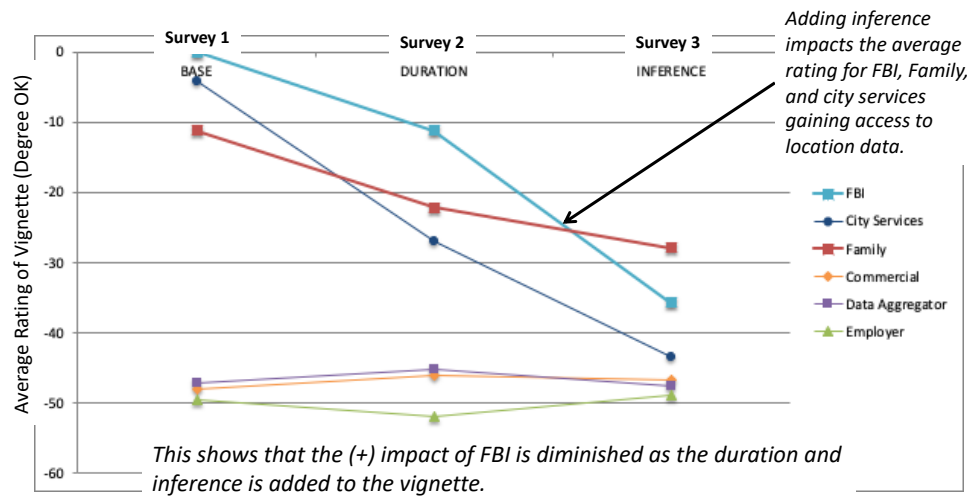
The actor acquiring the location data significantly affected respondents' judgments. As shown in the regression results in Figure 1, for each actor—FBI, commercial location-based service, city planner, and data—it was significantly less appropriate than the null condition for a family member to acquire location data.

Figure 1 also shows that adding inference impacts the average rating for FBI, family, and city services gaining access to location data. Figure 1 further reflects a significantly more negative rating of FBI, city services, and family collecting location data when the vignettes reveal the duration of the collection (Survey 2), and the nature of what the actor can infer about the individual (Survey 3). Even initially, the positive glow surrounding the FBI is extinguished when duration and inference are included in vignettes.

Table 5, with the main regression results, shows that initial differences in ratings between the FBI or city services versus a commercial entity (e.g., Yelp) diminishes as duration and inference are included. The FBI is favored above a commercial actor a when place only is included (+47;  $Ave_{FBI} = -0.04$ ,  $Ave_{Bus} = -47.14$ ). But when duration is added (+35;  $Ave_{FBI} = -11.20$ ,  $Ave_{Bus} = -46.09$ ) and inferences are drawn, the difference is diminished (+11;  $Ave_{FBI} = -35.67$ ,  $Ave_{Bus} = -46.80$ ). While the collection and use of location data by commercial actors such as Yelp or data aggregators is consistently not “okay,” the appropriateness of the FBI collecting location data is negatively impacted by

the mere mention of duration and the mere mention of the inferences drawn about the individual surveilled.

Figure 1: Average Vignette Rating by Actor

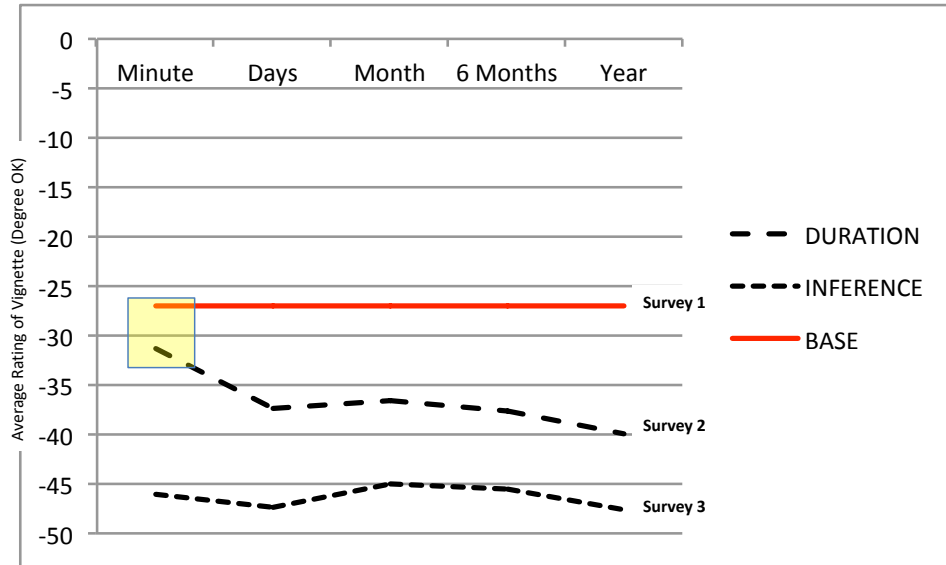


#### b) Duration

The duration of the tracking of location data was significant with -5.16 lower vignette rating (less “okay”) for each incremental step in additional time of tracking as shown in Figure 2. The impact of duration is lessened, (i.e., the slope is shallower) in Figure 2, for Survey 3 where the inferred information is also included.

Importantly, respondents appear to assume the shortest duration when no duration is included in the vignette, as in the base scenario in Survey 1. The average rating for a vignette with the duration set to “a few minutes” is the same as the baseline when no duration mentioned (see the yellow box in Figure 2).

Figure 2: Average Vignette Rating by Duration Period



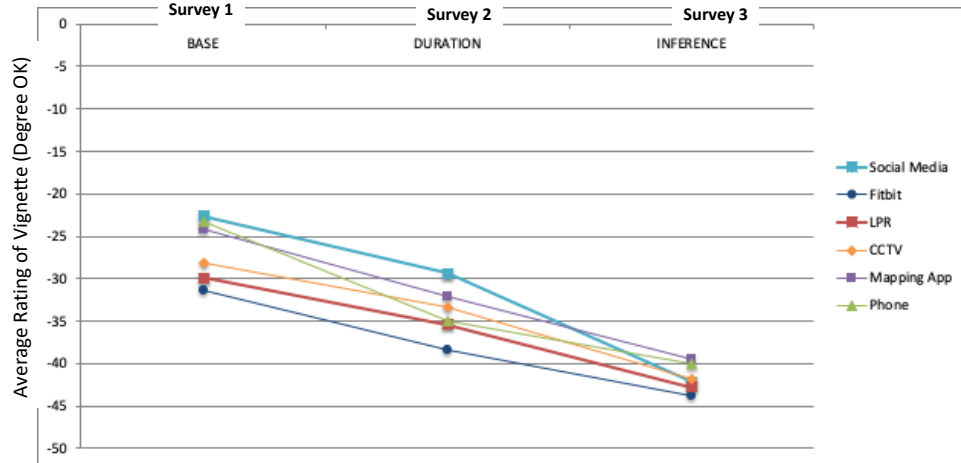
## c) Source

How the location data was gathered—through a social media post, a mapping app, a license-plate reader, a phone, or a CCTV with facial recognition—affected the degree to which the vignette was rated “okay,” as is shown in Figure 5. Capturing location through a phone or CCTV was rated the lowest, or least “okay”; capturing through social media and a mapping app was the highest rated source (although still negative).<sup>103</sup> In other words, respondents did not significantly differentiate across the different sources of gathering location data, particularly in comparison to the importance of who receives the information. This is shown by how the average rating is actually clustered for each survey across the sources and is also evident in Table 5 above in the general regression, where the coefficients are significantly different at times across types of sources, but not large (e.g., the difference between gathering location data via a phone versus a social networking app is -7.26 in Survey 2 (out of a 200-point scale) and not significant for Surveys 1 and 3.

103. Given the attention to the collection of location data from phones, the difference in respondents’ ratings across sources is significant but not a main driver of the appropriateness rating.



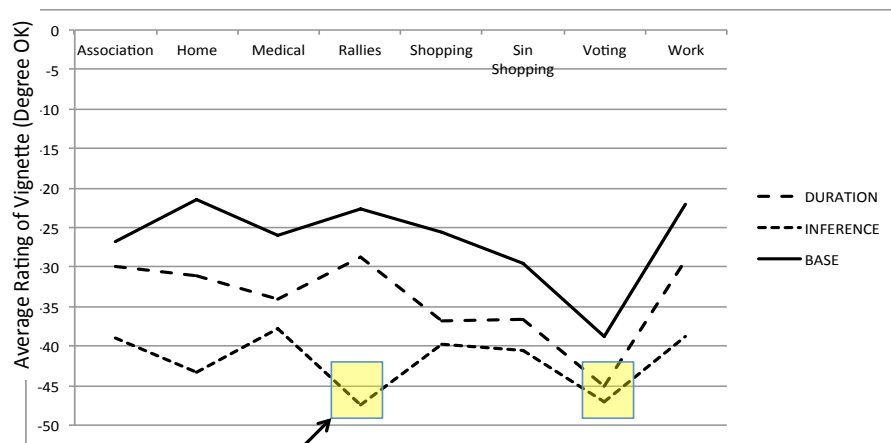
Figure 3: Average Vignette Rating by Data Collection Medium



## d) Inferred Information

Using location information to identify whether someone voted or attended a rally was rated the lowest among the different inferences to be drawn, with an average rating of about -50. See Figure 4, with voting and attending a rally highlighted with yellow boxes.

Figure 4: Average Vignette Rating by Inferred Information



*Note the degree that tracking location about voting and rallies are outside privacy expectations*

## 2. Interactions

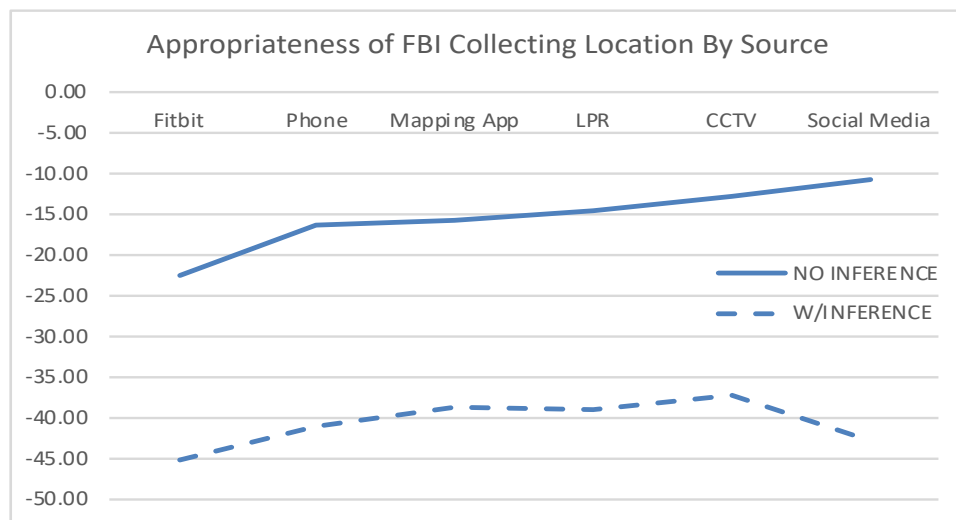
We were interested in whether the source, duration, or inferred information is perceived differently depending on the actor involved. For example, does the importance of the duration of location data collection depend on whether the actor acquiring the data is the FBI versus a family member? Does the importance of how the location data is collected depend on the actor collecting it? In order to identify if the actor modified the importance of the other contextual factors, we calculated the average vignette rating (the degree to which the vignette is rated “okay”). The results are in the Figures below.

### a) Appropriateness of Source by Actor

To illustrate this point, we compared the FBI with data aggregators. Figure 5a, below, illustrates that the importance of the source (how the location data was gathered) was relatively stable for each actor aside from when the collecting actor was the FBI. The degree to which the scenario was appropriate was greatest for the FBI acquiring the location data through social media and least for the FBI accessing the location through a FitBit (with no inference explained).

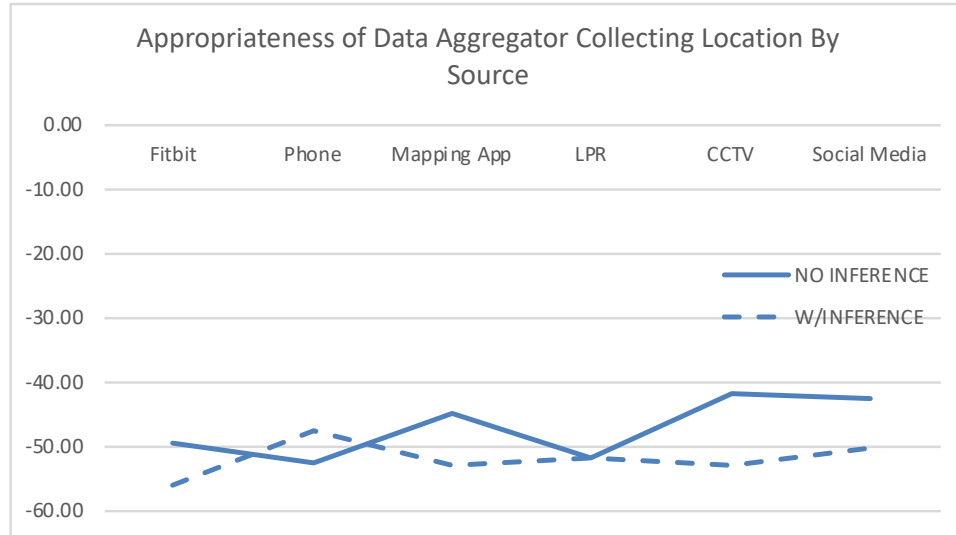
However, when the inference drawn about the individual is added to the vignettes (in Survey 3), the degree the collection of location data is appropriate decreases precipitously, and the manner in which the location data is collected (via phone versus FitBit versus social media) is statistically insignificant.

**Figure 5a: Average Vignette Rating for Each Source for FBI**



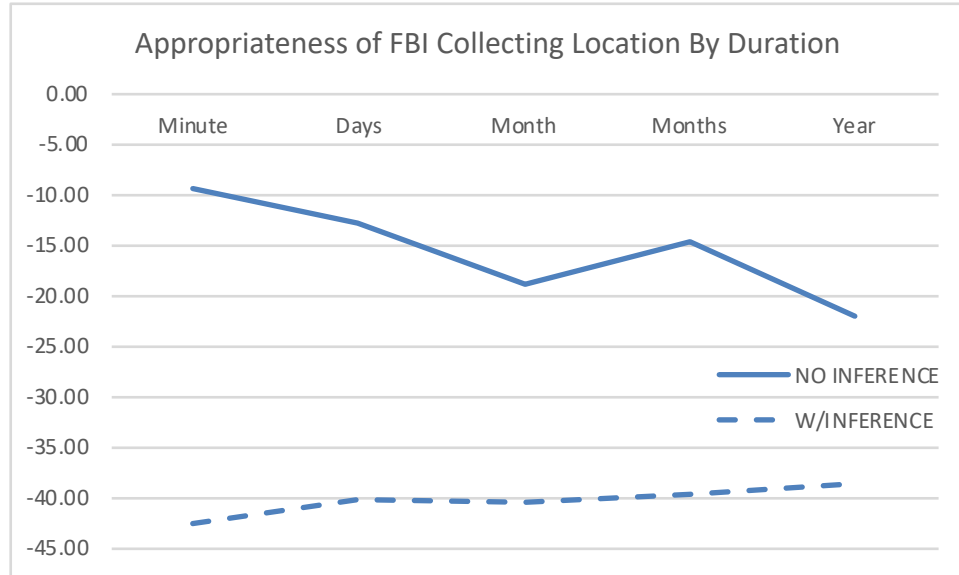
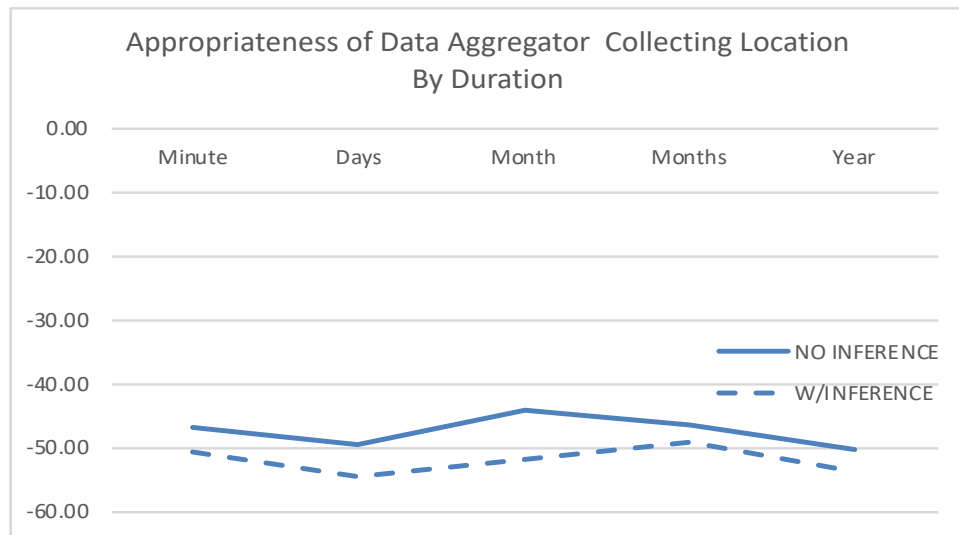
For data aggregators, the collection of location data from any source and either with inferences or without inferences included is not appropriate: the average vignette rating across sources being approximately -50.

**Figure 5b: Average Vignette Rating for Each Source for Data Aggregator**



#### b) Appropriate Duration by Actor

Figure 6a illustrates that duration is significant for particular actors, FBI in particular, compared to data aggregator in Figure 6b. The slope of each line is equal to the relative importance of duration to the rating task: a steeper negative slope is equivalent to the duration being more important to the rating task. For a data aggregator, the rating task is about the same regardless of the duration of the tracking. However, for the FBI, the duration of the surveillance is significant when the inference drawn is not included in the vignette, but disappears when inferences are included. This suggests that the inference drawn about the individual mediates the relationship between duration and the degree to which the location gathering is “okay.” In other words, when individuals are concerned about the duration of surveillance, they are actually concerned about what inferences can be drawn from longer-term surveillance.

**Figure 6a: Average Vignette Rating for Each Duration by FBI****Figure 6b: Average Vignette Rating for Each Duration by Data Aggregator**

### C. DISCUSSION OF MAIN STUDY

Varying the actor in the vignette who gathers the location information affects the degree to which the collection of location data is acceptable. However, the difference between the FBI or city services and a commercial entity (e.g., Yelp) diminishes as duration and inference are added. Importantly,

considering the attention given to the collection of location data from mobile phones, the difference in respondents' ratings of the appropriateness of collecting data across sources is significant but not the central driver of the appropriateness rating. All else being equal, gathering location data via a phone is statistically equivalent to gathering location data from a mapping app or social media but judged more acceptable than license-plate readers, CCTV, and FitBits. This finding may be significant for law and regulation that single out phones for distinctive treatment merely in their capacity to track location; these results suggest that individuals do not differentiate location information gathered via phone versus other mechanisms (CCTV, FitBits, etc.) as having different privacy expectations. The results suggest that the mechanisms for tracking location information, by themselves, do not drive privacy expectations.

The significance of duration disappears when inferences about an individual are also cited. This suggests that it is the potential for drawing inferences that mediates the relationship between duration and the assessments of appropriateness of location tracking. In other words, concerns over surveillance duration are actually concerns over inferences that longer-term surveillance facilitates.

## VI. FOLLOW-UP STUDY

Picking up on an issue we raised in Part I, the findings of our main study revealed one further aspect that needs attention. In particular, we sought greater clarity on how people conceive location in relation to how it is represented in technical systems and the policies that regulate them, either proclaimed by owners or imposed by others. This could tell us something about the match (or mismatch) between what concerns people when they say no to location tracking and the action a company takes to respect this: for example, ceasing to collect GPS data.<sup>104</sup>

Drawing on the finding from the pilot study that people respond to "location" and "GPS" in similar ways, we were interested in the impact of giving meaning or semantics to this numeric value. Pushing a step further toward this Article's driving question—"what is it about location?"—we sought to pinpoint the effects of naming a *place* by comparing it with references to generic *location*. In terms of CI, this follow-up study supplements the main study and pilot study with more specific insights on the parameter of information type and the ontologies that populate its parametric values.

---

104. Or other numerically represented location markers, such as nearest Wi-Fi coordinates or inferred location based on triangulation with nearby cell tower signals.

#### A. FOLLOW-UP STUDY DESIGN

To isolate the importance of adding place to vignettes describing a generic location, we ran two factorial vignette surveys. These allowed us to examine if the meaning of location matters to the respondents by not including the duration or the inference drawn. Otherwise, the same factors and levels were used as in the live survey.

##### 1. Base Survey: Actor-Source

- {Actor} acquires location data from {Source}

For example, the vignette under the first condition would be:

- The FBI acquires location data from the signal of a mobile phone.
- An employer acquires location data from a mapping app (e.g., Google Maps).
- A city emergency service (ambulance, fire) acquires location data from geo-tagged posts on social media (e.g., Twitter, Facebook, Instagram).

##### 2. Base + Place Survey: Actor-Source-Place

- {Actor} acquires location data from {Source} and uses this data to figure out if a person was at {Place}.

The vignette under the second condition would be,

- The FBI acquires location data from the signal of a mobile phone and uses this data to figure out if a person was at a liquor store.
- An employer acquires location data from a mapping app (e.g., Google Maps) and uses the data to figure out if a person was at a liquor store.
- A city emergency service (ambulance, fire) acquires location data from geo-tagged posts on social media (e.g., Twitter, Facebook, Instagram) and uses the data to figure out if a person was at a liquor store.

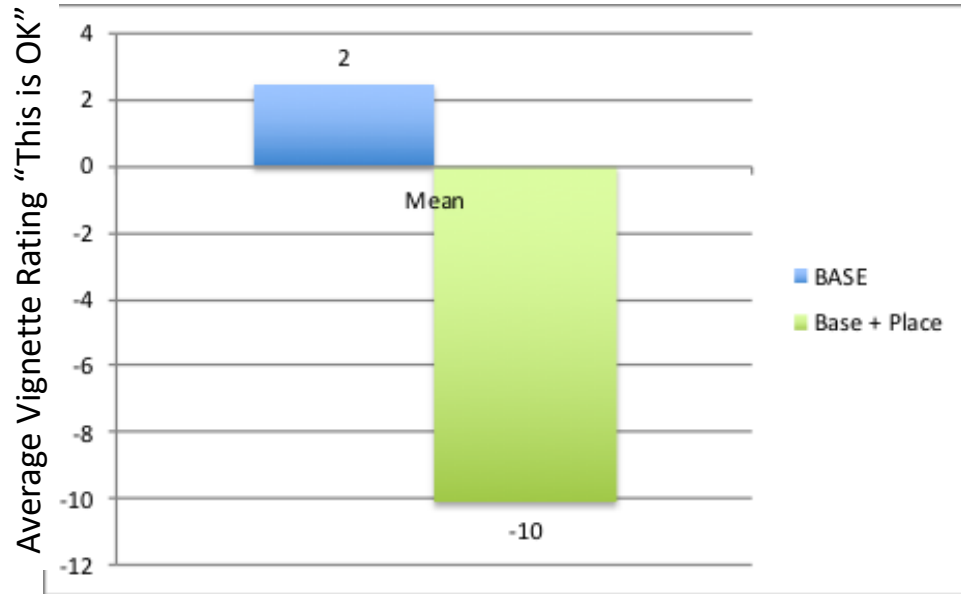
The survey was deployed on approximately 300 Amazon Mechanical Turk respondents, who each rated twenty vignettes.

#### B. FOLLOW-UP STUDY RESULTS

##### 1. *Average Rating Vignette Is “Okay”*

Results were quite stark: adding meaning to location data significantly drives down the average vignette rating, as shown in Figure 8.

Figure 8: Average Vignette Rating for Both Conditions

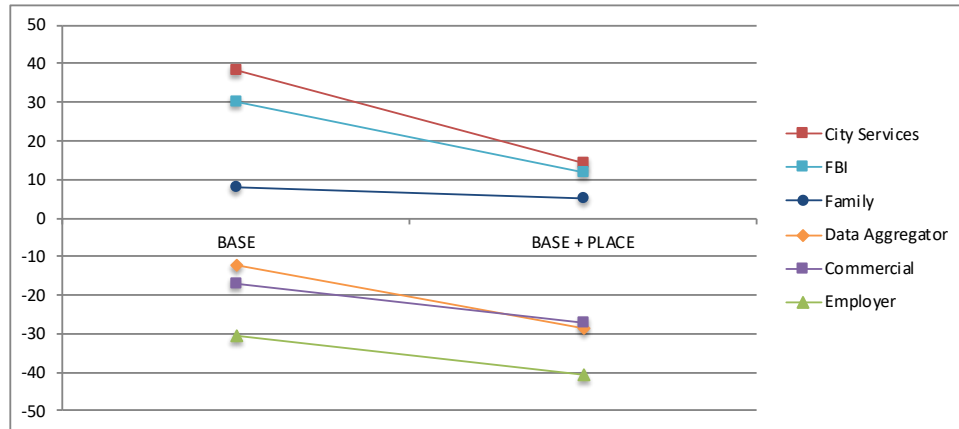


## 2. Actors

In addition, adding place to the vignette affects the collection of location data by the FBI and city services (which were relatively high) disproportionately more than other actors, as shown in Figure 9. The average vignette rating for the FBI drops from +30 to +10.<sup>105</sup>

105. No duration or inferences drawn about the individual were included in this survey. This isolates the impact of adding merely place to location data.

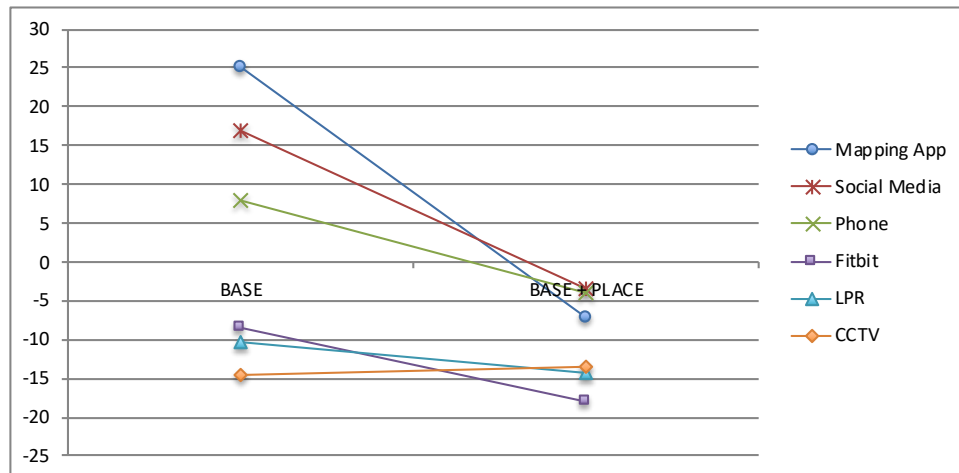
Figure 9: Average Vignette Rating for Each Actor by Condition



### 3. Source

Finally, three sources are disproportionately impacted when the place is given meaning in the vignette: collecting location data from a phone, social media, and mapping app is positive when no meaning for place is provided but negative once the vignette included the place inferred from the location data. This suggests that asking respondents about data collection via these sources normally does not evoke particular places, and it would need to be made explicit in any survey.

Figure 10: Average Vignette Rating by Data Collection Method (Source)





### C. FOLLOW-UP STUDY DISCUSSION

In sum, we found the following.

- Adding place to a generic location negatively affects the degree to which the scenario is “okay” overall, and particularly for the FBI (from +30 to +12) and city services (from +38 to +14) as actors.
- Adding place significantly decreases the degree to which the scenario is “okay” for three sources: Mapping App: from +25 to -7; Social Media: from +17 to -4; and Phone: from +8 to -4.

## VII. SIGNIFICANCE FOR TECHNOLOGY, REGULATION, AND LAW

Amidst growing concerns about the steep rise of location tracking technologies and the widespread infiltration of location into data analytics, this Article asked, “what is it about location?” that worries us, the subjects of tracking. Our results shed light on how people understand location data, and how contextual factors affect people’s reactions to others’ knowing their whereabouts. Among many interesting and actionable findings, the results once and for all debunk the fiction that no expectations of privacy apply in public locations. To the contrary, we found not only that people have definite expectations, but that these expectations are nuanced and are systematically linked to the contextual factors for which we tested. Further, it is also clear from our findings that many common practices in which government and commercial entities engage are at odds with the expectations and attitudes that our studies reveal. Some of those findings are listed below:

- The collection of location data across actors and sources was judged “Not Okay” by respondents. The average ratings for each survey ranged from -29.7 (with place included) to -46.3 (when inferences were included);
- The results suggest that the word “location” is synonymous with “GPS” in judging the scenario as appropriate with no significant difference between the two levels ( $p=0.95$ ). Further, the less precise measurement of locating someone at a street address or within a city was only a small improvement in the appropriateness of collecting location data.
- Duration was significant to the appropriateness of collecting location data when the inference drawn about the individual was not included. Interestingly, neither the length of storage nor the frequency of collection was similarly significant.

- As CI predicts, we found the actor collecting data to be a significant factor affecting people's attitudes. While respondents judged the collection of location data by all actors as "Not Okay" (with a negative rating), they did differentiate between actors. The relatively high approval of the FBI and city services as recipients of location diminished when duration is added as a factor, as well as inferences drawn.
- We anticipated differences, but the results showed that the source of location data (phone, Fitbit, social media) was not a significant predictor of respondents' judgments.
- The simple act of including place (at home, at work, etc.) had an outsized influence on responses. Adding place to a generic location negatively affects the degree to which the scenario is "okay" overall, and particularly for the FBI (from +30 to +12) and city services (from +38 to +14) as actors. Also, adding place significantly decreases the degree to which the scenario is "okay" for three sources: Mapping App: from +25 to -7; Social Media: from +17 to -4; and Phone: from +8 to -4.
- Across all variants, third-party data aggregators were among the most reviled among actors. With or without inferences drawn across sources, the average vignette rating was approximately -50. The juxtaposition of these findings with a hyperactive marketplace of third-party location data brokers siphoning up—buying and selling—is unsettling.<sup>106</sup>
- Finally, it is worth noting the degree of resentment people express about employers collecting location data, except when the location in question happens to be the workplace. Our findings are compatible with important work on employee surveillance by Professor Karen Levy as well as Professors Ifeoma Ajunwa, Kate Crawford, and Jason Schultz.<sup>107</sup> Our findings reinforce their arguments that even lawful employer surveillance of employees contravenes robust expectations. Our study shows clearly that there is a serious need to calibrate the existing letter of law with reasonable expectations of privacy.

---

106. See Jennifer Valentino-De Vries et al., *supra* note 27.

107. Karen E.C. Levy, *The Contexts of Control: Information, Power, and Truck-Driving Work*, 31 INFO. SOC'Y 160 (2015) (examining the impact of monitoring employees); Ifeoma Ajunwa et al., *Limitless Worker Surveillance*, 105 CALIF. L. REV. 735 (2017) (discussing the implications of employee surveillance by an employer in the context of the U.S. trucking industry).

## A. TECHNOLOGY

Our results are directly relevant to privacy-by-design. Here, we return to the intriguing question of how well people's understanding and preference for location privacy is represented (or modeled) in technology architecture and technical mechanisms. Our results flatly contradict the proposition that location privacy can be achieved by simply not collecting one of the technical markers, such as GPS coordinates. Technical research has shown that a smartphone user can be located using publicly available information, even when their location services are turned off.<sup>108</sup> Indeed, Google has admitted to tracking individuals with location services turned off (i.e., no GPS coordinates tracked), by triangulating an individual's whereabouts via nearest cell towers.<sup>109</sup> A further challenge comes from the ability to infer location from ostensibly non-location sensor data, collected by mobile devices where users' permission is not even needed.<sup>110</sup>

These results show that how location data is collected is not important to the privacy expectations. Further, the format of the data, whether as GPS coordinates or a street address, is not as important as locating someone at a "place." For companies, identifying individuals' location via other means, such as a data aggregator, a Wi-Fi sniffer, or a social media post is still not considered "okay." Further, asking individuals about the collection of GPS coordinates or even "location" data may not be precise enough for individuals to make a judgment—for, as noted, the simple inclusion of place (at home, at work, etc.) had an outsized influence on judgments of appropriateness. Furthermore, unless users are informed about the types of inferences that may be drawn, general questions about location privacy are ambiguous.

---

108. See, e.g., Arsalan Mosenia et al., *PinMe: Tracking a Smartphone User around the World*, 4 IEEE TRANSACTIONS ON MULTI-SCALE COMPUTING SYS. 420 (2017) (demonstrating that minimal information is required to track a smartphone user's location even when GPS is turned off); Lindsey Barrett, *Model(ing) Privacy: Empirical Approaches to Privacy Law & Governance*, 35 SANTA CLARA HIGH TECH. L.J. 1, 20 (2018) (noting the different studies which have found that, despite privacy policies, applications on phones can still access the information even when not in use).

109. Shannon Liao, *Google Admits it Tracked User Location Data Even When the Wetting was Turned Off*, VERGE (Nov. 21, 2017), <https://www.theverge.com/2017/11/21/16684818/google-location-tracking-cell-tower-data-android-os-firebase-privacy> [<https://perma.cc/K98P-ZMTF>]; Keith Collins, *Google Collects Android Users' Locations Even When Location Services are Disabled*, QUARTZ (Nov. 21, 2017), <https://qz.com/1131515/google-collects-android-users-locations-even-when-location-services-are-disabled/> [<https://perma.cc/A9ZL-FLX9>].

110. Sashank Narain et al., *Inferring User Routes and Locations Using Zero-Permission Mobile Sensors*, 2016 IEEE SYMP. ON SECURITY & PRIVACY 397 (explaining how a user's travel route can be inferred with high accuracy from gyroscope, accelerometer, and magnetometer information).

Beyond the challenge of stopping end runs around technical location markers, our results raise the question of how to represent location semantics through technical variables, and whether it is even possible. Protecting against inferences drawn from an individual's whereabouts (particularly patterns of movement over time) may be beyond purely technical means; similarly, limiting access to an individual's "place" may be challenging. As noted, place data (home, work, shopping, etc.) may be available outside of GPS and GIS systems via natural language communication on social media.

Finally, the collection of location data by third-party data aggregators was consistently judged inappropriately, no matter the duration or inferences drawn. This finding is consistent with our previous work showing strong disapproval of third-party brokers and aggregators, even when collecting data from public records.<sup>111</sup> This means that the common practice of integrating code from external libraries that either shares or integrates location data from third-party aggregators flies in the face of privacy expectations and significantly undermines trust.<sup>112</sup>

#### B. SIGNIFICANCE FOR REGULATION

The GDPR has introduced new privacy requirements for the data processing practices of firms doing business with European individuals. The processing of identifiable information, including location data, is limited: data processors must meet one of a few criteria, including that the processing of location data is necessary to complete a contract obligation, to protect vital interests of data subject or other person, to perform a task in the public interest, to comply with a law or regulation, or "for the legitimate interests" pursued by the data controller or a third party.<sup>113</sup>

Our findings are helpful for defining a company's legitimate interests. According to GDPR Article 6, there is a three-part test for identifying exceptions to a data processor's legitimate interests.<sup>114</sup> First, is there a legitimate interest behind the processing of location data? Second, is that processing necessary for that purpose? Finally, is the legitimate interest overridden by the individual's interests, rights, or freedoms? According to the GDPR, individuals' interests are defined in terms of reasonable expectations

---

111. See Martin & Nissenbaum, *Privacy Interests in Public Records*, *supra* note 7.

112. Martin, *Breaking the Privacy Paradox*, *supra* note 90; Martin, *The Penalty for Privacy Violations*, *supra* note 90.

113. *Lawful Basis for Processing*, *supra* note 40.

114. *What Is the 'Legitimate Interests' Basis?*, INFO. COMM'R OFF., <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/legitimate-interests/what-is-the-legitimate-interests-basis/> [https://perma.cc/8TDY-5UF9] (last visited Dec. 30, 2019).

around the collection and use of location data.<sup>115</sup> In other words, if individuals have strong expectations that location data will not be processed for that purpose, the interests of the individual are superseded by the ‘legitimate’ interests of the data processor.

The findings here illustrate under what conditions individuals find the gathering of location data to be appropriate and the appropriate inferences to be drawn about them. While using location data to identify if an individual is at home is deemed appropriate for a family member, the majority of scenarios were deemed inappropriate, particularly for commercial actors such as an employer, a data aggregator, or a commercial service (such as Yelp). Taking reasonable expectations seriously, our results suggest that the number of uses for location data aligning with a company’s purposes are considerably fewer than companies may seek to claim.

Outside of GDPR, the study suggests that regulations should focus on the type of information, rather than how the information is collected, to protect the interests of consumers and users. Location data was deemed inappropriate to collect regardless of how the information was gathered. Therefore, regulations that mistakenly narrowly focus on types of collection mechanisms (e.g., only based on trackers or Wi-Fi sniffers) would allow companies to collect location data that people deem inappropriate; such regulations would do little to actually protect the privacy interests of individuals. Our findings reinforce the idea that regulations should not follow specific technologies, but instead map onto values for the nature of the information, the recipients, and the flow constraints (collection, use, sharing, etc).

Instead, privacy regulations should look to limit who has easy access to location data after the initial collection. The results here show that particular actors, such as data aggregators, were consistently deemed not appropriate to collect location data. The current focus in the United States is to heavily regulate the handoff or initial disclosure of information through adequate notification and user consent. After disclosure, regulations are silent. Regulations should instead shift to focus on the sharing, aggregation, and use of information, including location data, after initial disclosure.

---

115. *Id.* (“The GDPR is clear that the interests of the individual could in particular override your legitimate interests if you intend to process personal data in ways the individual does not reasonably expect.”); Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC, (General Data Protection Regulation) (GDPR) art. 6, 2016 O.J. (L 119) 1.

C. SIGNIFICANCE FOR LEGAL DECISIONS

This study fills the need to better understand societal privacy expectations as a means to ascertain whether actual expectations are reasonable. Professors Matthew Kugler and Lior Strahilevitz clearly summarize why actual beliefs (as measured in surveys) are relevant to court opinions. They state, “[w]e show how scientific polling can alleviate concerns that, in undertaking such an inquiry, judges will place undue weight on their own beliefs or on the beliefs of people in their social orbits.”<sup>116</sup> Around location data specifically, Kugler and Strahilevitz note that reasonable expectations of privacy are “the average person’s expectations” or “popular expectations.”<sup>117</sup> These studies empirically examine the privacy expectations of individuals around the collection of location data in public.

Many of our findings could be useful in the courts. For example, the precision of location data (GPS coordinates, street address, city block, etc.) is less important than *who* is collecting the information. Further, the level of precision was far less important to respondents than whether location was identified in terms of a meaningful place (work, rally, home, etc.). Finally, the precision was less important than the type of inferences drawn or type of knowledge created by the breadth of location data collected.

In the past, courts have focused on GPS data gathered from a phone or GPS device in a car. Going forward, they would do well to highlight the nature of the collecting actor (“recipient” in CI terms; boss, the FBI, parents, etc.), rather than the source alone (phone, CCTV, mapping app, etc.). An exception we found was FitBit, which provoked greater disapprobation.

The appropriate duration of collecting location data before a warrant is needed is in flux. Previous work by Kugler and Strahilevitz found duration is not significant in affecting judgments concerning when a warrant is necessary.<sup>118</sup> In our study, duration seems to matter only insofar as it mediates inferences and ceases to play an explanatory role once an inference or place has been declared. In other words, respondents cared about duration only when no meaning was given to the location data. This conforms to what others have dubbed “mosaic theory,”<sup>119</sup> which is an awareness that insignificant bits of information, aggregated, may create a fine-grained picture that can threaten privacy. Interestingly, in the cases of inferring voting and attendance at a

---

116. Kugler & Strahilevitz, *supra* note 1, at 220.

117. *Id.* at 207.

118. *Id.* at 245, 248.

119. Ohm, *supra* note 43, at 373.

political rally from location, respondents disapproved across the board for all the recipients we listed in our study.

D. SIGNIFICANCE FOR HOW LOCATION IS LABELED IN SURVEYS AND LAW

How we ask about location in surveys matters to the normative judgments of individuals. As noted above, previous empirical work has asked about the collection of GPS coordinates. However, adding details such as the duration of collection, the place, and the inferences drawn about the individual decreases the degree to which the vignettes are “okay” and can change the relative importance of the actors, source, and place in determining whether the scenario is acceptable. This means that surveys about GPS location data will not capture privacy expectations regarding location unless surveys also include the type of knowledge created by the aggregated data and the purpose of the collection of the data in terms that are meaningful.

Based on the impacts on judgments when place is specified, researchers should start with the assumption that *place* is independent of the numerical GPS measure of latitude-longitude. Thus, it warrants independent study, in addition to interactions with the identities of collecting actors (recipients).

Three sources are disproportionately affected when place is given meaning in the vignette: collecting location data from a phone, social media, and mapping app is positive when no meaning for place is provided but negative once the place inferred from the location data is included in the vignette. This suggests that asking respondents about data collection via these sources would need to be made explicit in any survey.

Further, respondents appear to assume a short duration of collecting data when no duration of collection is mentioned. The vignette with no duration included was rated the same on average as a vignette with location data stored for one minute only. This is important, as respondents in a survey make assumptions about the given scenario when the researcher is silent on the matter. From our research here, respondents assume no particular “place” as being inferred from location data and assume the duration is short. By remaining silent on those factors, surveys might mistakenly be thought to support the collection of location data under a variety of conditions, when, in fact, this is an artifact of respondents assuming a best-case scenario.

Finally, more research is needed in order to explain consistently negative reactions to data aggregators or data traffickers.<sup>120</sup>

---

120. See Lauren Henry Scholz, *Privacy Remedies*, 94 IND. L.J. (forthcoming 2019) (manuscript at 12–13), available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3159746](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3159746).

## VIII. CONCLUSION

It is important to clarify the scope of our research and highlight its key contributions while acknowledging its limits. Our work set out to address gaps and shortcomings in how location privacy is conceived, which has led, in our view, to flawed technical, regulatory, and legal responses. Because *reasonable expectation of privacy* has served as a critical linchpin in all three of these domains, we approached these gaps and shortcomings through large-scale empirical studies structured around the theory of CI to provide a solid basis for revising these responses. The results of these studies are most striking in roundly debunking assumptions that have impeded privacy policy and practice generally, and in this instance, for location privacy.

What are some of these debunked truisms? First is a misplaced faith in the decisive influence of the public-private dichotomy. The studies reported in this Article (as with those in the preceding two) confirm that “public is public” is plain wrong; our respondents revealed strong, nuanced, and systematic privacy expectations in spaces and places typically considered public. Although we have been early proponents of this view and are no longer alone in holding it,<sup>121</sup> our studies offer compelling empirical backing. Second, our studies reveal that purely mathematical, non-semantic location markers (e.g., GPS coordinates) do not adequately model location privacy expectations. Misleading labels in device and system interfaces may, therefore, deceive users about underlying data practices. Finally, like it or not, respondents were highly discriminating on the question of recipients. From the perspective of CI, this is unsurprising, but, once again, this finding exposes how poorly the public-private dichotomy models expectations of privacy. Respondents were varied in their judgments of appropriateness for family, FBI, etc., though were consistently and deeply negative about third-party location aggregators and brokers.

With these and other general findings, our studies demonstrate the need for further and more detailed investigations. Our findings were obviously rooted in our own intuitions, particular interests, and controversies reported in mainstream media. Clearly, they do not offer anything close to a complete picture of location privacy expectations, particularly under the assumption that five parameters, at least, are simultaneously relevant to these expectations. With some of the gaps filled and a few key misconceptions debunked, this Article is an argument for more detailed and better studies of location privacy that will serve sounder court decisions, regulation, and technology design.

---

121. See Woodrow Hartzog, *The Public Information Fallacy*, 99 B.U. L. REV. 459 (2019).



## APPENDIX A – PILOT STUDY FOR SURVEY DESIGN

Important findings from the pilot studies:

- While the frequency of tracking did not significantly affect the degree to which a vignette was rated “okay,” the duration of tracking did. In particular, duration of tracking was important when the FBI was the actor acquiring the data. These results are suggestive of further lines of inquiry into variation across individuals—for example, whether individuals with high privacy despair (low trust, high privacy is important) are particularly sensitive to the duration factor in relation to the FBI.
- We found, inadvertently, two potential framing effects: a) the order in which subjects were presented control questions, whether before or after vignettes, affected judgments; and b) whether vignettes expressed in second or third person, also had an impact. These findings suggest a need for future work to refine privacy survey methodology.

We focused on the collection of location data due to how easily the data can be used to identify other information (e.g., who you are with) and because new technology supports tracking location data in new ways. Even narrowing the focus to location, grappling with the complexity of studying privacy in public called for a pilot study. In particular, the pilot study helped us to settle preliminary choices along three dimensions: first was the selection and articulation of factors of the factorial vignette survey; second was to guide our choices of respondent control questions; and third was to guide our choices of survey features, including the order of presentation of the parts, and survey voice (“you” or “a person”).

In the pilot survey, we experimented on these three elements. For example, in order to simplify the vignette, we tested a few of the factors (location, frequency, storage) included in the respondent controls both before and after the vignettes, and tested three different ‘voices’ for the vignette. In the end, we were able to simplify by using the word “location” only, dropping frequency, using the third-party vignette voice, and including control questions after the vignettes.

A respondent’s degree of extroversion<sup>122</sup> also impacted the privacy expectations in public in the study conducted by Matthew Kugler and Lior

---

122. The format of the surveys and controls are the same as the live survey in the Article above. We added one control (personality), explained here.

Strahilevitz.<sup>123</sup> Based on the five personality factors used across disciplines,<sup>124</sup> we used the scale for extroversion with the degree to which the respondent saw themselves as “Extroverted, enthusiastic (that is, sociable, assertive, talkative, active, NOT reserved, or shy).”

A. PILOT STUDY SURVEY DESIGN

Two facets of the survey design initially tested: (1) the ordering of the control questions; (2) the voice of the vignettes, and two parameters of the factorial vignette; (3) the precision of the location data described; and (4) significance of frequency of tracking. The results of this pilot study were used to design the main surveys described later.

1. *Features Tested*

- a. **Ordering of Controls and Vignettes.** Did placing the controls before or after the vignettes matter to (i) the rating of the vignette or (ii) the respondents’ ratings of the controls? To ensure the ordering did not impact the vignette ratings, we ran the pilot survey with the respondent controls both before and after asking the respondents to rate the vignettes. Table 1 illustrates the respondent controls being asked after the vignettes.
- b. **Vignette Voice** (“you” versus “a person”). We tested if the ‘voice’ of the vignette (second person, third person, or third person plural) mattered to the judgment of whether the information flow was appropriate as has been suggested before.<sup>125</sup> The survey was run three times with each type of voice and as depicted in Table 2.
- c. **Location.** Did the operationalization of “location” as GPS, location, street address, or city matter to the appropriateness of the scenario offered?
- d. **Storage versus Frequency of Data Collection.** In order to test if the frequency of the data collection or the time the data was stored impacted appropriateness of the information flow, we included both factors in the vignette.

---

123. Kugler & Strahilevitz, *supra* note 1, at 251–55.

124. Robert R. McCrae & Paul T. Costa, Jr., *Validation of the Five-Factor Model of Personality Across Instruments and Observers*, 52(1) J. PERSONALITY & SOC. PSYCHOL. 81, 83 (1987).

125. Slobogin & Schumacher, *supra* note 10, at 759.

## 2. *Vignette Factors in Pilot Study*

Table A1 below includes the vignette factors included in the pilot study to identify the importance of voice, storage, frequency, and precision of location. These factors are later ‘fixed’ in the subsequent study.

**Table A1: Factors for Pilot Survey**

<b>Factor</b>	<b>Levels</b>	<b>Operationalized in Vignette</b>
<b>Frequency</b> of data gathering	Continuous	Continuously, every hour, every day.
<b>Storage</b> How long the information is retained	Continuous	Indefinitely, 1 year, 1 month, 1 day, 1 hour, 10 minutes and then discarded.
<b>Actor</b>	Government	The local police
	Federal Government	FBI
	Phone	The operating system of a phone/device (e.g., Google Android or Apple iOS)
	Commercial	Companies offering a location-based service (local reviews or recommendations)
	Family	Family members (e.g., parents, spouse, or sibling)
<b>Voice</b>		1st person, <b>3rd Person Singular</b> , <i>3rd Person Plural</i>
<b>Precision</b> How specific is the location data	Location	Your location, <b>a person’s location</b> , <i>individuals’ location</i>
	City	Which city you are in, <b>which city a person is in</b> , <i>which city individuals are in</i>
	Street Address	Your nearest street address, <b>a person’s nearest street address</b> , <i>individuals’ nearest street address</i>
	GPS	Your GPS coordinates, <b>a person’s GPS coordinates</b> , <i>individuals’ GPS coordinates</i>

### 3. *Vignette Template for Pilot Study*

[Actor] collects [Precision] [Frequency] and stores that data [Storage]. E.g.,

- Second person: Companies offering a location-based service (local reviews or recommendations) collect your location every 15 minutes and store that data for 1 year.
- Third person: The FBI collects a person's nearest street address continuously and stores that data for 1 hour.
- Third person plural: The operating system of a phone/device (e.g., Google Android or Apple iOS) collects which city individuals are in continuously and stores that data for 1 hour.

### 4. *Vignette Rating Task*

For each vignette, respondents were instructed to indicate the degree to which they agreed with the question “Is this okay?” with a slider. The left side of the slider indicated “Definitely Not Okay” and the right side of the slider indicated “Definitely Okay.” The slider was on a scale of -100 to +100 with the number suppressed so the respondents saw only the labels “Okay” and “Not Okay.”

## B. PILOT RESULTS

### 1. *Ordering of Controls and Vignettes*

In order to test if the order in which the respondents were asked to rate the vignettes and control questions mattered to the results, the surveys were run first with the vignettes after the control questions and a second time with the controls asked after the vignettes. Table A2 includes the sample statistics of both surveys run.

The average vignette rating did not change when the control questions were asked first versus after the vignettes. The average rating remained about -36 (“Not Okay”). Interestingly, the ratings for certain control variables did change when the controls were asked after the vignettes, as shown in Table A2.

Specifically,

- The authoritarian score decreased from -13.32 to -20.58 when the question is asked after the vignettes. In other words, the respondents are less authoritarian after rating scenarios about commercial and governmental tracking.

- The average trust in business rating also decreases from -12.12 to -25.95 when the question is asked after the vignettes are rated. This is consistent with previous work on trust and privacy.<sup>126</sup>

Table A2: Sample Statistics for Surveys with Control Questions Before and After Vignettes

	Average Sample Statistics	
	<u>Controls 1<sup>st</sup></u>	<u>Controls 2<sup>nd</sup></u>
N Respondents	444	<b>406</b>
Authoritarian Scale	-13.32	<b>-20.48</b>
Trust Scale	0.42	<b>-8.80</b>
Female	1.53	1.41
AgeOver35	0.55	0.37
Privacy Important	72.56	71.85
Trust Government	-23.14	-29.63
Trust Business	-12.12	<b>-25.95</b>
_eq2_R2	0.77	0.77
DV Mean	-36.72	-35.82

### 2. *Vignette Voice (“You” Versus “A Person”)*

To test the importance of the vignette voice, the survey was run three times. Voice did make a difference. When the vignettes included a reference to the respondent (“you”), the vignettes were rated less “okay” (-35.32) compared to a third-person voice (-27.05) or a third-person plural voice (-30.45).

### 3. *Location*

In order to examine how respondents’ make sense of the precision of the location data collected, the rating task was regressed on the vignette factors and the results are in Table A3 below. The results suggest that the word “location” is synonymous with “GPS” in judging the scenario as appropriate, with no significant difference between the two levels ( $p=0.95$ ). Precision does matter to the vignette rating with a reference to only the street address (+5.69) and city (+8.19), as both considered improvements for the respondents over collecting GPS-level data.

126. See Martin, *supra* note 95 (examining the impact of the introduction of privacy notices on consumer’s trust).

Table A3: Regression of Vignette Rating Task on Vignette Factors for Pilot Design Survey

	Coef.	p
<b>Location</b>		
Street	5.69	0.00
City	8.19	0.00
GPS	0.09	0.95
(null = location)		
<b>Actor</b>		
CommercialActor	-14.13	0.00
FBIActor	-30.14	0.00
PhoneOSActor	-19.78	0.00
PoliceActor	-34.15	0.00
(null = family)		
FrequencyScale	0.86	0.26
StorageScore	-8.80	0.00
<b>Controls:</b>		
HighExtroversion	0.70	0.64
HighAuthoritarian	12.09	0.00
HighTrustDisposition	5.48	0.00
HighPrivacyImport	-20.90	0.00
HighTrustBusiness	20.32	0.00
_cons	9.88	0.00

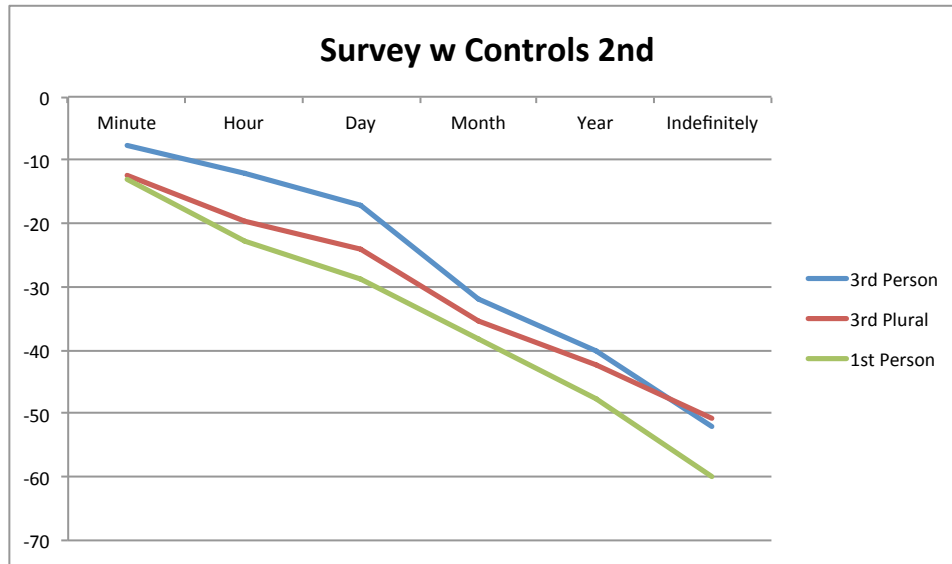
#### 4. *Storage Versus Frequency of Data Collection*

In order to focus on whether the storage of data or the frequency with which the location data is collected impacted the outcome, we examined the relative importance of both factors in the regression results in Table A1 and the average rating task (the degree to which the scenario is “okay”) for each amount of storage in Figure A1.

The length of storage time is inversely related to the rating of the vignette as “okay,” indicated by the steep negative slope in Figure A1. Frequency, by

contrast, was not significant; respondents did not rate the vignette any differently as the frequency levels changed.<sup>127</sup>

Figure A1: Average Vignette Rating for Each Level of Storage by Survey Voice



### 5. Discussion of Pilot Survey

In sum:

- We used the term “location” in the later studies, knowing that the term is equivalent to “GPS” for the respondent;
- We dropped the use of frequency;
- We shifted to the term “duration” for the duration of tracked location information;
- We used the third-person plural in the later vignettes and asked the control questions after the vignettes.

127. Because this result was somewhat surprising, we ran another vignette survey without storage included as a factor to allow the respondent to focus on frequency (from every five seconds to once per day). However, frequency was still not significant; the only difference was the average vignette rating decreased from -35.52 to -31.57 when storage was removed as a factor.

## APPENDIX B – FOLLOW-ON STUDY

We tested the impact of adding place to a vignette with just actors and source. This would allow us to see the effect of explaining the place that the location data provides. In other words, does it matter to respondents if we describe location tracking as gathering location data versus gathering location data to figure out someone is at a particular place?

### A. FOLLOW-ON STUDY #1: ADDING PLACE TO A SURVEY ABOUT LOCATION

To isolate the importance of adding place to vignettes describing a generic location, we ran two factorial vignette surveys. This allowed us to pilot if giving meaning to the location would matter to respondents. The same factors and levels were used in the live survey. The table is provided below as Table B1.

#### 1. Base Survey: Actor-Source

- {Actor} acquires location data from {Source}.

#### 2. Base + Place Survey: Actor-Source Place

- {Actor} acquires location data from {Source} and uses this data to figure out if a person was at {Place}.

For example, the vignette under the first condition would be,

- The FBI acquires location data from the signal from a mobile phone.
- An employer acquires location data from a mapping app (e.g., Google Maps).
- A city emergency service (ambulance, fire) acquires location data from geo-tagged posts on social media (e.g., Twitter, Facebook, Instagram).

Whereas the vignette under the second condition would be,

- The FBI acquires location data from the signal from a mobile phone and uses this data to figure out if a person was at a liquor store.
- An employer acquires location data from a mapping app (e.g., Google Maps) and uses the data to figure out if a person was at a liquor store.
- A city emergency service (ambulance, fire) acquires location data from geo-tagged posts on social media (e.g., Twitter, Facebook, Instagram) and uses the data to figure out if a person was at a liquor store.



Table B1: Factors Used in Pilot

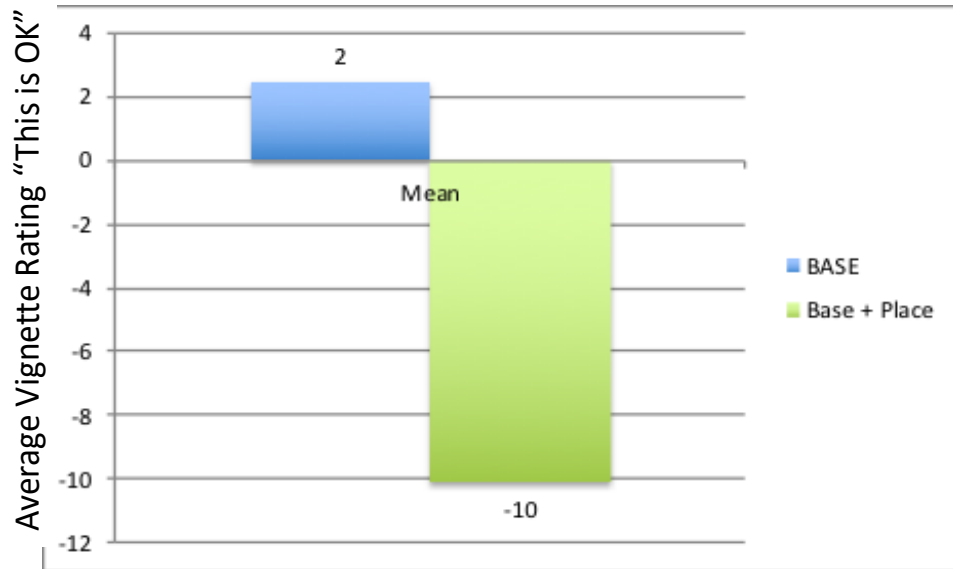
Concept	Description	As operationalized in Vignette
<b>Duration</b>	Length of tracking	A year, about 6 months, a month, a few days, a few minutes
<b>Actor</b>	Government	A city emergency service (e.g., ambulance or fire)
	Federal Government	The FBI
	Employer	An employer
	Commercial data aggregator	A commercial data broker
	Commercial	A commercial location-based service (e.g., Yelp)
	Family	A family member (e.g., parents, spouse, or sibling)
<b>Source</b>	License-plate reader	License-plate readers
	CCTV	CCTV cameras with facial recognition
	Phone	The signal from a mobile phone
	Fit Bit	A fitness app (e.g., FitBit or Stava).
	Social media	From geo-tagged posts on social media (e.g., Twitter, Facebook, or Instagram)
	Mapping app	A mapping app (e.g., Google Maps)
<b>Place</b>	Association	A restaurant or cafe
	Protests/rallies	The National Mall
	Sin Shopping	A liquor store
	Shopping	A shoe store
	Home	Home
	Work	Work
	Medical	A medical clinic
	Voting	A voting site

The survey was deployed using Amazon Mechanical Turk. Approximately 150 respondents each rated twenty vignettes.

1. *Average Rating Vignette Is “Okay”*

Adding place to the vignette and giving meaning to what location data could mean drives down the average vignette rating, as shown in Figure B1.

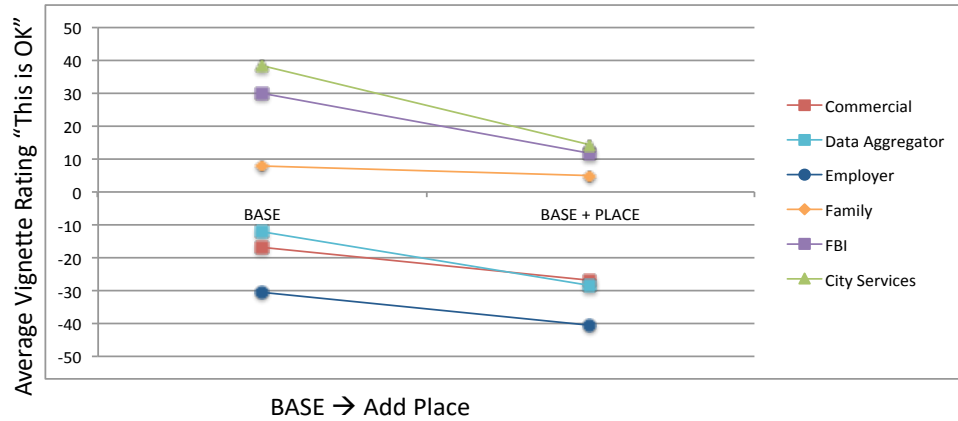
Figure B1: Average Vignette Rating for Both Conditions



2. *Actors*

In addition, adding place in condition 2 impacts the collection of location data by the FBI and city services more than other actors (although all were impacted), as shown in Figure B2. The average vignette rating for the FBI drops from +30 to +10.

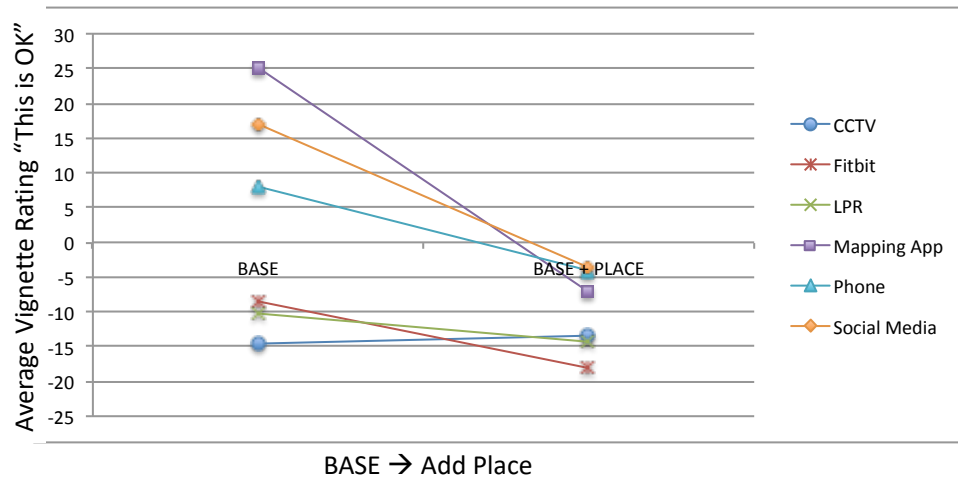
Figure B2: Average Vignette Rating for Each Actor by Condition



### 3. Source

Finally, three sources are disproportionately impacted when the place is given meaning in the vignette: collecting location data from a phone, social media, and a mapping app is positive when no meaning for place is provided but negative once the place inferred from the location data is included in the vignette. This suggests that asking respondents about data collection via these sources normally does evoke particular places. This would need to be made explicit in any surveys.

Figure B3: Average Vignette Rating for Each Source by Condition



B. FOLLOW-ON STUDY #2: ADDING PLACE TO SURVEY WITH DURATION INCLUDED

We then ran the survey two more times, with the duration factor added to both a base survey (actor, source, duration) and the survey with place included (actor, source, duration, and place). This allowed us to isolate the importance of place when duration is also included. In addition, this allowed us to measure how important duration is when the meaning of the location is also included.

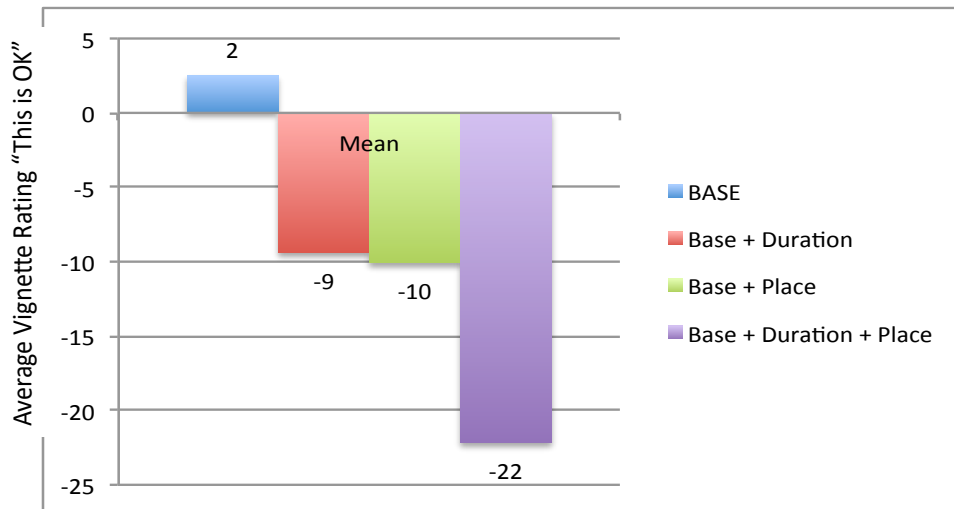
In other words, it is possible that when people are concerned about the duration of data collection, they are actually worried about what someone could find out about them. This would suggest that duration could be mediated by place.

The surveys were run again on Amazon Mechanical Turk with approximately 150 respondents for each condition.

1. *Average Rating Vignette Is "Okay"*

Adding place to the survey with duration already included did impact the average vignette rating.

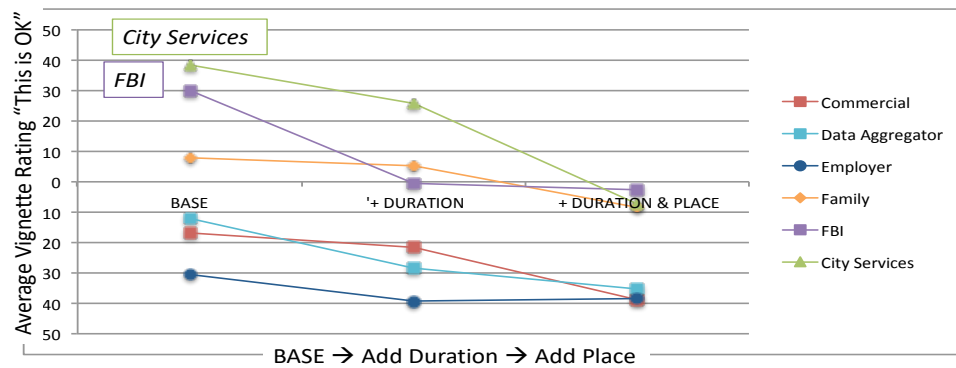
Figure B4: Average Vignette Rating for All 4 Conditions



## 2. Actor

To see how actor is judged when the place is added, we can track how the average rating (“This is Okay”) changes across surveys. The FBI and city services are impacted the most by including duration and place in the vignette.

Figure B5: Average Vignette Rating by Actor for Place Condition

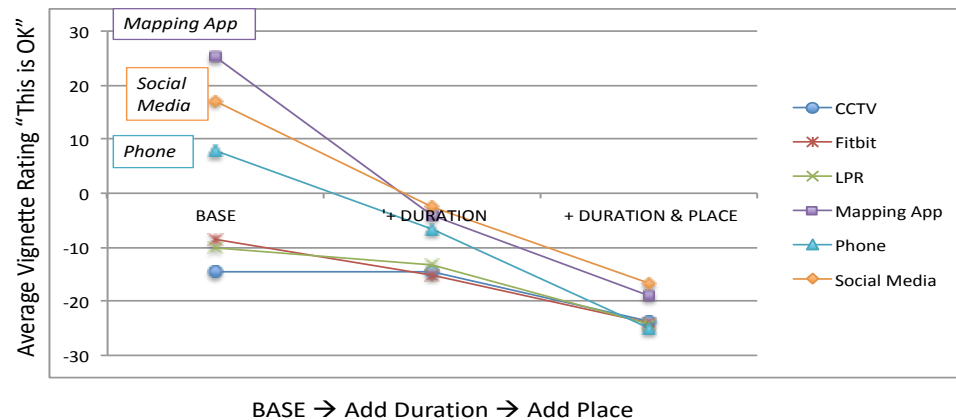


## 3. Source

Mapping App, Social Media, and Phone are impacted the most by including duration and place in the vignette.

- Mapping App: +25 → -19
- Social Media: +17 → -16
- Phone: +8 → -25

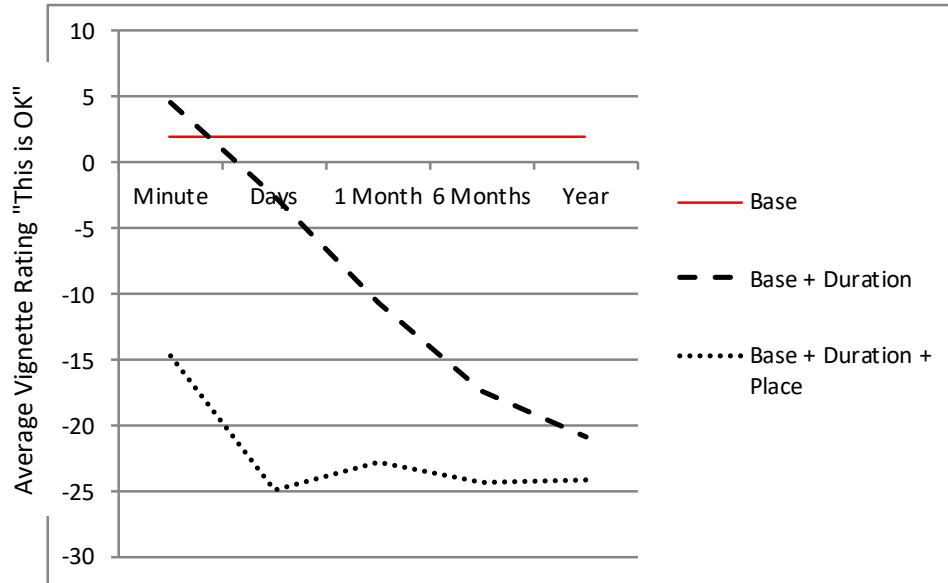
Figure B6: Average Vignette Rating by Source for Place Condition



#### 4. Duration

Finally, we can isolate the importance of duration when place is included in the vignette.

Figure B7: Importance of Duration when Adding Place



In sum, we found the following:

- Adding place to a generic location negatively impacts the degree the scenario is “okay” overall, particularly for the FBI and the city services as actors.
- In addition, the importance of duration is diminished if the place inferred is included. This suggests that “place” explains what respondents were worried about.
- Adding duration and place significantly decreases the degree the scenario is “okay” for three sources:
  - Mapping App: +25 → -19
  - Social Media: +17 → -16
  - Phone: +8 → -25

## APPENDIX C – QUALITY OF SAMPLES

The main survey was deployed through KnowledgeNetworks for a nationally-representative sample. Approximately 1,500 respondents took one of three possible vignette surveys. KnowledgeNetworks is an online research panel representative of the entire U.S. population. KnowledgeNetworks panel members are randomly recruited through probability-based sampling. Households are provided with access to the internet and hardware if needed.

At the same time, the survey was deployed through Amazon's Mechanical Turk where 1,200 respondents rated a total of 12,600 vignettes; 43% were female and 39% were over thirty-five years old. The sample was United States-only and each respondent was paid \$1.70 for taking the survey.

In a separate survey on privacy expectations for websites, Kirsten Martin has compared results from Amazon Mechanical Turk with results from a nationally representative sample from KnowledgeNetworks. The survey from the Amazon Mechanical Turk sample produces the same theoretical generalizations as the survey from the KnowledgeNetworks survey, illustrating the ability to build generalizable theory from Amazon Mechanical Turk samples in online privacy studies.<sup>128</sup>

Figure C1

### Organization of Data

Question 5 of 38 (11%)

Please read the following short vignette and answer the question below.

An employer acquires location data from a fitness app (e.g., FitBit or Stava) for a period of a year and uses the data to figure out if a person was at work.

Please move the slider towards the left for "strongly disagree" or to the right for "strongly agree"

**This is OK**

Strongly Disagree      Neutral      Strongly Agree

Previous Question   Save and Continue

User #1

User #2

Vignette Rating #1  
Vignette Rating #2  
Vignette Rating #3  
.  
.  
.  
Vignette Rating #30

Vignette Rating #1  
Vignette Rating #2  
Vignette Rating #3  
.  
.  
.  
Vignette Rating #30

Options:

- Time to take survey
- Respondent R2
- Respondent Std Dev
- Respondent 'range'
- Last5Qs v. First5Qs

128. Martin, *Privacy Notices as Tabula Rasa*, *supra* note 90, at 16; Martin, *The Penalty for Privacy Violations*, *supra* note 90, at 108.

The sample was analyzed for nonresponsive respondents. Since the respondents each rated thirty independently-generated vignettes, the pattern of their rating on a sliding scale of -100 to +100 for each vignette could be analyzed. We marked two types of surveys as unresponsive: those that rated over twenty of the thirty vignettes as “0” (never moved the slider) and those that rated over twenty-five vignettes at one of the end points (moved the slider to one end almost every time). For the KnowledgeNetworks sample, this resulted in 10% of Survey 1 respondents, 13% of Survey 2 respondents, and 16% of Survey 3 respondents being removed from the pool. The number of respondents discarded from non-responsive ratings. For the Amazon Mechanical Turk sample, 2% of Survey 1 respondents, 5% of Survey 2 respondents, and 11% of Survey 3 respondents were found to be unresponsive. The Amazon Mechanical Turk sample was higher quality than the KnowledgeNetworks sample with the same theoretically generalizable findings.

Table C1

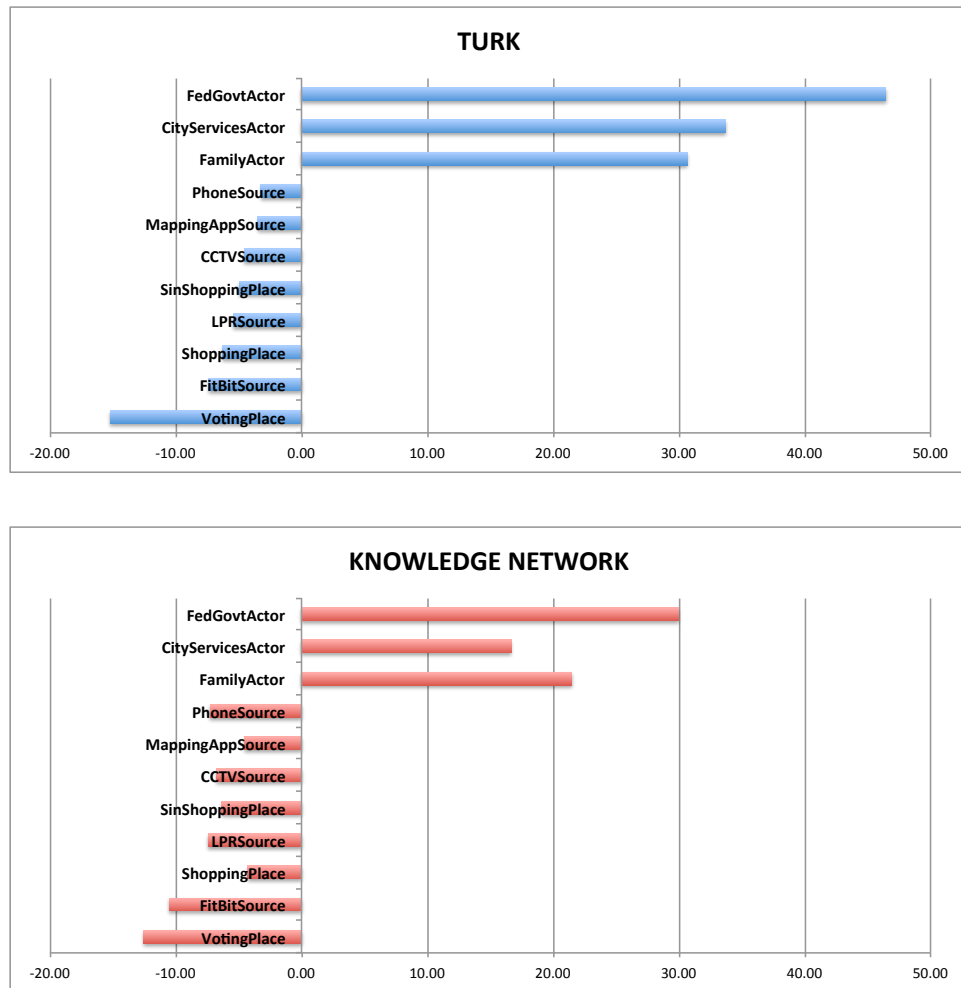
	N	Bad Resp	Very Bad Resp	EndPts > 20 Bad EndPts	EndPts > 25 Very Bad EndPts	0s > 15 Bad 0s	0s > 20 VeryBad 0s
<b>Mechanical Turk</b>							
Base	396	5.5%	2.3%	5.5%	2.3%	0.0%	0.0%
Duration	407	5.6%	5.4%	10.2%	5.4%	0.2%	0.2%
Inference	400	21.6%	11.3%	21.1%	11.3%	0.5%	0.0%
<b>Knowledge Networks</b>							
Base	502	19%	10%	11.0%	3.7%	10.8%	7.6%
Duration	524	22%	13%	13.7%	6.4%	12.1%	8.8%
Inference	509	30%	16%	23.1%	11.5%	9.1%	6.7%



Table C1 (continued)

<b>M Turk</b>	<b>N</b>	<b>Average</b> <b>OKDV</b>	<b>Female</b>	<b>Age</b> <b>Over 35</b>	<b>Trust</b> <b>Business</b>	<b>Privacy</b> <b>Important</b>	<b>Trust</b> <b>Gov't</b>	<b>Trust</b> <b>Disposition</b>	<b>Authoritarian</b> <b>Scale</b>	<b>Trust</b> <b>Scale</b>
Base	396	-23.25	46%	43%	3.58	74.78	-16.49	33.82	-16.42	6.97
Duration	407	-27.40	38%	47%	3.73	73.30	-18.21	31.22	-15.38	5.58
Inferen	400	-51.18	47%	46%	5.23	78.78	-26.63	34.73	-22.26	4.45
<b>KN w/o Bad</b>	<b>N</b>	<b>Average</b> <b>OKDV</b>	<b>Female</b>	<b>Age</b> <b>Over 35</b>	<b>Trust</b> <b>Business</b>	<b>Privacy</b> <b>Important</b>	<b>Trust</b> <b>Gov't</b>	<b>Trust</b> <b>Disposition</b>	<b>Authoritarian</b> <b>Scale</b>	<b>Trust</b> <b>Scale</b>
Base	407	-26.64	49%	75%	3.22	71.94	-20.93	37.77	7.42	6.54
Duration	408	-33.95	54%	74%	4.02	69.62	-20.51	36.22	7.60	7.81
Inference	356	-41.83	50%	73%	4.41	69.75	-23.88	35.99	2.09	5.73
<b>KN w/o Very Bad</b>	<b>N</b>	<b>Average</b> <b>OKDV</b>	<b>Female</b>	<b>Age</b> <b>Over 35</b>	<b>Trust</b> <b>Business</b>	<b>Privacy</b> <b>Important</b>	<b>Trust</b> <b>Gov't</b>	<b>Trust</b> <b>Disposition</b>	<b>Authoritarian</b> <b>Scale</b>	<b>Trust</b> <b>Scale</b>
Base	453	-29.70	49%	75%	2.85	72.32	-23.08	36.40	5.57	5.44
Duration	455	-36.60	53%	75%	2.53	70.97	-22.93	35.52	6.83	5.94
Inference	427	-46.30	50%	75%	4.10	72.14	-27.67	35.88	2.76	4.57

Figure C2: Theoretical Generalizations



# CAN YOU PAY FOR PRIVACY?

## CONSUMER EXPECTATIONS AND THE BEHAVIOR OF FREE AND PAID APPS

*Kenneth A. Bamberger,<sup>†</sup> Serge Egelman,<sup>††</sup> Catherine Han,<sup>†††</sup>  
Amit Elazari Bar On<sup>‡</sup> & Irwin Reyes<sup>‡‡</sup>*

### ABSTRACT

“Paid” digital services have been touted as straightforward alternatives to the ostensibly “free” model, in which users actually face a high price in the form of personal data, with limited awareness of the real cost incurred and little ability to manage their privacy preferences. Yet, the actual privacy behavior of paid services, and consumer expectations about that behavior, remain largely unknown.

This Article addresses that gap. It presents empirical data both comparing the true cost of “paid” services as compared to their so-called “free” counterparts, and documenting consumer expectations about the relative behaviors of each.

We first present an empirical study that documents and compares the privacy behaviors of 5,877 Android apps that are offered both as free and paid versions. The sophisticated analysis tool we employed, AppCensus, allowed us to detect exactly which sensitive user data is accessed by each app and with whom it is shared. Our results show that paid apps often share the same implementation characteristics and resulting behaviors as their free counterparts. Thus, if users opt to pay for apps to avoid privacy costs, in many instances they do not receive the benefit of the bargain. Worse, we find that there are no obvious cues that consumers can use to determine when the paid version of a free app offers better privacy protections than its free counterpart.

We complement this data with a second study: we surveyed 1,000 Android mobile app users as to their perceptions of the privacy behaviors of paid and free app versions. Participants indicated that consumers are more likely to expect the paid version to engage in

---

DOI: <https://doi.org/10.15779/Z38XP6V40J>

© 2020 Kenneth A. Bamberger, Serge Egelman, Catherine Han, Amit Elazari Bar On & Irwin Reyes. This research was supported by the Center for Long-Term Cybersecurity (CLTC) at UC Berkeley, the National Science Foundation (NSF) under grant CNS-1817248, and the National Security Agency’s (NSA) Science of Security Program under contract H98230-18-D-0006. Sanjana Parikh provided superb research assistance.

<sup>†</sup> The Rosalinde and Arthur Gilbert Professor of Law, UC Berkeley; Faculty Co-Director, Berkeley Center for Law and Technology.

<sup>††</sup> Research Director, Usable Security & Privacy Group, International Computer Science Institute (ICSI); Research Scientist, Electrical Engineering and Computer Sciences, UC Berkeley.

<sup>†††</sup> Undergraduate student, UC Berkeley.

<sup>‡</sup> Lecturer, UC Berkeley School of Information.

<sup>‡‡</sup> Researcher, Usable Security and Privacy Group, International Computer Science Institute (ICSI).

privacy-protective practices, to demonstrate transparency with regard to its data collection and sharing behaviors, and to offer more granular control over the collection of user data in that context.

Together, these studies identify ways in which the actual behavior of apps fails to comport with users' expectations, and the way that representations of an app as "paid" or "ad-free" can mislead users. They also raise questions about the salience of those expectations for consumer choices.

In light of this combined research, we then explore three sets of ramifications for policy and practice.

First, our findings that paid services often conduct equally extensive levels of data collection and sale as free ones challenge understandings about how the "pay for privacy" model operates in practice, its promise as a privacy-protective alternative, and the legality of paid app behavior.

Second, our findings offer important insights for legal approaches to privacy protection, undermining the legitimacy of legal regimes relying on fictive "notice" and "consent" that do not reflect user understandings as bases for the collection, sale, and processing of information. They fortify demands for a privacy law that focuses on vindicating actual consumer expectations and prohibiting practices that exploit them, and strengthen the argument for ex ante regulation of exploitative data practices where consumers are offered no opportunity for meaningful choice or consent.

Third, our work provides technical tools for offering transparency about app behaviors, empowering consumers and regulators, law enforcement, consumer protections organizations, and private parties seeking to remedy undesirable or illegal privacy behavior in the most dominant example of a free vs. paid market—mobile apps—where there turns out to be no real privacy-protective option.

## TABLE OF CONTENTS

<b>I.</b>	<b>INTRODUCTION .....</b>	<b>330</b>
<b>II.</b>	<b>PAYING FOR PRIVACY: OUR RESEARCH IN CONTEXT.....</b>	<b>336</b>
A.	PAYING FOR PRIVACY .....	336
B.	APP BEHAVIOR AND CONSUMER EXPECTATIONS STUDIES .....	340
1.	<i>Paid and Free App Behavior Study</i> .....	340
a)	Static Analysis .....	342
b)	Dynamic Analysis .....	342
2.	<i>Consumer Expectations Survey</i> .....	343
a)	Open-Ended Questions .....	344
b)	Likert-Scale Questions.....	345
3.	<i>Limitations</i> .....	346
<b>III.</b>	<b>FINDINGS.....</b>	<b>347</b>
A.	APP BEHAVIOR .....	347
1.	<i>Declared Android Permissions</i> .....	348
2.	<i>Bundled Third-Party Packages</i> .....	349
3.	<i>Network Transmissions</i> .....	350
B.	CONSUMER EXPECTATIONS SURVEY DATA .....	351
1.	<i>Open-Ended Questions</i> .....	352
a)	Expected Differences .....	352
b)	User Preference .....	352
2.	<i>Expectations About Privacy Behaviors</i> .....	353
<b>IV.</b>	<b>IMPLICATIONS FOR PRIVACY PROTECTION .....</b>	<b>354</b>
A.	INSIGHTS FOR LAW AND POLICY .....	355
1.	<i>Meaningful Consent</i> .....	355
2.	<i>Consumer Expectations and Privacy Enforcement</i> .....	356
3.	<i>Risk Salience, Transactional Salience, and Privacy Protection</i> .....	359
B.	OPPORTUNITIES FOR CONSUMER EMPOWERMENT AND ENHANCED OVERSIGHT.....	363
<b>V.</b>	<b>CONCLUSION.....</b>	<b>364</b>

## I. INTRODUCTION

Users pay a high price to enjoy “free” digital services as they engage in the most prominent quid pro quo of the digital age: the exchange of personal information and privacy for utility and comfort. At the same time, amid ensuing media attention to growing data abuses, businesses have come to recognize that users are often willing to expend a small monetary sum for an “ad-free” experience. Many companies have promoted such “paid” services as a straightforward alternative, in which users can choose to pay with money, instead of with personal privacy.<sup>1</sup>

The free model has dominated many provinces of digital space. In aggregate, more than 90% of available mobile applications are free.<sup>2</sup> Rather than charging consumers directly, free app developers generate significant revenue in other ways, such as partnering with advertising networks to provide ads to users. Google’s AdMob, for instance, is found in more than 1 million apps, and has yielded more than \$1 billion collectively to developers.<sup>3</sup>

Scholarship has increasingly critiqued the free model, revealing its true cost.<sup>4</sup> Properly understood, “free” transactions are anything but. They are, instead, exchanges between consumers and services collecting their data<sup>5</sup>—unequal exchanges, moreover, in which users possess limited awareness of the real costs incurred, and little ability to manage their privacy preferences.<sup>6</sup> This scholarship documents “the many reasons consumers find it impossible to account for the risk of harm from online data collection.”<sup>7</sup> The opacity of app

---

1. See *infra* Section II.A.

2. Mansoor Iqbal, *App Download and Usage Statistics (2019)*, BUSINESS OF APPS (Apr. 24, 2020), <https://www.businessofapps.com/data/app-statistics/> [<https://perma.cc/42NU-GZ84>].

3. *Get paid to show relevant ads from over a million advertisers with Google AdMob*, GOOGLE, <https://developer.android.com/distribute/best-practices/earn/show-ads-admob> [<https://web.archive.org/web/20181129004421/https://developer.android.com/distribute/best-practices/earn/show-ads-admob>] (last visited Jan. 7, 2020).

4. See, e.g., John M. Newman, *The Myth of Free*, 86 GEO. WASH. L. REV. 513, 520 (2018); Chris Jay Hoofnagle & Jan Whittington, *Free: Accounting for the Costs of the Internet’s Most Popular Price*, 61 UCLA L. REV. 606, 613 (2014); Katherine J. Strandburg, *Free Fall: The Online Market’s Consumer Preference Disconnect*, 2013 U. CHI. LEGAL F. 95, 95 (2013).

5. Hoofnagle & Whittington, *Free*, *supra* note 4, at 608; see also Jan Whittington & Chris Jay Hoofnagle, *Unpacking Privacy’s Price*, 90 N.C. L. REV. 1327 (2012) (applying transaction cost economics to define the relationship between consumers and social networks as an exchange).

6. See Jonathan A. Obar & Anne Oeldorf-Hirsch, *The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services*, 23 INFO., COMM. & SOC. 1, 16 (2018) (noting that of more than 500 surveyed users, 93% accepted a “first-born child assignment” term and 98% ignored or missed it).

7. See, e.g., Paul Ohm, *Free for the Taking (or Why Libertarians are Wrong about Markets for Privacy)*, JOTWELL (May 26, 2014), <http://cyber.jotwell.com/free-for-the-taking-or-why>

behavior obscures the extensive processing of personal information and the fact that the collection and sale of user information “is the main business proposition.”<sup>8</sup> The misdirection of privacy policies and the framing effects of the “myth” of free, moreover, exacerbate the “privacy paradox,”<sup>9</sup> by which consumers behave inconsistently with their actual privacy-protective preferences when it comes to decisions about personal information.

The paid model of digital services has been touted as a means to solve these inadequacies, and enhance consumer choice regarding privacy.<sup>10</sup> User fees offer a substitute for ad revenues, and—potentially—from the intrusive data collection that fuels targeted advertising.<sup>11</sup> Users paying for apps generally expect them to be of higher quality compared to free versions, and the removal of ads in the paid version may be understood (rightly or wrongly) as implying freedom from the associated extensive data collection.<sup>12</sup> Media outlets,

---

-libertarians-are-wrong-about-markets-for-privacy/ [https://perma.cc/C4LB-KMJQ] (“*Free* and *Free Fall* document the many reasons consumers find it impossible to account for the risk of harm from online data collection.”).

8. Chris Jay Hoofnagle & Jan Whittington, *Free*, *supra* note 4, at 606, 613, 628; *see also id.* at 620 (“[B]ut firms online are like firms offline—both spend money to generate their products and both must recoup costs to survive.”); *see also id.* at 628 (“For these types of businesses and the third parties they market to, the collection and sale of personal information about consumers is the main business proposition.”); *see also id.* at 633 (“If, for any reason, a firm is unable to earn enough revenue from either ads or paying customers, the firm can simply sell the personal information on the market. The firm may not even have intended to capitalize on the personal information it collected with each transaction, free or otherwise.”).

9. *See, e.g.*, Alessandro Acquisti et al., *Privacy and Human Behavior in the Age of Information*, 347 *SCIENCE* 509, 510 (2015) (“This discrepancy between attitudes and behaviors has become known as the ‘privacy paradox.’”).

10. *See, e.g.*, David Z. Morris, *Sheryl Sandberg Says Facebook Users Would Have to Pay for Total Privacy*, *FORTUNE* (Apr. 7, 2018), <http://fortune.com/2018/04/07/sheryl-sandberg-says-facebook-users-would-have-to-pay-for-total-privacy/> [https://perma.cc/J3KH-6NEY].

11. *See* Stacy-Ann Elvy, *Paying for Privacy and the Personal Data Economy*, 117 *COLUM. L. REV.* 1369, 1373, 1373 n.16 (2017) (describing the pay-for-privacy (PFP) approach, “which requires consumers to pay higher fees to avoid data collection and targeted advertisements while offering discounts to consumers who consent to these practices”) (citing Letter from Senator Elizabeth Warren to Tom Wheeler, Chairman, FCC 2 (June 21, 2016), [http://www.warren.senate.gov/files/documents/2016-6-21\\_Letter\\_to\\_FCC\\_re\\_Privacy\\_Rulemaking.pdf](http://www.warren.senate.gov/files/documents/2016-6-21_Letter_to_FCC_re_Privacy_Rulemaking.pdf) [http://perma.cc/9WWT-7362] (describing [i]nternet service provider discount plans as ‘requir[ing] consumers to pay hundreds of dollars extra each year so that [a company] does not collect and sell information on the websites they visit, the ads they see, and the terms they enter into search engines’ ”)); *see also* Michael R. Hammock & Paul H. Rubin, *Applications Want to be Free: Privacy Against Information*, *COMPETITION POL’Y INT’L* (2011) (suggesting that regulators ignore customer concerns and allow the market to provide solutions to concerned consumers to enhance their privacy).

12. *See* Matthew Panzarino, *Why You Should Want to Pay for Apps*, *NEXT WEB* (Apr. 23, 2011), <https://thenextweb.com/apps/2011/04/24/why-you-should-want-to-pay-for-apps/> [https://perma.cc/QE9K-GGCE]; Max Van Kleek et al., *X-Ray Refine: Supporting the*

moreover, have reflected such expectations, crediting paid apps with having better security and privacy assurances than free apps.<sup>13</sup>

Yet, unlike the free business model, the paid model for digital services has largely evaded scholarly attention. Moreover, the actual privacy behavior of paid apps, and consumer expectations about that behavior, have largely evaded scholarly attention. This Article addresses that gap. It presents empirical data both comparing the true cost of “paid” services as compared to their so-called “free” counterparts, and documenting consumer expectations about the relative behaviors of each.

We first present an empirical study that applied a scalable analysis framework to document and compare the privacy behaviors of thousands of Android apps that are offered both as free and as paid versions.<sup>14</sup> Our method employed static and dynamic analysis: static analysis to determine the third-party Software Development Kits (SDKs) bundled with each app and the permissions that each app requests; and dynamic analysis to monitor what sensitive data is collected by which remote services in real-time, as each app is executed. From a random sample of free apps listed on the Google Play Store’s category-level top charts, we examined thousands of pairs of free apps and their paid counterparts. The sophisticated analysis tool we employed, AppCensus, developed through a collaboration at the International Computer Science Institute (ICSI),<sup>15</sup> allowed us to detect exactly which sensitive user data is accessed by each app and with whom it is shared.<sup>16</sup>

Utilizing our framework to compare 5,877 paired apps in their “free” and “paid” versions, we examined whether the cost paid by the user (in privacy terms) was in fact lower in paid services, challenging the common conception

---

*Exploration and Refinement of Information Exposure Resulting from Smartphone Apps*, ACM CHI CONF. ON HUM. FACTORS COMPUTING SYSS. (2018) (“Free apps (or freemium versions of apps) were naturally expected to send data to more companies than their paid counterparts because paid apps were perceived to need less ad support.”).

13. Sara Angeles, *Are Free Apps Safe?*, BUS. NEWS DAILY (Aug. 2, 2013), <https://web.archive.org/web/20181129010454/businessnewsdaily.com/4868-free-app-security-risk.html> [<https://perma.cc/9A93-5RRS>].

14. See *infra* Section II.B(1).

15. AppCensus was started as a research project at the International Computer Science Institute (ICSI), which is a research institute affiliated with UC Berkeley, and has since been spun off as an independent startup. See *How This Works*, APPCENSUS, <https://search.appcensus.io/about> [<https://perma.cc/JNG8-PFQ3>] (last visited Jan. 7, 2020); APPCENSUS, <https://www.appcensus.io/> [<https://perma.cc/7D5Z-LH8A>] (last visited Jan. 7, 2020).

16. AppCensus AppSearch analyzes free publicly-available Android apps and reports the private and personally identifiable information that different apps access and share with other parties over the internet, what personal data is being accessed by an app, and then with whom that app shares it. The results reflect the actual behavior of the apps when they are used. See *How This Works*, APPCENSUS, *supra* note 15.



that paid services are more secure, offer stronger privacy protections, and share less personal data. Our results show that paid apps often share the same implementation characteristics and resulting behaviors as their free counterparts: 48% of the paid apps we examined carried all of the same third-party code (e.g., for advertising, analytics, graphics rendering, logging, etc.) as their free versions; 56% of paid apps had all the same privileges to access sensitive system resources; and 38% of paid apps collected all the same personal and tracking information about users as their free versions. Thus, if users opt to pay for apps to avoid privacy costs,<sup>17</sup> in many instances they do not receive the benefit of the bargain. Worse, we find that there are no obvious cues that consumers can use to determine when the paid version of a free app offers better privacy protections than its free counterpart.<sup>18</sup>

We complement this data with a second study: we surveyed a large sample of mobile app users as to their perceptions of the privacy behaviors of paid and free app versions.<sup>19</sup> The survey explored consumers' expectations around privacy with regard to different app versions, and specifically considered whether, when apps are advertised as "ad-free," most consumers believe that this is synonymous with "better privacy." Our findings suggest that consumers not only expect more privacy-preserving behaviors from the paid app in practice, but are more likely to assume a higher level of transparency from the paid version about its data collection and sharing behaviors, as well as more granular control over the collection of their data. Thus, the mere act of paying for apps—regardless of how much—is associated with receiving better privacy. Nonetheless, when asked in open-ended questions at the beginning of the survey (before being primed to consider privacy and security), a large majority (83%) of participants indicated that they would choose to buy the free app version, and nearly none referenced privacy or security behaviors as a driving consideration. So while our results indicate that consumers do care about privacy and believe that paying for apps yields better privacy, this is far from the only factor that consumers consider when choosing whether or not to purchase a given app.

---

17. *Data Privacy: What the Consumer Really Thinks*, ACXION (Feb. 2018), [https://dma.org.uk/uploads/misc/5a857c4fdf846-data-privacy---what-the-consumer-really-thinks-final\\_5a857c4fdf799.pdf](https://dma.org.uk/uploads/misc/5a857c4fdf846-data-privacy---what-the-consumer-really-thinks-final_5a857c4fdf799.pdf) [<https://perma.cc/W4J5-TNRM>] ("[A] sizeable proportion of consumers indicate that they would prefer to pay for online services so that they do not have to share any personal data.").

18. See generally *FPF Mobile Apps Study*, FUTURE PRIVACY F., (June 2012), <http://www.futureofprivacy.org/wp-content/uploads/Mobile-Apps-Study-June-2012.pdf> [<https://perma.cc/WVV8-HQ2J>] (finding that only 48% of free apps and 32% of paid apps provide in-app access to a privacy policy).

19. See *infra* Section II.B(2).

Our combined research identifying ways in which the actual behavior of apps fails to comport with users' expectations has important ramifications for policy and practice.

First, our findings that paid apps often conduct equally extensive levels of data collection and sale as free ones offers important information about how the "paid" model operates in practice. This model is already critiqued for privileging those with more resources by letting them "buy out" of certain exploitative practices to which others are subject.<sup>20</sup> If, in fact, the economics of this model also relies on widespread data collection and sharing, its promise as a privacy-protective alternative is questionable.

So, moreover, may be its legality. The treatment of data in ways that are not necessary for the provision of an app's service triggers the European GDPR's requirement that there be another legal basis for this processing of personal data (e.g., explicit consent).<sup>21</sup> While commentators have suggested that these and similar provisions in other recent privacy laws could prove the death knell for the free model and privilege the position of paid apps,<sup>22</sup> our data suggests that much paid app behavior also fictionalizes notions of consent, calling into question its legality.

Second, by providing empirical foundations for better understanding both corporate behavior and consumer expectations, our findings offer important insights for legal approaches to privacy protection.<sup>23</sup> Specifically, they further undermine the legitimacy of legal regimes relying on fictive "notice" and "consent" that do not reflect user understandings as bases for the collection, sale, and processing of information. At a minimum, they fortify demands for a privacy law that focuses on vindicating actual consumer expectations, and prohibiting practices that exploit them.<sup>24</sup> More broadly, they strengthen the

---

20. See, e.g., Sophia Cope & Jeremy Gillula, *AT&T is Putting a Price on Privacy. That is Outrageous*, GUARDIAN (Feb. 20, 2015), <https://www.theguardian.com/commentisfree/2015/feb/20/att-price-on-privacy> [<https://perma.cc/E6M2-8J5J>] (describing AT&T's practice of charging customers extra money for increased privacy protections).

21. Recital 43, Commission Regulation 2016/679, 2016 O.J. (L 119) (EU) 1 (repealing Directive 95/46/EC) (General Data Protection Regulation) ("Consent is presumed not to be freely given if it does not allow separate consent to be given to different personal data processing operations despite it being appropriate in the individual case, or if the performance of a contract, including the provision of a service, is made dependent on the consent despite such consent not being necessary for such performance.").

22. *Will Free Apps Soon be Dead in Europe?*, MINTZ (Feb. 9, 2016), <https://www.mintz.com/insights-center/viewpoints/2826/2016-02-will-free-apps-soon-be-dead-europe> [<https://perma.cc/S5NQ-4Z9W>].

23. See *infra* Section III.B.

24. Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 667 (2016) (identifying actions by the agency that could form a possible basis for a move in enforcement towards preventing "broken expectations of

argument for ex ante regulation of exploitative data practices where consumers are offered no opportunity for meaningful choice or consent. Amid the adoption of a new statutory privacy regime in California and increased discussion regarding omnibus federal privacy legislation, these findings can inform policymakers as they seek to implement broader privacy protections for users.<sup>25</sup>

Finally, building on our evidence that users often misunderstand technological models, to their detriment, this Article demonstrates the need for technical tools that offer transparency about app behaviors, and empower consumer choice.<sup>26</sup> Our study demonstrates that, at least in the most dominant example of a free versus paid market—mobile apps—there turns out to be no real privacy-protective option. The failures of transparency or auditability of app behaviors, moreover, mean that consumers have no way to identify instances in which paid apps might actually be limiting data collection, and actually offering more privacy-protective options. Those failures also deprive users, regulators, and law enforcement of any means to keep developers accountable. Despite the touted potential for paid models to support consumer choice and privacy protection, then, these information failures destroy any opportunity for a meaningful privacy market. Without information about app privacy practices, privacy is removed as a salient concern for users faced with a free or paid choice.

Accordingly, this work demonstrates how dynamic analysis of the type we performed in this study could serve as a tool for empowering app users. Building in such tools can allow users to go online and test, in real-time, an app's privacy behavior, revealing collection practices formerly obscured within the "black box." This technology could therefore empower users to be advocates, and inform their choices to better align their expectations with reality. More systemically, these tools could facilitate the creation of third-party certification markets, which could then test apps and label them based on their observed behaviors, thus offloading the burden from consumers. The same tools, moreover, could equip regulators, law enforcement, consumer protections organizations, and private parties seeking to remedy illegal privacy behavior through the civil system.

---

consumer privacy"); *see generally* KENNETH A. BAMBERGER & DEIRDRE K. MULLIGAN, *PRIVACY ON THE GROUND: DRIVING CORPORATE BEHAVIOR IN THE UNITED STATES AND EUROPE* 64 (2015) (discussing an evolving orientation among privacy leaders towards consumer expectations).

25. *See* California Consumer Privacy Act of 2018, CAL. CIV. CODE § 1798.100 (West 2018) (CCPA).

26. *See infra* Section IV.B.

While automation and technology are often viewed as counter-beneficial to privacy in the context of data collection, we suggest they can serve a vital function in streamlining and scaling both user decision making and regulatory enforcement. Wedding laws strengthening consumer protection with the types of privacy-enhancing technologies often overlooked by privacy regulations should be front and center in any consumer-focused legislative effort.

## II. PAYING FOR PRIVACY: OUR RESEARCH IN CONTEXT

### A. PAYING FOR PRIVACY

When Facebook Chief Operating Officer Sheryl Sandberg said in interviews last year that allowing users the option to completely opt-out of tracking and data profiling would require a “paid product,”<sup>27</sup> her comments echoed increased traction for the proposition of offering users the option of paying (or paying more) to limit the use and dissemination of their personal information.<sup>28</sup> Broadband Internet Service Providers (ISPs) have piloted offerings that permit customers to opt out of surveillance for an extra fee.<sup>29</sup> The landmark California Consumer Privacy Act (CCPA)—the furthest-reaching privacy legislation in the United States—opens the door for differential pricing based on the right to sell or share some kinds of data.<sup>30</sup>

The availability of fees from paid versions of digital services provides a substitute for targeted advertising revenues that drive the free services model.<sup>31</sup> Commentators have therefore touted it as an important means to empower

---

27. Andrew Albanese & Annie Coreno, *Are We Headed for a Pay-for-Privacy World?*, PUBLISHERS WKLY. (Apr. 6, 2018), <https://www.publishersweekly.com/pw/by-topic/industry-news/libraries/article/76530-are-we-headed-for-a-pay-for-privacy-world.html> [<https://perma.cc/XZR2-ULYS>].

28. Josh Constine, *How ad-free subscriptions could save Facebook*, TECH CRUNCH (Feb. 17, 2018), <https://techcrunch.com/2018/02/17/facebook-subscription/> [<https://perma.cc/K79Q-E8RR>].

29. See, e.g., Cope & Gillula, *supra* note 20 (describing AT&T’s piloting of a service which allows gigabit service customers to opt out of surveillance for \$29 per month).

30. Allen St. John, *How California’s New Privacy Law Could Affect You (Even If You Don’t Live There)*, CONSUMER REP. (June 29, 2018), <https://www.consumerreports.org/privacy/how-californias-new-privacy-law-could-affect-you/> [<https://perma.cc/APJ8-64KW>] (“The most controversial provision of the new law allows companies to provide a discount in exchange for the right to sell or share some kinds of data.”); see also Adam Schwartz, *The Payoff From California’s “Data Dividend” Must Be Stronger Privacy Laws*, ELECTRONIC FRONTIER FOUND. (Feb. 15, 2019), <https://www.eff.org/deeplinks/2019/02/payoff-californias-data-dividend-must-be-stronger-privacy-laws> [<https://perma.cc/8BBU-X99K>] (cautioning against moves towards “pay-for-privacy” in the CCPA).

31. See Benjamin Edelman, *Priced and Unpriced Online Markets*, 23 J. ECON. PERSP. 21, 34 (2009) (discussing the economic reality that zero pricing can be sustainable “when there are adequate profits in complementary businesses like advertising or technical support”).

consumer choice regarding the use of their personal information.<sup>32</sup> Indeed, some argue, preserving a choice between free and paid options is important in increasing access for low-income consumers, who might otherwise be priced out of services.<sup>33</sup> Especially in light of the fact that most users do not actually choose to pay for services despite their articulated privacy concerns,<sup>34</sup> and given the overall consumer surplus from online applications,<sup>35</sup> commentators argue that this is exactly the type of market-driven solution to which policymakers should defer in protecting the privacy of concerned consumers.<sup>36</sup>

At the same time, the turn to paid models as an antidote to the information abuses by providers of free services has received significant criticism. Two sets of concerns—distributional and operational—have received significant attention.

The fundamental distributional concern with relying on price-based market models (especially in place of government regulation) involves the disproportionate barriers placed on exercising a fundamental right.<sup>37</sup> As Julie Cohen noted nearly twenty years ago, “[i]f data privacy costs money—or, conversely, if surrendering privacy saves money—access to privacy will be

---

32. Omri Ben-Shahar, *Your Internet Privacy Should Be Up for Sale*, FORBES (Aug. 8, 2016), <https://www.forbes.com/sites/omribenshahar/2016/08/08/your-internet-privacy-should-be-up-for-sale/#2ec1f66d7ef2> [<https://perma.cc/CK37-HLZR>].

33. See *id.*; Thomas M. Lenard, ‘Pay-for-Privacy’ Internet Actually Benefits Low-Income Consumers, HILL (Aug. 16, 2016, 7:34 AM), <https://thehill.com/blogs/pundits-blog/technology/291549-pay-for-privacy-internet-actually-benefits-low-income-consumers> [<https://perma.cc/48QG-YKAJ>].

34. Anthony Spadafora, *Americans reluctant to pay for privacy*, TECHRADAR (Jan. 16, 2019), <https://www.techradar.com/news/americans-reluctant-to-pay-for-privacy> [<https://perma.cc/A8LD-4UGM>] (discussing a Center for Data Innovation study finding that only 27% of the surveyed would pay a monthly subscription fee in exchange for less data collection, despite the fact that 80% of respondents said they wanted online services to collect less data, and 63% of the surveyed opposed receiving free applications or services in exchange for more intensive data collection).

35. Hammock & Rubin, *supra* note 11, at 1 (“The costs of online privacy-related harm (such as identity theft) and of protective activities are small relative to the benefits from applications that are supported by online advertising, which depends on the collection of personal information.”).

36. *Id.* at 2 (“If consumers do have valid privacy concerns, markets can and do respond to them.”).

37. See Schwartz, *supra* note 30 (“Pay-for-privacy schemes undermine this fundamental right. They discourage all people from exercising their right to privacy. They also lead to unequal classes of privacy ‘haves’ and ‘have-nots,’ depending upon the income of the user.”); Elvy, *supra* note 11, at 1400 (“[U]se of this model is likely to contribute to the divide between those that can afford privacy and those that cannot.”).

more unequal than if it did not.”<sup>38</sup> Stacy-Ann Elvy’s foundational work on the pay-for-privacy model, moreover, explores the ways that this divide is particularly pernicious, in that it can lead to the collection of more data about precisely those consumers who are particularly susceptible to predatory and discriminatory behavior.<sup>39</sup>

The operational concerns derive from longstanding research into the practical and cognitive barriers skewing the market for privacy. Because of the inscrutability of the behavior of companies that collect, process, and share data—and the inability to predict and understand how it will be used in the future—consumers have little access to a true understanding of the ways that their data will be used, and the resulting personal implications (indeed, in many cases companies themselves do not know how they will use the data in the future).<sup>40</sup> Consumers, then, simply do not have the capacity to estimate the value of their own data, or the costs of handing it over.

These information asymmetries are compounded by the opacity of privacy policies,<sup>41</sup> and—especially under such conditions of uncertainty—the malleability of consumer choices in light of the framing of “consent” to data use and sharing, and of very small transaction costs, or frictions.<sup>42</sup> Moreover, even those who may wish to take more active measures to protect their privacy are often unable to succeed because of technical barriers, collective action problems, monitoring costs, third-party leakage, and data interconnectedness.<sup>43</sup> Thus, privacy—even for consumers who might feel strongly about protecting their data—might not actually be a salient feature of

---

38. Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1398 (2000).

39. See Elvy, *supra* note 11, at 1421–28 (providing examples of the ways that companies use consumer lifestyle data and data analytics in making choices about services offered to consumers and of predictive data to discriminate against individuals deemed “less valuable” or “risky”).

40. See Strandburg, *supra* note 4, at 143, 148, 150.

41. Kevin Litman-Navarro, *We Read 150 Privacy Policies. They Were an Incomprehensible Disaster*, N.Y. TIMES (June 12, 2019), <https://www.nytimes.com/interactive/2019/06/12/opinion/facebook-google-privacy-policies.html> [<https://perma.cc/2RYY-357A>]; see also Aleecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, PROCS. TECH. POL’Y RES. CONF. 565 (2008) (concluding that if the average internet user reads every word of all privacy policies they come across, the user would spend 201 hours reading, worth roughly \$3,534 annually).

42. See Susan Athey et al., *The Digital Privacy Paradox: Small Money, Small Costs, Small Talk* 12 (Stanford Inst. Econ. Policy Research Working Paper No. 17-032, 2017) <https://siepr.stanford.edu/sites/default/files/publications/17-032.pdf> [<https://perma.cc/4LBE-LYZR>] (analyzing the choices of a group of MIT undergraduates on data use and sharing); see also *supra* notes 86–90 and accompanying text (discussing framing effects).

43. See Strandburg, *supra* note 4, at 156–57, 164.

a free-or-paid choice, because it is frequently difficult for consumers to negotiate for privacy protections in the current market. This lack of transparency and choice is why analyses of “revealed preferences” are unlikely to accurately explain consumers’ true preferences.

Generally, the lack of information and the malleability of consumer choices has provided some explanation for the “privacy paradox,” by which consumers care about personal privacy, yet frequently exhibit privacy-compromising behaviors.<sup>44</sup> More specifically, it suggests the deep shortcomings of the pay-for-privacy model for offering consumers a meaningful option and an informed choice. In the absence of necessary information, consumers may rationally be unwilling to pay for more privacy because they cannot accurately value it, or assess whether or how it may be protected. Nonetheless, when consumers are given clear privacy indicators, the privacy paradox disappears: they *do* pay for increased privacy, thereby better aligning their behaviors with their stated preferences.<sup>45</sup>

Elvy’s work, moreover, raises two additional questions about the paid services model, regarding which far less empirical research exists.<sup>46</sup> First, she points out, companies using a paid model may still not actually refrain from monetizing consumer data.<sup>47</sup> Understanding actual company behavior is particularly important in the context of mobile apps, for example, that offer ad-free paid versions, but whose data collection, processing, and sharing behaviors still remain hidden from the consumer.

Second, and in light of this possibility, Elvy notes the possibility that “[p]rivacy-conscious consumers who elect to pay for” services may be misled about data practices.<sup>48</sup> Research has found consumers’ privacy choices to be highly susceptible to suggestions pre-disclosure: when treated with a very modest privacy-enhancing accommodation, subjects counter-intuitively increased disclosure incommensurately.<sup>49</sup> In light of the broader pay-for-privacy discourse, and the implicit trade-off between payment with money and payment with data, then, consumers might reasonably (and mistakenly) expect paid services to be more privacy-protective.

---

44. See Alessandro Acquisti, *Privacy in Electronic Commerce and the Economics of Immediate Gratification*, ACM ELECTRONIC COM. CONF. 21–29 (2004).

45. Janice Y. Tsai et al., *The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study*, 22.2 INFO. SYS. RES. 263–64 (2011); Serge Egelman et al., *Timing Is Everything? The Effects of Timing and Placement of Online Privacy Indicators*, SIGCHI CONF. ON HUM. FACTORS COMPUTING SYSS. 324–25 (2009).

46. See Elvy, *supra* note 11, at 1413–19.

47. See *id.* at 1419.

48. *Id.*

49. Athey et al., *supra* note 42, at 17–18.

## B. APP BEHAVIOR AND CONSUMER EXPECTATIONS STUDIES

We present two studies that fill this empirical gap, and begin to answer these two questions—whether companies using a paid model actually refrain from monetizing consumer data, and whether consumers might expect paid services to be more privacy-protective—with data from the mobile app context. The first study employs a dynamic analysis tool to track the actual privacy behaviors of paid and free apps. This tool, currently employed by regulators and watchdog groups, provides crucial information largely unavailable to consumers making privacy decisions, and—as we discuss in Part III—can provide a broader means for empowering consumers in making more informed decisions about the use and dissemination of their data. Almost half of the paid apps that we examined in this study shared the same types of personal information with the same third parties as free ad-supported versions.

The second study consists of an online survey of mobile app users. In this survey, we presented 1,000 respondents with screenshots of two apps from the Google Play Store: a “free” app and its paid counterpart. We then asked them questions about which app they would be more likely to install, as well as what differences they would expect to exist between the two versions. Our results demonstrate that consumers are significantly more likely to expect strong privacy protections when purchasing apps, as compared to installing their free ad-supported counterparts.

### 1. *Paid and Free App Behavior Study*

In this analysis, we generalize different app monetization models into two overarching categories: we define “free apps” as those that are available for download on the app store at no up-front cost; and we define “paid apps” as apps that require a one-time payment to download. Our focus is on paid apps in which the consumer pays for the app as a single discrete product, rather than for a continuously renewed service. We acknowledge that apps may employ other monetization strategies, such as the “freemium” or “paidmium” models, in which potentially recurring in-app purchases generate revenue for the developer. Though we are aware that some apps do offer in-app purchases to disable ads, these are beyond the scope of this study.

The Google Play Store does not reliably link free apps to their paid versions, or even indicate if a corresponding paid version exists at all. Therefore, we first developed our own method to identify pairs of free and paid versions of the same general app (e.g., “Quick PDF Scanner FREE” and



“Quick PDF Scanner PRO”). We evaluated and compared the behavior of these pairs using both static and dynamic analysis techniques.<sup>50</sup>


We formed our app corpus by consulting the AppCensus database,<sup>51</sup> which is regularly updated by crawling the “Top Free” charts in each of the Play Store’s categories. We then created a labeling task on Amazon Mechanical Turk. We presented workers with a free app and a list of all paid apps from the same developer, asking them to select the paid version that most closely resembled the given free app (Figure 1). In order to increase the likelihood of valid free and paid pairings, we only presented workers with free apps whose titles or package names contained the words “free” or “lite,” as those keywords would suggest that a “paid” or “full” version exists. If the free app did not have a corresponding paid version, workers were instructed to select the “Paid version does not exist” option. We presented each free app to three different workers, then manually adjudicated the responses for agreement and correctness. We paid workers \$0.10 for each match in consensus with the others, yielding a corpus of 5,877 pairs of apps.

**Figure 1: Amazon Mechanical Turk task in which participants identified the paid counterpart of the given free app**



Black and White Photo Editor

Which of the following, if any, is the paid version of the free app listed above?

☐  Bead Template Creator Premium

☐  Bitwise binary calculator

☐  KP Forecast for aurora

☐ Paid version does not exist

50. “Static analysis” refers to the examination of programs without executing them in order to rapidly detect whether they contain certain instructions or data. Static analysis often yields false positives because some detected instructions may not ever get executed in practice. “Dynamic analysis” refers to the examination of program behavior by executing it to monitor what it does, which may lead to false negatives if certain functionality is not executed during the testing period.

51. See generally APPCENSUS, *supra* note 15.

We looked for similarities across pairs of free and paid apps along three dimensions: (1) the portion of Android permissions designated by the operating system as “dangerous”—signifying that they control access to sensitive data or personally identifiable information—declared by the free app that are also declared by the paid app; (2) the portion of third-party packages (i.e., SDKs) found in the free app that are also included in the paid version; and (3) the portion of sensitive network transmissions performed by the free app also performed by the paid app. We believe these three aspects are a good representation of apps’ data collection and sharing behaviors. We employed the following methods to evaluate these:

a) Static Analysis

We used the Android Asset Packaging Tool to extract the permissions apps request for various device resources.<sup>52</sup> We then identified differences in dangerous permissions within pairs of free and paid apps. Additionally, we relied on Apktool to examine apps’ file structures for the package names that comprise the app.<sup>53</sup> We identified third-party libraries by eliminating package names that shared the same first two levels as the app package (i.e., disregarding code belonging to the core app). This revealed what third-party libraries—possibly used for monetization and data collection—are shared between free apps and their paid counterparts.

b) Dynamic Analysis

We used dynamic analysis methods derived from earlier work to automatically evaluate apps by executing them in an instrumented environment (deployed on identical Nexus 5X smartphones) that captures apps’ network traffic.<sup>54</sup> We relied on the Android SDK’s Application Exerciser Monkey tool to automatically explore apps without user intervention.<sup>55</sup> Although there is no guarantee that paired apps have identical user interfaces, we controlled for differences in app execution by providing both apps with the same random input stream at the same time. This increases the likelihood that observed differences in app behavior arose from implementation differences, rather than differences in input.

---

52. See generally *AAPT2*, ANDROID, <https://developer.android.com/studio/command-line/aapt2> [https://perma.cc/996A-NJ3N] (last visited Jan. 7, 2020).

53. See generally *A Tool for Reverse Engineering Android APK Files*, APKTOOL, <https://ibotpeaches.github.io/Apktool/> [https://perma.cc/U4F6-9ERA] (last visited Jan. 7, 2020).

54. See Irwin Reyes et al., “Won’t Somebody Think of the Children?” *Examining COPPA Compliance at Scale*, PROC. ON PRIVACY ENHANCING TECHS. 63 (2018).

55. See generally *UI/Application Exercise Monkey*, ANDROID, <https://developer.android.com/studio/test/monkey> [https://perma.cc/CGS9-K4G5] (last visited Jan. 7, 2020).

At the end of each paired execution, we analyzed the captured network data to identify which sensitive data types were sent to which remote services—services that could be for advertising, profiling, crash reporting, etc. We focused on detecting the transmission of sensitive data that can be used to uniquely track a user over time and across different services: persistent identifiers, such as the Android Advertising ID (AAID), International Mobile Equipment Identity (IMEI), and Wi-Fi MAC address; as well as personally identifiable information (PII), such as geolocation, name, and phone number. In order to detect the transmission of sensitive data, we not only used simple string matching, but also relied on methods from previous work for the decoding of obfuscated network traffic, which uses regular expressions formed from the manual inspection of different data encoding schemes.

Most current approaches to detecting suspicious application activity on mobile platforms rely on static analysis<sup>56</sup> or dynamic analysis.<sup>57</sup> However, previous approaches fall short because they either do not observe actual violations—instead only detecting when a program *might* contain violative code (in the case of static analysis), or do not scale (in the case of prior dynamic analysis approaches).

Our dynamic analysis framework allows us to monitor actual program behavior in real time and at scale. The AppCensus platform allows us to examine how often and under what circumstances apps and third-party libraries access sensitive resources guarded by permissions. By combining this infrastructure with a modified version of Lumen (previously known as Haystack),<sup>58</sup> an advanced network monitoring tool, we obtain a sophisticated holistic view of when sensitive data is accessed and where it gets sent.

## 2. Consumer Expectations Survey

In addition to examining the differences in behaviors between each free app and its paid counterpart, we also wanted to understand consumer expectations surrounding the two versions of each app. Specifically, our

---

56. See, e.g., Clint Gibler et al., *AndroidLeaks: Automatically Detecting Potential Privacy Leaks in Android Applications on a Large Scale*, PROC. TRUST (2012); Michael I. Gordon et al., *Information-Flow Analysis of Android Applications in DroidSafe*, PROC. NDSS SYMP. (2015); Jinyung Kim et al., *ScanDal: Static Analyzer for Detecting Privacy Leaks in Android Applications*, IEEE MOST (2012); Sebastian Zimmeck et al., *Automated Analysis of Privacy Requirements for Mobile Apps*, PROC. NDSS SYMP. (2017).

57. See, e.g., William Enck et al., *TaintDroid: An Information-flow Tracking System for Realtime Privacy Monitoring on Smartphones*, PROC. USENIX OSDI (2010).

58. See generally HAYSTACK PROJECT, <https://haystack.mobi> [<https://perma.cc/BT4T-NUZR>] (last visited Jan. 7, 2020) (providing an app that “analyzes mobile traffic and helps to identify privacy leaks inflicted by apps and the organizations collecting this information”).

research questions about people's expectations and beliefs about mobile app privacy when presented with a free app and its paid alternative were as follows:

- What differences do consumers expect when downloading an app for free versus purchasing it?
- Given these differences, which app would users be more likely to install?
- Do users expect different privacy behaviors from a free version and its paid counterpart?

We recruited 1,000 participants from the Prolific Academic survey platform,<sup>59</sup> limiting participation to those within the United States who successfully completed at least 95% of the previous tasks that they had undertaken. The survey took approximately five minutes to complete, for which we compensated participants \$1.00 for their time. This study was reviewed and approved by the UC Berkeley Institutional Review Board.

We conducted our study during May 2019. We piloted our study with 100 participants, and then ran the main study with 1,000 participants. Our data is drawn only from the latter 1,000 responses. Based on the pilot, we did not make any changes to the survey. Our sample was gender-balanced, with 50% self-identifying as male; the median reported age was 30, with the reported ages ranging from 18 to 76. In addition, approximately 54% of our sample had at least a bachelor's degree, and 56.5% of our sample reported themselves as single.

Our survey was composed of several sections: a mix of open-ended responses, multiple-choice questions, and five-point Likert-scale questions<sup>60</sup> (listed below)—ranging from “Definitely A” to “Definitely B,” concluding with a series of demographics questions.<sup>61</sup>

#### a) Open-Ended Questions

After obtaining participants' consent, we first presented respondents with two images of a free version of an app and its paid counterpart, controlled to have the same rating, where the key differences were in the installation price and the title of the app (“free” or “lite” for the free versions, “pro” or

---

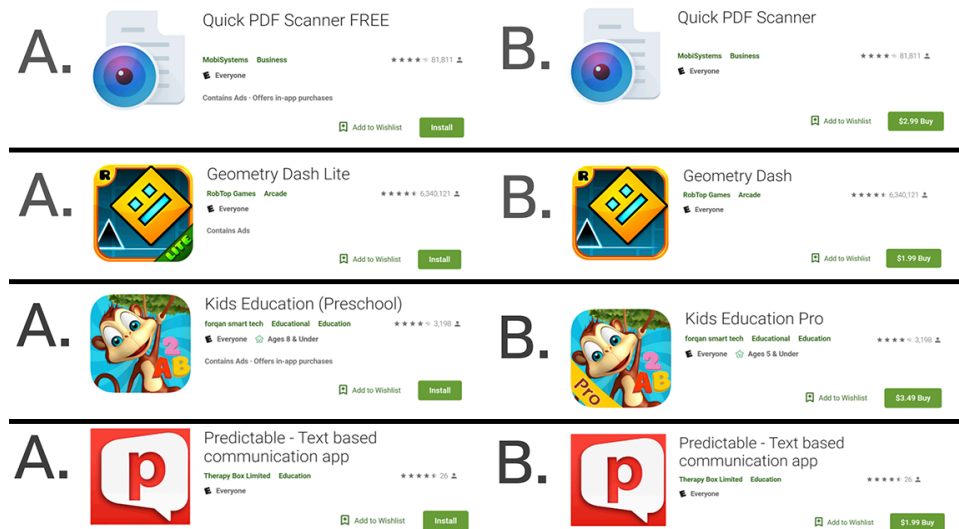
59. PROLIFIC, <https://prolific.ac> [<https://perma.cc/4JGC-QTLS>] (last visited Jan. 7, 2020).

60. In questionnaire research using Likert scales, respondents specify their level of agreement or disagreement on a symmetric agree-disagree scale for a series of statements. See Susan Jamieson, *Likert scale*, BRITANNICA, <https://www.britannica.com/topic/Likert-Scale> [<https://perma.cc/WA7Y-DEDN>] (last visited Jan. 7, 2020).

61. For detailed wording of individual questions, see *infra* Section II.B.2.b.

“premium” in the paid versions). We randomly selected a pair of apps from four possibilities (Figure 2).

**Figure 2: Participants were randomly shown one of four pairs of apps, labeled A and B**



We randomized how the two apps were presented to participants across two metrics: (1) what app they were shown (selected from four possibilities shown in Figure 2): a PDF scanning app, a game app, a children’s education app, and a text-based communication app, and (2) whether the free app was labeled as A or B. We later recoded this so that in our analysis App A always referred to the free version, while App B referred to the paid version.

After displaying the randomly-selected pair of apps, we asked participants, “In what way, if any, would you expect the above two apps to differ?” Responses to this question were collected using an open-ended text field. Two independent coders later coded these responses as binary values based on whether participants mentioned privacy or related concepts. Next, we asked participants to specify which app they would be more likely to install and why. As before, this question was coded by two independent coders based on whether concepts pertaining to privacy were mentioned.

#### b) Likert-Scale Questions

After participants answered these open-ended questions, they proceeded to the next page of the survey. The top of this page once again displayed the same pair of apps as the previous page, and asked participants to answer several Likert-scale questions using the following five-point scale: “Definitely

A (1),” “Likely A (2),” “Equally A and B (3),” “Likely B (4),” and “Definitely B (5).” The statements participants rated were as follows:

*Consider the same two apps once again. Based on the images, which app do you believe is more likely to...*

- share your data with third-party services?
- share your data with advertisers?
- share your data with law enforcement agencies?
- encrypt your data to protect it from potential breaches?
- be transparent with you about its data collection and sharing behaviors?
- comply with privacy laws and regulations?
- delete all your data from its servers after you uninstall the app?
- keep your data on their servers when no longer needed for the functionality of the app?
- have effective privacy controls (features that allow you to specify which data types you do not want the app to collect)?
- access more resources than it needs for its functionality (i.e., more permissions)?
- protect the data you gave it permission to access?

Following these Likert-scale questions, we included a few related questions for a separate study (which we do not discuss in this Article), and then concluded by collecting demographic information.

### 3. *Limitations*

While our findings (see Part III, below) show that on average, paid apps may provide fewer privacy protections than consumers expect, we note several limitations of our methodology. First, all apps were executed by randomly generating user interface events (i.e., random taps, swipes, etc.), which means that certain app functionality may not have been executed during the testing period. Thus, it is possible that under more realistic testing circumstances, some of the apps might exfiltrate data to additional third parties. Similarly, if the user interfaces between free and paid apps differ substantially, different

functionality may have been executed between the two during the testing period, confounding our results. Related to this, another confounding factor is simply due to the stochastic nature of mobile advertising: the same app executed multiple times is likely to contact multiple ad networks, which in turn load ad content from different advertisers and attribution trackers. Thus, it is possible that some pairs of free and paid apps were not observed contacting all of the same third parties for this reason. This suggests that our empirical observations may be lower bounds for privacy-invasive behaviors.

Regarding the expectations survey (see Part III, below), we observed that many participants said they would have chosen the free app despite later indicating that they believed that it would have worse privacy practices. Given that we randomly assigned apps to participants so that they could answer questions about the same free and paid pair, it is likely that many participants would not have chosen either of these apps to install under normal circumstances. Similarly, because participants were not exposed to privacy risks (nor financial costs), our results only show relative stated intentions, rather than revealed preferences. Thus, we believe that further study is needed to better understand participants' decision making between free and paid apps (as distinct from their expectations about the apps' privacy behaviors) under more realistic circumstances.

### III. FINDINGS

#### A. APP BEHAVIOR

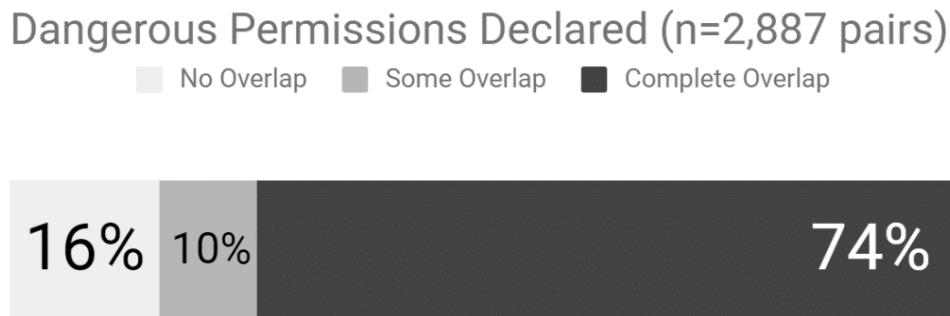
This work focuses on measurable differences in privacy between free and paid versions, so all presented comparisons are conditioned on the free app having at least one observation for any of the corresponding metrics. That is, in each of the following analyses, we disregard pairs in which the free app had no third-party packages, no permission requests, or no sensitive data shared with a third-party service, respectively. As a result, the total sample size in each analysis fluctuates slightly.

We note that there are indeed some paid apps that have observations along these dimensions that were not seen in their free counterparts. However, these represent only a small portion of our corpus: out of the 5,877 studied, 350 paid apps requested *dangerous* permissions not declared by their free versions, and 255 paid apps transmitted data not observed in the free release. We stress that our analysis quantifies the degree to which free apps' behaviors along these three metrics are carried over to their corresponding paid versions.

### 1. *Declared Android Permissions*

The Android permission system serves to protect user privacy. Apps must hold appropriate permissions to use various device resources (e.g., internet access, the camera, etc.) and access sensitive user data (e.g., phone number, location data, various persistent identifiers, etc.). A subset of Android's permissions are deemed “dangerous” because they guard sensitive resources that directly affect user security and privacy, such as the contact list or location information.<sup>62</sup> All of the resources categorized as dangerous permissions require user consent at runtime (though upon approval, the user is never prompted again).

**Figure 3: Frequency of dangerous Android permissions inherited between free and paid versions, where the free app requested at least one dangerous Android permission**



Of the 5,877 pairs in our corpus, 2,877 had free versions that declared at least one Android-defined dangerous permission. In 74% of these pairs, the paid version (Figure 3) declared all of the same dangerous permissions held by the free version. That is, paid apps held all the same access to sensitive resources as free versions in a majority of the cases where any dangerous permissions were declared. Since third-party libraries (SDKs) share the same permissions as the main app code, any third-party libraries for tracking and/or user profiling present in the paid version could have access to the same user data as the free counterparts. The most common dangerous permissions that both the paid and free versions requested were those that use disk storage shared between apps, get information about the phone's state (e.g., phone number, cellular network information, call status), and access to the device's geolocation.

62. *Protection Levels*, ANDROID, <https://developer.android.com/guide/topics/permissions/overview#normal-dangerous> [https://perma.cc/GG2Q-2YU9] (last visited Jan. 7, 2020).

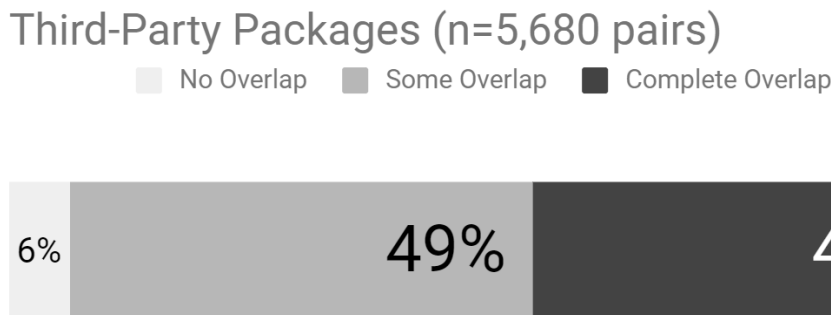


We also note that 16% of the pairs in our study had paid apps that did not request any of the dangerous permissions declared by their corresponding free versions. This suggests potential over-permissioning of free apps in these cases, in which free apps held access to dangerous permissions that may not have been necessary for those apps' core functionality. Overall, this implies that free apps will likely have access to permissions that the paid app does not, putting users' privacy at higher risk by requesting permissions that are unnecessary to the apps' core functionality.

## 2. *Bundled Third-Party Packages*

It is common practice in software engineering to use third-party code to expedite development. That is, developers do not need to “reinvent the wheel” and can instead integrate functionality into their programs written by others. In mobile apps, third-party libraries allow for pre-built functionality like graphics rendering, advertising, and analytics, among others. Third-party code bundled in apps has the same privileges as the host app, and can access all the same device resources and personal data available to the host app.

**Figure 4: Frequency of third-party package reuse among free and paid pairs, where the free app had at least one third-party package**



Of the 5,877 pairs in our corpus, 5,680 had at least one third-party package in the free version. Of these pairs, as Figure 4 shows, 45% of paid apps contained the same third-party libraries as the free versions, while 6% of paid apps showed no third-party libraries carried over from their free versions. The remaining 49% of paid apps had varying degrees of third-party library reuse from the free version to the paid version. This data suggests that paid apps are likely to contain most, if not all, of the same third-party libraries as the free versions. Although we acknowledge that our analysis did not account for third-party libraries included but not actually executed (i.e., dead code), these results show that developers of paid apps have little motivation to remove externally-

produced code in paid apps. This may leave paying consumers exposed to the same potential for third-party data collection as found in free apps.

Based upon the library categorizations of LibRadar,<sup>63</sup> we analyzed the types of third-party libraries present in free and paid versions of apps, focusing our attention on libraries labeled as “Advertising” and “Analytics”; some of the common libraries categorized as Advertising included advertising companies such as *Unity*, *AppLovin*, Google’s *AdMob*, and *Chartboost*.

Focusing on advertising libraries specifically, LibRadar detected at least one ad library present in either the free or paid release (or both) in 3,043 pairs. Of these, 2,918 free apps contained ad libraries, while 1,320 paid apps contained ad libraries. Furthermore, 209 paid apps even bundled at least one advertising library that was not present in its free counterpart, suggesting that some paid apps will not only share some of the same advertising libraries included in the free version, but also introduce new ones. Thus, although ad libraries are certainly monetizing most free apps, paying for an app only reduces the likelihood of encountering ad software by half.

### 3. *Network Transmissions*

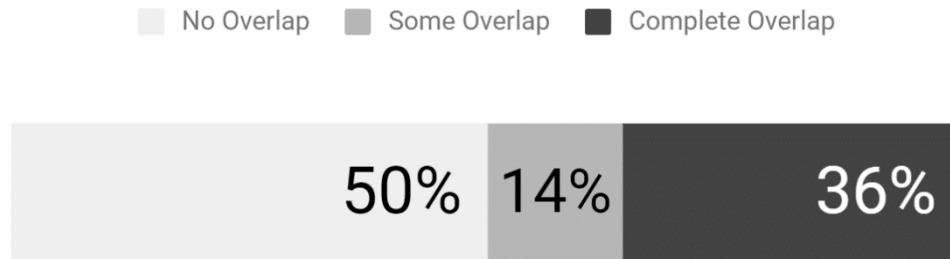
Third-party libraries bundled in apps routinely collect various data from users and their devices, sending it back to the app companies’ servers. For example, crash reporting services often gather hardware specifications and usage telemetry to help developers debug their apps, while advertising networks collect persistent identifiers and personal information to better target users with relevant ads. By observing all of the network traffic associated with an app, we can discern the types of sensitive data being transmitted (e.g., Android Advertising ID, e-mail addresses, geolocation information, etc.) and the recipient of that data.

---

63. Ziang Ma et al., *LibRadar: Fast and Accurate Detection of Third-party Libraries in Android Apps*, 2016 IEEE/ACM 38TH INT’L CONF. ON SOFTWARE ENGINEERING COMPANION (ICSE-C) 653 (describing the function and method of LibRadar).

**Figure 5: Frequency of unique domain destinations shared between free/paid pairs, where the free app transmitted sensitive data to at least one domain**

### Destinations with Sensitive Data (n=1,599 pairs)



Among the 5,877 pairs of apps that we examined, 1,599 pairs' free version transmitted sensitive data to online services over the internet. Out of these 1,599 pairs, we observed that 50% of these pairs' paid versions (Figure 5) did not communicate with any of the domains that the free version did, while 14% shared some destinations with the free version. Conversely, 36% of these pairs' paid versions communicated with all of the same domains as the free version. We found that overall, the most frequently-observed sensitive data types shared by both free and paid apps were indeed those that enable persistent tracking, such as Advertising ID (651 pairs), Android ID (570 pairs), device IMEI (65 pairs), and location (39 pairs).

#### B. CONSUMER EXPECTATIONS SURVEY DATA

We directly tested the null hypothesis that consumers are likely to believe that free and paid versions of the same app offer the same privacy and security protections. To test this hypothesis, we constructed a survey with a mix of open-ended questions, multiple-choice questions, and five-point Likert-scale questions, concluding with a series of demographics questions. We observed that when provided with an open-ended question about the differences between the free and paid versions of an app, few participants mentioned privacy unprompted. This suggests that privacy behaviors may not be their *primary* consideration. However, when explicitly asked to compare the privacy behaviors of the two app versions, we were able to reject the null hypothesis. This indicates that while privacy differences may not be among the participants' primary considerations, they are nonetheless an important secondary consideration for a significant proportion of study participants.

### 1. *Open-Ended Questions*

#### a) Expected Differences

To probe further into consumers' expectations of free and paid app behavior, we began with an open-ended question to avoid priming. After being presented with a free version of an app A, and its paid counterpart, app B, respondents were asked, "In what way, if any, would you expect the above two apps to differ?" Approximately half of the responses (49.4%) mentioned the inclusion or exclusion of ads between versions. One participant stated, "The free one will have ads and the paid one will not" (P77). Another wrote, "Option A will have ads, but I'm not paying for it (it will likely be very annoying). Option B will not, but I would have to pay \$3 up front" (P857).

Many responses mentioned differences in app features (48.1%). "The first will not have all the features of the second," one participant stated (P152). Another suggested the possibility of upgrades: "The first one would be free to install with limited features, but could be upgraded by paying" (P146).

However, few participants—without being primed to privacy/security—mentioned security and privacy differences between the versions (1%): "I think that the B app would have less intrusive permission requests than the A app would. The B app would be more trustworthy than the A app" (P457). Others even suspected malicious intent as a difference between the app versions, stating, "I might worry that A has a higher risk of viruses upon download, but otherwise they seem the same" (P174). Similarly, another wrote, "App A will be ad-supported and unlocking the full features would require payments. I would also be more concerned about it having spyware / malware aspects" (P287). These responses suggest that while participants think about the presence of ads unprompted, many do not immediately jump to the privacy implications of those ads.

#### b) User Preference

We found that most users (81.6%) would be more likely to install the free version of an app over its paid counterpart. To avoid priming, we followed up with an open-ended question: **why?** Of the participants that would be more likely to install the free version, the most prevalent reasoning (58.5%) was simply that the app was free. One participant even noted their willingness to trade security for price, stating, "I hate spending money on apps, especially if the service they offer is simple. So even though there's a higher chance of a virus for app A, I would download it to save the money" (P174). Echoing this sentiment, another participant wrote, "Because I'm broke, and while I suspect the free app will harvest data beyond what would be appropriate for its function, I don't expect any better of the paid app" (P167). Another common

reason was that users would prefer a low-risk option to try an app before committing to purchasing the paid version based on perceived usefulness, quality, and desire to support the developer. Thus, it is likely that the responses to this hypothetical question were influenced by the fact that we were asking participants to choose between two versions of an app, when in reality they may have had little interest in installing either.

However, of the participants that were more likely to state that they would purchase and install the paid version of the app, the most common reason (30.3%) given in the responses was the removal of advertisements. Many users mentioned how advertisements adversely affected the overall user experience, some even describing the advertisements as “annoying” or “disruptive.” Almost 6% of the participants who preferred to pay for the app were more likely to say they would purchase it because of perceived privacy and security risks in free apps. A participant wrote, “I would review it for permissions it wants and any customer complaints of spyware, but generally paid apps are safer and I would want the full set of features right away” (P296). Another even stated, “[the paid app] would be less susceptible to security breaches and data mining” (P457). In addition, the responses revealed another facet of consumer preference—advertising to children. Out of the 103 participants who were randomly assigned children’s education apps as the focus of the survey, many (28.2%) of the participants expressed that they would purchase the paid version of the app because they would not want advertisements displayed to their children. One participant even made the distinction between their purchasing preferences based on if the advertisements were directed at children or not, stating, “[the app] looks like a child’s app and I would want my child not to be bothered by ads. Even if it were for me, I might chose [sic] the ad free version” (P894).

## 2. *Expectations About Privacy Behaviors*

For a quantitative perspective on users’ expectations on the privacy behaviors between free and paid versions of an app, we asked participants to evaluate on a Likert scale the differences between the apps based on the provided statements.<sup>64</sup> Overall, while the price of an app did not have any observable effect on whether participants believed an app was likely to request more information than it actually needed to function ( $p = 0.68$ ), we did find that participants were more likely to expect that the free version would share their data with advertisers ( $p < 0.0001$ ) and law enforcement agencies ( $p <$

---

64. See *infra* Section II(B)(2) for scale and provided statements; all comparisons were made using the one-sample Wilcoxon signed rank test to evaluate the observed data against the null hypothesis.

0.0001). In addition to this, users were also more likely to expect the free version to keep their data on the app's servers when no longer needed for the functionality of the app ( $p < 0.0001$ ).

Paralleling this theme, users were more likely to expect the paid version to encrypt their data, protecting it from potential breaches ( $p < 0.0001$ ), and comply with privacy laws and regulations ( $p < 0.0001$ ). Similarly, participants were also more likely to expect the paid version to both protect the data they gave the app permission to access ( $p < 0.0001$ ) and delete all their data from its servers after they uninstall the app ( $p < 0.0001$ ). Not only were they more likely to expect more privacy-preserving behaviors from the paid app in practice, but they also were more likely to expect a higher level of transparency from the paid version with regard to its data collection and sharing behaviors ( $p < 0.0001$ ) and more granular control over the collection of their data ( $p < 0.0001$ ). This suggests that the mere act of paying, regardless of how much, is associated with receiving better privacy.

#### IV. IMPLICATIONS FOR PRIVACY PROTECTION

Our research shows that while consumers expect that paid apps are likely to have better privacy and security practices than their free counterparts, those expectations may not comport with reality. Specifically, almost half of the paid apps that we examined shared the same types of personal information with the same third parties as the free ad-supported versions. Worse, there was no obvious way for a consumer to understand when paying for an app was likely to lead to better privacy and when it was not.

These findings have important implications for our understandings of the ways that users develop expectations about the privacy behaviors of digital services, and for law and policy intended to protect personal data. Specifically, they underscore existing research about the ways that consumers develop understandings (and misunderstandings) about data usage, and the ways in which regulatory regimes purportedly premised on consent can thwart user intentions and vitiate that notion.

More specifically, they point to a wide divergence between expectations and reality and provide important evidence about the pervasive and unanticipated collection and sharing of data. While some of that behavior might be captured in evolving legal regimes, notably the GDPR, much is not. This failure suggests the importance of expanding regulatory approaches that on one hand reflect the ways that inaccurate expectations might lead to user deception, while on the other hand recognize that consumer confusion can vitiate notions of consumer choice by eliminating privacy as a salient factor in consumer decision making altogether.

Finally, our research suggests the important role that the tools developed in our research can play in empowering and protecting users by making app behavior more transparent. These tools can reduce the opacity of app behavior and information asymmetries that corrupt the market for privacy. This could better align consumers' expectations with the actual behaviors of the apps they use, increasing the possibility that consumers might have the capacity to make meaningful choices about the digital services they use and the information they share. It also might clarify for policymakers and citizens ways that consumer choice fails as a means to govern information use meaningfully, and the contexts in which repairing market failures and addressing market abuses will require regulation.

A. INSIGHTS FOR LAW AND POLICY

In this light, our findings point to a number of important implications for law and policy.

1. *Meaningful Consent, and the GDPR*

As an initial matter, the ways that paid apps frequently collect and share personal data likely run afoul of privacy initiatives focused on ensuring that consumers be sufficiently informed about information practices such that consent to use their data is meaningful—notably the GDPR.<sup>65</sup> The finding that paid apps frequently share data in similar ways as their free counterparts indicates that, where free apps would violate the European law—insofar as they have no legal basis for the processing of personal data—many paid apps would as well. Commentators have predicted that the GDPR's presumption against consent as a legal basis for data processing when the data use in question is not “necessary” for the provision of a service will prohibit many of the information practices engaged in by free services, ending the “Internet's Grand Bargain” by which data is traded for services, and leading to a widespread replacement of free offerings with those that will require payment. Yet paid apps that engage in the behavior we documented would also violate such requirements.

---

65. See also California Consumer Privacy Act Regulations, Proposed Text of Regulations § 999.305(a)(3) (Oct. 20, 2019), <https://oag.ca.gov/sites/all/files/agweb/pdfs/privacy/ccpa-proposed-regs.pdf> [<https://perma.cc/HM8G-L3QW>] (“If the business intends to use a consumer's personal information for a purpose that was not previously disclosed to the consumer in the notice at collection, the business shall directly notify the consumer of this new use and obtain explicit consent from the consumer to use it for this new purpose.”).

## 2. *Consumer Expectations and Privacy Enforcement*

The divergence our research demonstrates between user expectations and the privacy and security of the apps they use further underscores the shortcomings of legal regimes that purport to rely on consent as the basis for privacy protection, yet fail to contend with the barriers to accurate consumer understandings of the ways technology implicates personal privacy. In the case of the apps we studied, both free and paid, users simply had no capacity to pierce the opacity of app behavior, and the apps themselves provided few clues to understand the relative levels of privacy protection necessary for informed consent.

Our findings resonate with broader research into the ways that users develop their (frequently inaccurate or incomplete)<sup>66</sup> understandings of technology, and their interactions with it. The data behaviors of digital services providers are largely black boxes, preventing users from understanding the details of the ways that their information is collected, used, and shared. Privacy policies, moreover, are often deceptive and misleading.<sup>67</sup> Even when they do reflect accurate data practices, they are lengthy, often inscrutable, and—despite their length—often fail to disclose behaviors with sufficient granularity as to provide users with material information.<sup>68</sup> Even disclosures required by legal regimes seeking to mandate notice with sufficient personalization and specificity to warn consumers about problematic information practices have been plagued with problems of ambiguity that result in confusion about whether the recipient of the communication was at risk and should take action.<sup>69</sup>

Especially relevant in such uncertain contexts, consumers, constrained by human limits on attention and cognition, are “boundedly rational”

---

66. Lindsey Barrett, *Model(ing) Privacy: Empirical Approaches to Privacy Law and Governance*, 35 SANTA CLARA HIGH TECH. L.J. 1, 17–20 (2018) (summarizing studies that find, *inter alia*, that most consumers don’t understand basic facts including what privacy policies are, that applications continue to run in the background when the user is not directly engaged with them, and that an app can still access their information when not in use).

67. See Ehimare Okoyomon et al., *On The Ridiculousness of Notice and Consent: Contradictions in App Privacy Policies*, IEEE WORKSHOP ON TECH. & CONSUMER PROTECTION (2019), <https://www.ieee-security.org/TC/SPW2019/ConPro/papers/okoyomon-conpro19.pdf> [<https://perma.cc/YX3B-JPK3>] (discussing the gaps between disclosed data collection practices as articulated in privacy policies, and de facto data collection practices as observed using dynamic analysis tools).

68. Barrett, *supra* note 66, at 17–18 (“Few people read privacy policies, and those who do are left with little basis to understand the uses of their data.”).

69. Yixin Zou et al., *You ‘Might’ Be Affected*, PROC. 2019 CHI CONF. ON HUM. FACTORS IN COMPUTING SYSS. 11 (2019) (presenting a study concluding that 97% of sampled data breach notifications were difficult or fairly difficult to read based on readability metrics).



decisionmakers.<sup>70</sup> Even if they had the capacity, they simply do not have the incentive to invest the time required to discern and evaluate all the terms of an agreement,<sup>71</sup> nor would it be rational for them to do so.<sup>72</sup> This is especially true—and creates particular opportunities for consumer exploitation—when issues are “nonsalient” to consumer decisions to engage in a transaction. Data use and privacy usually are “nonsalient,” in that consumers face a “lack of meaningful choice” about them, or they are “hidden,” or “unduly complex.”<sup>73</sup> Put differently, they are not policed by the market because they do not impact the decision making of consumers, who lack the ability to evaluate them.

Without access to reliable knowledge about app behavior, most users we surveyed understood payment for an app to suggest improved data privacy and security practices. This finding underscores research in the field of computer-human interaction that has begun to identify ways in which, in the absence of easily-accessible understandings, consumer expectations about technology instead result from “folk theories”: intuitive causal explanations that people construct to explain the world. Drawing on whatever clues are available from the framing of the technology, the discourse around it, and their interface with it, users develop “non-authoritative conceptions of the world” that can diverge significantly from the designers’ views regarding—and the reality about—what a technology system is, and how it works.<sup>74</sup> These understandings, in turn, can

---

70. Russell Korobkin, *Bounded Rationality, Standard Form Contracts, and Unconscionability*, 70 U. CHI. L. REV. 1203, 1204–06 (2003). In fact, it is because of the bounded rationality of consumers that the “market . . . will often include terms that are socially inefficient, leav[ing] buyers as a class worse off.” *Id.* at 1206.

71. Aaron Perzanowski & Chris Jay Hoofnagle, *What We Buy When We Buy Now*, 165 U. PA. L. REV. 315, 323–27 (2017) (demonstrating that consumers lack an understanding of what it is they are buying when purchasing online digital media); Obar & Oeldorf-Hirsch, *supra* note 6, at 1 (finding that participants who joined a fictitious social network spent fifty-one seconds on average reading the Terms of Service, with a 93% acceptance rate, and 98% of participants missed the intentional “gotcha clauses” like the assignment of their first-born child).

72. *See, e.g.*, Obar & Oeldorf-Hirsch, *supra* note 6, at 1, 7 (finding that 98% of those agreeing to conditions of using a fictitious social network platform missed the intentional “gotcha clauses” the researchers implemented in the terms specifically mentioning users’ data will be shared for the purpose of assessing eligibility for “employment, financial service (bank loans, insurance, etc.), university entrance, international travel, the criminal justice system, etc.” and that users’ first-born child will be assigned to the platform provided as payment for accessing the network).

73. *See* Amit Elazari Bar On, *Unconscionability 2.0 and the IP Boilerplate: A Revised Doctrine of Unconscionability for the Information Age*, 34 BERKELEY TECH. L.J. 567 (2019) (discussing standards for salience).

74. *See* Motahhare Eslami et al., *First I “Like” It, Then I Hide It: Folk Theories of Social Feeds*, PROC. CHI CONF. ON HUM. FACTORS COMPUTING SYSS. 2372–73 (2016) (discussing the development of folk theories about how the algorithms driving social media feeds operate, arising from “seams” in the system that are visible to users); *see also* Benjamin Toff & Rasmus

result in “inaccurate understandings” of how technology systems work, “mismatches between designer and user intent,” and “expectation violation.”<sup>75</sup> Users rely on such folk models, moreover, to justify ignoring expert advice, reinforcing behaviors that increase data vulnerability and exploitation.<sup>76</sup>

Relatedly, it has been demonstrated that consumers’ expectations regarding digital services’ use of their personal data, and consequent privacy choices, are shaped by their perceptions about the trustworthiness of those services,<sup>77</sup> foregrounding the question of ways that trust is generated in disclosure settings, and the relevance to these impressions of a “paid” versus “free” distinction.<sup>78</sup> In the context of the choice between free and paid digital services in particular, one recent study identified trust that the features would deliver privacy benefits, such as reduced data exploitation, as a significant influence on a customer’s attitude toward paying for the premium version.<sup>79</sup>

The divergence we demonstrate between app behavior and user understandings, then, further points to the importance of regulatory efforts targeted at vindicating consumer expectations and preventing their abuse. Daniel Solove and Woodrow Hartzog have documented ways that the FTC’s enforcement activity has laid the foundations for a robust privacy regulatory

---

Kleis Nielsen, “I Just Google It”: *Folk Theories of Distributed Discovery*, 68 J. COMM. 636 (2018) (identifying folk theories about the way news reaches consumers through digital platforms, and the way that shapes engagement with public affairs); Motahhare Eslami et al., *User Attitudes towards Algorithmic Opacity and Transparency in Online Reviewing Platforms*, PROC. CHI CONF. ON HUM. FACTORS COMPUTING SYSS. (2019) (surveying various users’ attitudes and “folk theories” concerning the operation of algorithms on the Yelp platform).

75. Michael A. DeVito et al., “Algorithms Ruin Everything”: #RIPTwitter, *Folk Theories, and Resistance to Algorithmic Change in Social Media*, PROC. CHI CONF. ON HUM. FACTORS COMPUTING SYSS. 3163 (2017).

76. See Rick Wash, *Folk Models of Home Computer Security*, PROC. SYMP. ON USABLE SECURITY & PRIVACY (2010) (identifying eight ‘folk models’ of security threats used by home computer users, and how these models are used to justify ignoring expert security advice).

77. See, e.g., Valentina Bali, *Tinkering Toward a National Identification System: An Experiment on Policy Attitudes*, 37 POL. STUD. J. 233, 250 (2009) (highlighting the role of trust in government institutions when it came to determining respondents’ concerns over personal identification); Kirsten Martin, *Transaction Costs, Privacy, and Trust: The Laudable Goals and Ultimate Failure of Notice and Choice to Respect Privacy Online*, 18 FIRST MONDAY (2013) (discussing methods to develop trust as an alternative to notice in protecting privacy).

78. See Patricia A. Norberg et al., *The Privacy Paradox: Personal Information Disclosure Intentions Versus Behaviors*, 41 J. CONSUMER AFF. 100, 119 (2007) (“Future research, on a practical level, must examine how factors such as the physical environment, the media of data collection and responses to human interaction impact our assessment of trust and risk.”).

79. Michel Schreiner & Thomas Hess, *Why Are Consumers Willing to Pay for Privacy? An Application of the Privacy-freemium Model to Media Companies*, 164 EUR. CONF. ON INFO. SYS. COMPLETED RES. PAPERS 5, 5–6, 12–13 (2015) (surveying German Facebook users regarding their willingness-to-pay for a premium version of Facebook with increased privacy).

regime prohibiting broken consumer expectations of privacy,<sup>80</sup> pursuant to their authority to prevent deceptive acts under Section 5 of the agency's enabling act.<sup>81</sup> While the Commission's enforcement actions have generally focused on "broken promises" in the form of violations of privacy policies and explicit representations, Solove and Hartzog have urged an expansion of the agency's strategy to prohibit company behaviors that have deceptive *effects*, taking into account expectations reflecting broader context.<sup>82</sup> Our findings document the existence of these deceptive effects based on the gap between consumers' expectations and reality, highlighting a need to address this issue, whether through the FTC's "common law" privacy jurisprudence, or through broader privacy legislation.

### 3. *Risk Salience, Transactional Salience, and Privacy Protection*

Finally, the considerable consumer confusion in the face of the complete app behavior opacity reflected in our findings might suggest taking bolder regulatory steps. Reliance on expectations as a legal backstop against privacy-intrusive behaviors is problematic enough in the face of increasingly widespread technical capacity on the one hand,<sup>83</sup> and market constraints on the options offered to consumers on the other.<sup>84</sup> Looking to unfounded expectations to set the appropriate boundaries of privacy protection is a sham. Alternatives could take the form of actively policing elements of user-app transactions in both the paid and free context that may be deceptive or exploitative, deeming them unconscionable.<sup>85</sup> Relatedly, they could manifest in

---

80. Solove & Hartzog, *supra* note 24, at 667 (identifying actions by the agency that could form a possible basis for a move in enforcement towards preventing "broken expectations of consumer privacy"); *see also* CHRIS JAY HOOFNAGLE, FEDERAL TRADE COMMISSION PRIVACY LAW AND POLICY 123–25 (2016) (discussing FTC use of surveys on consumer understandings, so as to better understand how consumers perceive statements or representations made to them by businesses); BAMBERGER & MULLIGAN, *supra* note 24, at 183–196 (detailing ways that the FTC has sought, through a variety of "soft" and "hard" regulatory approaches, to link legal standards to consumer expectations).

81. 15 U.S.C. § 45(a)(1) (declaring unlawful "unfair or deceptive acts or practices in or affecting commerce").

82. Solove & Hartzog, *supra* note 24.

83. *See generally* David Gray & Danielle Citron, *The Right to Quantitative Privacy*, 98 MINN. L. REV. 62, 127 (2013) (calling such an approach to the definition of "reasonable expectation" of privacy in the Fourth Amendment "technological determinism run amok").

84. *See* Kenneth A. Bamberger & Orly Lobel, *Platform Market Power*, 32 BERKELEY TECH. L.J. 1051, 1089 (2017) (discussing the ways that technology platform market power can restrict users' privacy choices).

85. *See* Bar On, *Unconscionability 2.0*, *supra* note 73, at 625 (urging the application of the unconscionability doctrine to boilerplate terms in technology transactions in light of the newly revised Restatement of The Law Consumer Contracts, Council Tentative Draft, which

the conclusion that reliance on consumer choice is ineffective against certain information collection, use, and sharing practices, which must be curbed instead by direct substantive prohibition.

a) Privacy Salience and Privacy Design

An extensive body of empirical literature has identified privacy's "salience"—its prominence in a person's awareness at the time they are faced with a privacy decision—as an important element in shaping whether or not that individual makes more or less privacy-protective choices.<sup>86</sup> Accordingly, users' privacy preferences are not stable and coherent, but rather highly dependent on context.<sup>87</sup> When users are primed regarding privacy concerns, they are less likely to disclose data.<sup>88</sup> Moreover, timing matters. In-app dialogs increase salience more than those shown before an app's installation,<sup>89</sup> and even a fifteen-second delay between data use disclosures and the relevant decision can generate measurable differences in privacy-protective behavior.<sup>90</sup>

Along with informational asymmetries between users and tech companies, cognitive limitations on individual ability to process privacy policies and fully understand data use and other decisional biases that discount risk,<sup>91</sup> the phenomenon of risk salience, and the resulting manipulability of consumer decisions it allows, has provided an explanation for the "privacy paradox" by

---

addresses the one-sidedness of a term that unreasonably undermines "the consumer's benefit from the bargain").

86. Meredydd Williams et al., *Privacy Salience: Taxonomies and Research Opportunities*, IFIP INT'L SUMMER SCH. ON PRIVACY & IDENTITY MGMT. 263, 263–278 (summarizing research and defining privacy salience "as whether an individual is currently considering the topic of informational privacy").

87. Leslie K. John et al., *Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information*, 37 J. CONST. RES. 858, 858–59 (2011); see also HELEN NISSENBAUM, *PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE* (2010) (discussing the role of context in privacy attitudes and behaviors).

88. John, *supra* note 87 (presenting four studies); Hazim Almuhiemedi et al., *Your Location Has Been Shared 5,398 Times! A Field Study on Mobile App Privacy Nudging*, 6 PROC. ACM CONF. ON HUM. FACTORS COMPUTING SYSS. 787 (2015) (providing real-time information about lax app data sharing practices prompted over half of studied users to change permissions).

89. Rebecca Balebako et al., *The Impact of Timing on the Salience of Smartphone App Privacy Notices*, PROC. ACM CCS WORKSHOP ON SECURITY & PRIVACY SMARTPHONES & MOBILE DEVICES 63 (2015).

90. Idris Adjerid et al., *Sleights of Privacy: Framing, Disclosures, and the Limits of Transparency*, PROC. NINTH SYMP. ON USABLE PRIVACY & SECURITY 2 (2013).

91. Alessandro Acquisti & Jens Grossklags, *What Can Behavioral Economics Teach Us About Privacy?*, in *DIGITAL PRIVACY: THEORY, TECHNOLOGIES AND PRACTICES* (Alessandro Acquisti et al., eds., 2007) (discussing the roles of information asymmetry and bounded cognition in explaining the privacy paradox).

which consumers behave in ways that undermine their stated privacy commitments and concerns.<sup>92</sup>

The findings of our expectations survey comport with these insights. Before being primed to the issue of privacy, only 1% of our respondents mentioned privacy or security as elements in which they would expect the paid and free versions of an app to differ. Yet when asked directly, over half indicated their belief that there would be a difference.

These findings suggest the importance of purposive policy efforts to build privacy “nudges” into the design of user interfaces and default configurations to assist users in overcoming hurdles to meaningful privacy choice,<sup>93</sup> especially in the mobile context in which multiple parties are often involved in data collection, and the small screen size presents display challenges.

b) Transactional Salience and Substantive Privacy Regulation

The combination of salience effects with other causes of the “privacy paradox,” moreover, suggests that another form of “non-salience” might also be at work—the non-salience of privacy as an element to the prospective bargain entered into by the vast majority of users surveyed—whether they choose paid or free services. Indeed, among our group, before being primed to think about privacy and security issues, only 6% cited them as motivators for their version choice.

Given the hurdles to accurate user comprehension about data practices, the opacity of actual app behaviors, and the way users shape expectations based on folk theories uninformed by necessary information, moreover, the non-salience of privacy (and to what practices they were “consenting”) is hardly surprising. Armed with only unsubstantiated and largely inaccurate intuitions about the behavior of apps that provide neither transparency nor

---

92. Meredydd Williams et al., *The Perfect Storm: The Privacy Paradox and the Internet-of-Things*, INT’L CONF. ON AVAILABILITY, RELIABILITY & SECURITY 2, 2–4 (2016) (discussing the roles of risk salience, user interface design, and default configurations in explaining the privacy paradox).

93. Alessandro Acquisti et al., *Nudges for Privacy and Security: Understanding and Assisting Users’ Choices Online*, 50 ACM COMPUTING SURVS. 1 (2017) (discussing research regarding design choices to overcome decision-making hurdles affecting individuals’ choices in the presence of privacy and information security tradeoffs); FED. TRADE COMM’N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE, RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS 59–60 (2012), <https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf> [<https://perma.cc/YF9V-M62M>] (recommending just-in-time disclosures and the obtaining of affirmative express consent before allowing apps to access sensitive content such as geolocation information through APIs).

clues to their treatment of data, users had no way of reliably factoring privacy into the choice before them.

These findings resonate with the development of the notion of “salience” in the context of standard-form contracts (often termed “contracts of adhesion”) more broadly. That context offers tools for preventing opaque app behavior that in practice exploits consumers, without relying on fictive consumer choice. Such contracts, including shrink-wrap licenses, software End User License Agreements (EULAs), and digital-platform Terms of Service, arise in contexts in which consumers face similar challenges to understanding inscrutable disclosures and possess no incentive to invest the time required to understand and evaluate terms—including those regarding privacy and data use—and no ability to negotiate them.<sup>94</sup>

Taking into account the actual hurdles faced by consumers in these contexts, scholars and policymakers informed by behavioral understandings have argued that because non-salient terms—those that, in the words of the draft Restatement of The Law of Consumer Contracts currently under consideration by the American Law Institute, do not “affect consumers’ contracting decisions,”<sup>95</sup> and are therefore not policed by market negotiation. They must thus be viewed with suspicion and either policed *ex post* by courts, or *ex ante* by legislation.<sup>96</sup>

In that vein, a recent complaint brought by the Los Angeles County Attorney under the California State Unfair Competition Law against a popular mobile weather app urged a court to disregard the company’s privacy policy in sanctioning its behavior, because the app provided in-app disclosures that inaccurately presented users with contradictory privacy information, and these in-app disclosures were much more likely to be read by users. Users, the government argued, “have no reason to seek [geolocation data collection] information by combing through the app’s lengthy [privacy policy], buried within which are opaque discussions of [the developer’s] potential

---

94. See generally Bar On, *Unconscionability 2.0*, *supra* note 73.

95. RESTATEMENT OF THE LAW CONSUMER CONTRACTS, COUNCIL DRAFT NO. 5 at 95 (AM. LAW INST. 2018) [https://www.ali.org/media/filer\\_public/05/30/053007a1-2b37-4142-b9c3-7a881e847d50/consumer\\_contracts\\_-\\_td\\_-\\_online.pdf](https://www.ali.org/media/filer_public/05/30/053007a1-2b37-4142-b9c3-7a881e847d50/consumer_contracts_-_td_-_online.pdf) [<https://perma.cc/U8DK-G4CH>]. The draft Restatement explains, in discussing the standards for consumer consent, that the “concept of salience underlies the metrics regularly used to determine whether a contract term is unconscionable.” *Id.* Thus consumers’ consent is vitiated when they face a “lack of meaningful choice” (if a term was non-salient because it did not “affect consumers’ contracting decisions”) or when a term constitutes an “unfair surprise,” was “hidden,” was “unduly complex,” or resulted from “uneven bargaining power”—and these tests “are either synonymous with, or direct results of, nonsalience.” *Id.*; see Bar On, *Unconscionability 2.0*, *supra* note 73 (discussing the role of salience in the unconscionability doctrines).

96. See Korobkin, *supra* note 70, at 1204–06.

transmission of geolocation data to third parties and use for additional commercial purposes.”<sup>97</sup> Indeed, the complaint recognized, “the vast majority of users do not read those sections at all,”<sup>98</sup> effectively invoking the principle of “salience.”

Where, as in the case of our findings, expectations diverge widely from reality, and mask pervasive and unanticipated collection and sharing of data, regulatory approaches must better reflect the ways that inaccurate expectations might lead to user deception. Regulatory approaches must also reflect the ways that structural and cognitive barriers can vitiate notions of consumer choice by eliminating privacy as a salient factor in consumer decision making altogether. In those contexts, notice and consent is merely a façade held up by an untethered fantasy of a rational, informed, and empowered consumer.

B. OPPORTUNITIES FOR CONSUMER EMPOWERMENT AND ENHANCED OVERSIGHT

Finally, the effectiveness of our tool in piercing the opacity of app data collection practices suggests the promise of technical mechanisms that can foster transparency, increase salience, and empower users’ decision making at scale by unmasking complex processes for empowering consumers and policing privacy-preserving design. While research often promotes auditing within the domain of technical experts,<sup>99</sup> by fostering dynamic analysis tools with accessible user-interfaces, such as the AppCensus tool, auditing could be scaled, and even crowd-sourced, to allow the average user to uncover data abuse practices.<sup>100</sup> Such tools could also enable the creation of a market for third-party assurance privacy seal and certification programs to set standards and enable companies to demonstrate privacy accountability and compliance.<sup>101</sup> Using dynamic analysis tools, these programs could centralize the testing of apps and label them based on their observed behaviors, relieving

---

97. Complaint at 3, *California v. TWC Prod. & Tech., LLC* (Cal. Super., Jan. 3, 2019), <https://int.nyt.com/data/documenthelper/554-l-a-weather-app-location/8980fd9af72915412e31/optimized/full.pdf> [<https://perma.cc/ZFY6-ALB9>] (seeking relief under California state Unfair Competition Law (Cal. Business and Professions Code, § 17200.)).

98. *Id.*

99. Jenna Burrell, *How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms*, 3.1 BIG DATA & SOC’Y 1, 3–4 (2016) (categorizing such mechanisms of opacity into various categories and suggesting that some opacity stems from “technical illiteracy,” due to the specialized technical skill set needed to evaluate algorithms).

100. See Eslami et al., *User Attitudes*, *supra* note 74, at 12–13 (proposing user-auditing enhancing tools and crowd-sourcing algorithmic auditing in the context of potentially abusive machine-learning processes).

101. See Kenneth A. Bamberger & Deirdre K. Mulligan, *Privacy on the Books and on the Ground*, 63 STAN. L. REV. 247, 263 (2011) (discussing the rise of online privacy seal programs).

consumers of the individualized information-gathering burden and facilitating market enforcement.

Used in this fashion, such tools would further expose abusive terms and behaviors to the attention of consumer advocacy groups—truly increasing their salience.<sup>102</sup> One can even envision how such tools can be used to train machine-learning algorithms to highlight and spot behaviors in the wild, and flag them for review by consumers, regulators, and lawyers—using code to spot abusive code.<sup>103</sup>

With growing attention to privacy concerns, moreover, regulators, developers, and platform providers, such as the Google Play Store, need better tools to monitor app behavior and hold app developers accountable. Dynamic analysis tools and privacy-enhancing technologies and innovations geared to support greater transparency into data collection practices are a key component to the privacy landscape. Similarly, the tools described in this Article could benefit regulators in investigating the market for noncompliance by making it easier for them to detect violations and bring enforcement actions. If these enforcement actions are brought publicly, it may motivate other app developers to pay more attention to the privacy behaviors of their apps.

## V. CONCLUSION

Our findings about consumer expectations and app privacy behavior strengthen the case for combining laws grounding consumer protection in behavioral realities with privacy-enhancing technologies that increase accountability. Privacy has seen its share of democratic degradation, where decades-long research has demonstrated the inability of consumers to comprehend lengthy privacy policies or notices and the ways that this failure

---

102. Cf. MARGARET JANE RADIN, *BOILERPLATE: THE FINE PRINT, VANISHING RIGHTS, AND THE RULE OF LAW* 16, 243 (2012) (“NGO can organize publicity campaigns to make known to the public what some of the onerous terms in the fine print actually mean. The can take the lead in organizing a rating site that will advise consumers which firms are using reasonable terms and which are not.”); see also *Ranking Digital Rights*, NGO <https://rankingdigitalrights.org/> [<https://perma.cc/LSW2-5LBV>] (last visited Jan. 7, 2020) (rating leading internet companies human rights accountability posture (on a variety of topics from free expression to privacy) based on their Terms of Service and Privacy Policies, *inter alia*).

103. See Irwin Reyes et al., “*Won’t Somebody Think of the Children?*” *Examining COPPA Compliance at Scale*, PROC. ON PRIVACY ENHANCING TECH. 63 (2018) (explaining how the AppCensus tool is allowing users to search a name of a mobile app and learn about its actual information collection practice thereby empowering users).



largely obviates market competition over the quality of privacy-related contractual clauses.<sup>104</sup>

The wide divergence between expectations and reality counsels the expansion of regulatory approaches that on one hand reflect the ways that inaccurate expectations might lead to user deception, and on the other recognize that consumer confusion can vitiate notions of consumer choice by eliminating privacy as a salient factor in consumer decision making altogether. Equipping users, third parties, and regulators with analytic tools once reserved only for experts and academic researchers, moreover, would go far to address this phenomenon. Dynamic analysis tools offer the means to audit and contest explanations and notices provided by private parties and uncover actual behavior in an accessible manner—crucial to efforts to introduce more transparency and explainability in the context of machine-learning processing and information collection, and bringing information to bear to shape market practices.

---

104. RADIN, *BOILERPLATE*, *supra* note 102, at 213 (explaining how boilerplate are acts of “democratic degradation”; they employ mass systems of contracts to restructure and supersede the rights given by legislators, taking away rights granted by the democratic process). In the context of privacy, see Alessandro Acquisti & Jens Grossklags, *Privacy and Rationality in Individual Decision Making*, 3.1 IEEE SECURITY & PRIVACY 26 (2005) (providing survey evidence as to how the bounded rationality of users affects their privacy decision-making processes and attitudes).

