UNC SCHOOL OF LAW

NORTH CAROLINA JOURNAL OF LAW & TECHNOLOGY

Volume 16 | Issue 3 Article 6

3-1-2015

How the Rise of Big Data and Predictive Analytics Are Changing the Attorney's Duty of Competence

Peter Segrist

NORTH CAROLINA JOURNAL OF LAW & TECHNOLOGY
VOLUME 16, ISSUE 3: MARCH 2015

HOW THE RISE OF BIG DATA AND PREDICTIVE ANALYTICS ARE CHANGING THE ATTORNEY'S DUTY OF COMPETENCE

Peter Segrist*

If the legal profession had been able to foresee in the late 1990s and early 2000s, prior to the meteoric rise and ensuing cultural ubiquity of social media, that every tagged spring break photo, 2:00 a.m. status update, and furious wall post would one day be vulnerable to potential exposure in the cold, unforgiving light of civil and criminal litigation, attorneys would have been well-advised to discuss the ramifications of such actions, statements, and disclosures with their clients. Today, a similar phenomenon is looming in the form of the collection, aggregation, analysis and sale of personal data, and it will be the prudent attorney who competently advises his clients to stay ahead of the curve.

From the standpoint of attorney competency, the emergence of the Internet has forced attorneys to confront unique and complex ethical problems in terms of advising clients as to which types of Internet activity may be off limits or ill-advised. The recent ethics opinions of several bar associations demonstrate how technological advances are effectively shaping the duty of competence, concluding that an attorney's duty of competence may include an obligation to advise clients regarding their posts to social media. Contemporaneously, the data broker industry has rapidly expanded into the digital sphere, daily collecting huge

* Associate, Sher, Garner, Cahill, Richter, Klein & Hilbert, L.L.C.; Law Clerk to the U.S. District Judge Patricia Minaldi, Western District of Louisiana, 2013–14; J.D., 2013, Loyola University New Orleans, College of Law; Editor-in-Chief, 2012–13, Loyola Law Review; B.S., 2006, Tulane University. The author would also like to express his sincere gratitude to the very talented editors at the University of North Carolina Journal of Law and Technology, and to Sarah Dawkins for her unrelenting encouragement, patience and support.

swaths of consumers' personal information. That information is now legally exchanged between entities for value, and sophisticated analytical tools have been developed that permit data holders to make meaningful, highly accurate and highly personal deductions and predictions from high volume, seemingly chaotic, datasets. This Article argues that the same rationale that supports the notion that attorneys should advise clients against irresponsible social media usage also supports the finding that, given the current lack of regulation on the collection, commoditization, aggregation and analysis of consumer data, there is an emerging ethical obligation to advise clients regarding the responsible, and, ideally, anonymous, use of the Internet.

TABLE OF CONTENTS

I. INTRODUCTION

> [T]he intensity and complexity of life, attendant upon advancing civilization, have rendered necessary some retreat from the world, and man, under the refining influence of culture, has become more sensitive to publicity, so that solitude and privacy have become more essential to the individual; but modern enterprise and invention have, through invasions upon his privacy, subjected him to mental pain and distress, far greater than could be inflicted by mere bodily injury.[1]

The creation of the Internet has forced attorneys to confront complex ethical problems when advising clients on appropriate Internet activity. Recent ethics opinions from several bar associations have concluded that an attorney's duty of competency extends to advising clients regarding their posts to social media, subject to substantive rules regarding spoliation, due to the huge potential impact that such postings can have on a client's position in both potential and ongoing litigation.[2] Activities on social media, general Internet browsing, and myriad other everyday activities now generate tremendous amounts of seemingly innocuous personal data.[3] Contemporaneously, the data broker industry, which has essentially commoditized information associated with the

---

[1] Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 196 (1890).

[2] *See, e.g.*, N.Y. State Bar Ass'n, Social Media Ethics Guidelines (2014), https://www.nysba.org/Sections/Commercial_Federal_Litigation/Com_Fed_PDFs/Social_Media_Ethics_Guidelines.html [hereinafter NYSBA Opinion]; Phila. Bar Ass'n, Prof'l Guidance Comm., Op. 2014-5, at 4–5 (2014), *available at* http://www.philadelphiabar.org/page/Opinions2010Present?appNum=4 [hereinafter PBA Opinion] (discussing discovery concerns related to social media); N.Y. Cnty. Lawyers Ass'n, Op. 745 (2013), *available at* http://www.nycla.org/siteFiles/Publications/Publications1630_0.pdf [hereinafter NYCLA Opinion] (discussing advising a client regarding posts on social media sites); Pa. Bar Ass'n, Formal Opinion 2014-300, *available at* https://www.pabar.org/members/catalogs/Ethics%20Opinions/formal/F2014-300.pdf#search=%222014-300%22 [hereinafter Penn. Opinion] (discussing ethical obligations for attorney using social media); N.C. State Bar Ass'n, Formal Ethics Op. 5 (2014) [hereinafter NCB Opinion], http://www.ncbar.com/ethics/ethics.asp?page= 5&from=7/2014; ABA MODEL RULES OF PROF'L CONDUCT R. 1.1, cmt. 8 (placing an affirmative duty upon attorneys to stay abreast of technological developments relevant to the practice of law as part of their competency obligation).

[3] *See generally infra* Part II (discussing private sector data collection practices).

individual, has rapidly expanded into the digital sphere, collecting huge swaths of consumers' personal information daily. This information is now legally exchanged between entities for value, and sophisticated analytical tools have been developed that permit data holders to make meaningful, highly accurate, and highly personal deductions and predictions from high volume, seemingly chaotic, datasets. Given the current lack of regulation on the collection, commoditization, aggregation, and analysis of consumer data, this Article argues that the same rationale that supports the notion that attorneys should advise clients against irresponsible social media usage also supports the finding that there is an emerging ethical obligation to advise clients regarding the responsible and, ideally, anonymous Internet use.

Since the first censuses were conducted and crop yields recorded in the ancient world, data collection and analysis have been crucial components of a wide array of societal and technological improvements.[4] Today, data storage and processing costs are plummeting, while data collection methods are increasing. Simultaneously, increases in the number and variety of data-producing devices—sensor technologies, GPS trackers, and the so-called "Internet of things,"[5]—as well as the number of individuals connected to the Internet, have given rise to a situation wherein the amount of data presently available to both governments and private industries to feed the machinery of information processing analysis, with regards to either an entire population, or a single individual, has become unimaginably huge.[6]

---

[4] EXECUTIVE OFFICE OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES 1 (2014), *available at* http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf [hereinafter WHITE HOUSE BIG DATA REPORT]; *see also* VIKTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, BIG DATA 21 (2014) [hereinafter CUKIER BIG DATA] (discussing early censuses and associated problems).

[5] See *infra* notes 61–69, and accompanying text, describing the Internet of things as the sum of all devices connected to the Internet, such as thermostats, heart monitors, car insurance company driving monitors, and the like.

[6] *See* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 1 ("The collection, storage, and analysis of data is on an upward and seemingly unbounded trajectory, fueled by increases in processing power, the cratering costs of

In the wake of these advancements, a relatively new industry has developed. In its recent report on Big Data, the Federal Trade Commission ("FTC") defined data brokers as "companies that collect consumers' personal information and resell or share that information with others."[7] Personal data has been commoditized, and is now sold and exchanged like any other good by these data brokers. As the practice of wide scale data collection grows, both private industry and governmental bodies and agencies are rapidly discovering that algorithmic[8] analyses of vast troves of information empower the data holders to make staggering deductions about anything from where the next flu epidemic is likely to strike in the United States, to whether a woman is pregnant and the date of conception, to a person's likelihood of committing a criminal act, and much more.[9] This is the essence of the emerging field of predictive analytics wherein technological advances have created

---

computation and storage, and the growing number of sensor technologies embedded in devices of all kinds."); Yafit Lev-Aretz, *Copyright Lawmaking and Public Choice: From Legislative Battles to Private Ordering*, 27 HARV. J.L. & TECH. 203, 253 (2013) (describing user activity in the billions at sites such as Google and YouTube); Amit Chowdry, *Samsung, Intel and Dell Launch "Internet of Things" Consortium*, FORBES (July 9, 2014), http://www.forbes.com/sites/amitchowdhry/2014/07/09/samsung-intel-and-dell-launch-internet-of-things-consortium/; CUKIER BIG DATA, *supra* note 4, at 9 ("The amount of stored information grows four times faster than the world economy, while the processing power of computers grows nine times faster.").

[7] FEDERAL TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY at i (2014), *available at* http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf [hereinafter FTC REPORT].

[8] *See Algorithm,* BLACK'S LAW DICTIONARY (9th ed. 2009) (defining "algorithm" as "[a] mathematical or logical process consisting of a series of steps, designed to solve a specific type of problem").

[9] *See generally infra* Part III (discussing big data and predictive analytics); *see also* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at pmbl. ("A significant finding of this report is that big data analytics have the potential to eclipse longstanding civil rights protections in how personal information is used in housing, credit, employment, health, education, and the marketplace.").

the ability to make detailed predictions and deductions from gigantic, and seemingly chaotic, datasets.[10]

No single technological advancement since the invention of the printing press has so dramatically affected nearly every aspect of human society as has the Internet. Its birth and development have resulted in dramatic and global shifts in countless professions, and the legal profession has not been immune to this transformation.[11] For instance, attorneys arguably have an ethical obligation to be able to competently navigate the Internet as part of their basic researching skills.[12] The American Bar Association ("ABA") now

---

[10] *See generally infra* Part III.C (discussing predictions and deductions that are capable of being made from large datasets).

[11] David Hricik, *Lawyers Worry Too Much about Transmitting Client Confidences by Internet E-mail*, 11 GEO. J. LEGAL ETHICS 459, 459 (1998) ("The Internet has changed American business, including the legal profession."); James Podgers, *Lawyers Struggle to Reconcile New Technology with Traditional Ethics Rules*, ABA J. (Nov. 1, 2014), http://www.abajournal.com/magazine/article/the_fundamentals_lawyers_struggle_to_reconcile_new_technology_with_traditio/; ABA Comm. on Prof'l Ethics 20/20, Introduction and Overview (2012), *available at* http://www.americanbar.org/content/ dam/aba/administrative/ethics_2020/20120508_ethics_20_20_final_hod_introdution_and_overview_report.authchec kdam.pdf ("[T]echnology has irrevocably changed and continues to alter the practice of law in fundamental ways."); Andrew Perlman, *The Twenty-First Century Lawyer's Evolving Ethical Duty of Competence*, 22 NO. 4 THE PROF. LAW. 1 (2014), *available at* http://www.americanbar.org/publications/professional_lawyer/2014/volume-22-number-4/the_twentyfirst_century_lawyers_evolving_ethical_duty_competence.html (identifying electronic discovery and cloud-based services as some of evolving facets of a lawyer's duty of competence).

[12] *See generally* Lawrence D. MacLachlan, *Gandy Dancers on the Web: How the Internet Has Raised the Bar on Lawyers' Professional Responsibility to Research and Know the Law*, 13 GEO. J. LEGAL ETHICS 607 (2000); *see also* Perlman, *supra* note 11, at 4 (discussing *Iowa Supreme Court Att'y Disciplinary Bd. v. Wright*, 840 N.W.2d 295, 301–04 (Iowa 2013), where a lawyer was disciplined for permitting his clients to fall for a well-known internet scam involving an inheritance from a distant Nigerian relative, because the lawyer failed to conduct a "'cursory internet search' that would have uncovered the truth," and *Johnson v. McCullough*, 306 S.W.3d 551, 558–59 (Mo. 2010), wherein the "Missouri Supreme Court recently held that lawyers should use 'reasonable efforts,' including Internet-based tools, to uncover the litigation

routinely releases ethics opinions addressing issues such as a lawyer's ability to review and research a juror's Internet presence[13] or what is required to protect confidentiality when sending an unencrypted email to a client.[14]

Squarely addressing such developments, the ABA recently added a comment to Rule 1.1 of the Model Rules of Professional Conduct modifying an attorney's competency obligations to include an affirmative duty to educate him or herself as to *technologies* relevant to the practice of law.[15] Addressing clients' social media postings, several bar associations' recent ethics opinions have noted that an attorney's competency obligations under Rule 1.1 could "give rise to an obligation to advise clients, within legal and ethical requirements, concerning what steps to take to mitigate any adverse effects on the clients' position emanating from the clients' use of social media."[16] This stems from the simple and obvious proposition that an individual should be careful in allowing personal information about one's self to be disseminated into the world, as that information may ultimately be used against its owner.

Applying the foregoing concept to big data collection and an attorney's duty to advise clients, consider the following: big data analytics are now capable of making increasingly personal deductions about individuals from large, seemingly random datasets and a largely unregulated for-profit industry has emerged for the purpose of personal data collection, commodification,

---

history of jurors prior to trial in order to preserve possible objections to the empanelment of those jurors").

[13] ABA Comm. on Ethics & Prof'l Responsibility, Formal Op. 466 (2014) (discussing an attorney's ability to review a juror's Internet "presence").

[14] ABA Comm. on Ethics & Prof'l Responsibility, Formal Op. 413 (1999); *see generally* Peter Geraghty, *Cybersecurity and the Use of Emerging Technologies, Part 1*, YOUR ABA (Dec. 2014), *available at* http://www.americanbar.org/publications/youraba/2014/december-2014/cybersecurity-and-the-use-of-emerging-technologies--part-1.html.

[15] MODEL RULE OF PROF'L CONDUCT R. 1.1, cmt. 8.

[16] NYCLA Opinion, *supra* note 2; *see also* NYSBA Opinion, *supra* note 2; PBA Opinion, *supra* note 2; Penn. Opinion, *supra* note 2; NCB Opinion, *supra* note 2.

aggregation, analysis and sale.[17] Moreover, anonymization of this data is not feasible as a long-term solution to privacy concerns.[18] Further, the technology exists today that would allow a data broker to use already legally-collected information pertaining to a given individual to make highly accurate deductions and predictions about that individual. Consider also the steadily increasing utility and ubiquity of informal discovery and online research in litigation—including the employment of data brokers by attorneys.[19]

The practical applications and advantages of having access to such information on one's adversaries for parties engaged in litigation, or potential litigation, are readily apparent, both in terms of leading to discoverable evidence to be used at trial as well as in discovering inadmissible information that is nevertheless advantageous to creative litigants engaged in contentious civil or criminal litigation. For instance, imagine the value to a corporate defendant engaged in settlement negotiations with an injured plaintiff-employee to know that that injured worker is facing severe financial constraints. Consider the degree to which that same defendant's negotiating position would be strengthened if it possessed information that the same plaintiff had a child at home with an expensive-to-treat chronic illness—how much more financially desperate and, thus, eager to settle, would such a plaintiff be? Consider further how advantageous it would be for a family law attorney to be able to employ data experts to determine the likelihood of a party's infidelity or the likelihood of a spouse's continued substance abuse in child custody proceedings. How might such information inform settlement negotiations, investigatory tactics, or trial strategy?

---

[17] *See, e.g.*, CUKIER BIG DATA, *supra* note 4, at 156 ("[F]irms of all stripes amass mountains of personal information concerning all aspects of our lives, share it with others without our knowledge, and use it in ways we could hardly imagine.").

[18] *See infra* Part III.D (discussing the shortcomings of anonymization strategies).

[19] *See, e.g.*, *infra* notes 111–13 and accompanying text (discussing attorneys' use of data brokers).

The field of predictive analytics,[20] described below, is becoming more sophisticated, and is capable of making increasingly accurate deductions from the personal data already available to data brokers. Data brokers already create profiles and dossiers of individuals for sale.[21] The primary constraints on the industry are self-regulatory in nature, enforced only by the public's capacity for outrage at private sector data collection and privacy intrusions.[22] Had legal practitioners been able to foresee twenty years ago the impact that public online postings and social media would have on future litigation, attorneys may have prevented their clients from voluntarily distributing highly personal, potentially damaging information out into the world. Since personal information, once gleaned—from browsing habits, online purchases, customer loyalty and reward cards, the use of the various devices that comprise the Internet of things, online surveys, GPS tracking

[20] *See infra* Part III.A (describing the field of predictive analytics).

[21] *See infra* notes 107–10 and accompanying text (describing "people search" and related products).

[22] *See* PEW RES. CTR., PUBLIC PERCEPTIONS OF PRIVACY AND SECURITY IN THE POST-SNOWDEN ERA 3 (Nov. 12, 2014), *available at* http://www.pewinternet.org/files/2014/11/PI_PublicPerceptionsofPrivacy_111214.pdf (noting that, even though a majority of Americans would like to see the government do more to regulate private sector data collection, fifty-five percent of respondents "'agree' or 'strongly agree' with the statement: 'I am willing to share some information about myself with companies in order to use online services for free.'"); *see also infra* notes 270–71 and accompanying text; *60 Minutes: Data Brokers* (CBS television broadcast Aug. 24, 2014); Shannon Pettypiece & Jordan Roberston, *Did You Know You Know Had Diabetes? It's All Over the Internet*, BLOOMBERG.COM (Sept. 11, 2014), http://www.businessweek.com/news/2014-09-11/how-big-data-peers-inside-your-medicine-chest. *But see* PEW RES. CTR., PUBLIC PERCEPTIONS OF PRIVACY AND SECURITY IN THE POST-SNOWDEN ERA 28 (Nov. 12, 2014), *available at* http://www.pewinternet.org/files/2014/11/PI_PublicPerceptionsofPrivacy_111214.pdf ("[P]ublic concern over the amount of personal information businesses are collecting has been growing."). *See generally infra* Part II.C. Given current research and trends, it would not be unreasonable for one to conclude that, although public concern over private sector data collection may be growing as Americans become more educated on the subject, few will demand comprehensive change simply due to the convenience offered by free online services. *See infra* note 270 and accompanying text (discussing "privacy fatigue").

information reported by mobile devices, and on and on—is often never truly deleted, this Article attempts to make the case that a new ethical obligation is emerging which makes it incumbent upon attorneys to advise their clients of what is necessary to prevent their personal data from leaking out into the digital world *today*, so that it might not be used against them tomorrow.

Part II describes some of the more common private sector data collection practices. Part III introduces big data and predictive analytics, and attempts to demonstrate how conclusions can be drawn from vast amounts of seemingly innocuous data. Part IV describes the current legal climate in which the industry operates. Part V comments on the impact these developments are having on the attorney's duty of competence. Part VI then sets forth some of the more basic suggested techniques of stopping or, at least, slowing and diluting the flow of one's personal data out into the digital realm.

II. THE HARVEST—HOW INFORMATION IS COLLECTED, AGGREGATED, AND PROCESSED

"*Our digital reach will soon approach nearly every Internet user in the U.S.*"[23]

Begin with this basic premise: deductions can be made from information. The more information one has, the more one can deduce. Additionally, when one has historical data for comparison, one can more accurately make predictions. The following section discusses some of the more common ways that private companies

[23] Judith Aquino, *Acxiom Prepares New 'Audience Operating System' Amid Wobbly Earnings*, AD EXCHANGER (Aug. 1, 2013 2:48 PM), http://www.adexchanger.com/analytics/acxiom-prepares-new-audience-operating-system-amid-wobbly-earnings/ (quoting Scott Howe, CEO of Acxiom); *see also* Richard Behar, *Never Heard of Acxiom? Chances Are It's Heard of You*, FORTUNE (Feb. 23, 2004), http://archive.fortune.com/magazines/fortune/fortune_archive/2004/02/23/362182/index.htm; ACXIOM CORP., ANNUAL REPORT (2013), *available at* d3u9yejw7h244g.cloudfront.net/wp-content/uploads/2013/09/2013-Annual-Report.pdf ("[O]ur capabilities include . . . multi-sourced insight into approximately 700 million consumers worldwide [with] [o]ver 3,000 propensities for nearly every U.S. consumer.").

are amassing information about virtually every person on earth—one click, purchase, or digital transaction at a time—in order to make those deductions and sell the results.

A. *Tracking Methodology*

The diversity of methods by which information is collected in today's increasingly digitized environment is difficult to overstate. Take Internet browsing: the typical person's daily Internet activity is collected, traced, logged, and analyzed by a dizzying number of entities in a variety of ways. The data-collecting entities themselves are often divided into two groups: so-called "first parties"—social media, news websites, online retailers, and other consumer websites—who collect information directly from users, often unbeknownst to the users themselves, and "third parties"—those to whom information is either passed by first parties or who conduct their own monitoring and tracking of one's browsing habits surreptitiously.[24]

Every individual computer, smartphone, and tablet currently connected to the Internet has a unique Internet Protocol ("IP") address, like a digital fingerprint.[25] While websites need a user's IP

---

[24] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 41; *see also* Daniel J. Solove, *Privacy & Power: Computer Databases and Metaphors for Information Privacy*, 53 STAN. L. REV. 1393, 1411 (2001) ("Currently, there are two basic ways personal information is collected in cyberspace: (1) by directly collecting information from users (registration and transactional data); and (2) by surreptitiously tracking the way people navigate through the Internet (clickstream data)."); *see also* Chris Jay Hoofnagle, et al., *Symposium: Privacy and Accountability in the 21st Century: Behavioral Advertising: The Offer You Cannot Refuse*, 6 HARV. L. & POL'Y REV. 273, 276 (2012).

[25] *See* Anne Klinefelter, *When to Research is to Reveal: The Growing Threat to Attorney and Client Confidentiality from Online Tracking*, 16 VA. J.L. & TECH. 1, 6 (2011); Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814, 1837 (2011) (describing the IP address as "a unique identifier that is assigned to every computer connected to the internet"); *see also* Joshua J. McIntyre, *Symposium: Trial 2010: A Look Inside Our Nation's Courtrooms: Twentieth Annual DePaul Law Review Symposium: Comment: Balancing Expectations of Online Privacy: Why Internet Protocol (IP) Addresses Should*

address to deliver content, some websites also use these unique addresses to track Internet users for purposes such as billing, customer service, tracking user preferences, and targeted marketing, among others.[26] It is also standard practice for websites to collect the web address or universal resource locator ("URL") of the page that linked to them (i.e. the referring URL) which itself can reveal the user's immediately prior-used search terms and websites visited, while also collecting the date and time that someone from a particular IP address visited their website.[27]

While IP addresses, being generally fixed, are typically associated with either a unique home or office Internet connection, and thus easily associated with a specific individual through their Internet Service Provider ("ISP"), they can also easily become associated with a specific individual through a rather simple analysis of a user's web traffic.[28] First-party websites also acquire personal information whenever a user willingly volunteers it, for instance, by opening an account and giving a full name, home address, email address, or taking an online quiz entitled "Are You Good in Bed?".[29] A recent article in the Washington Post reported that the analytics code used by BuzzFeed, a self-described "social

---

Be Protected as Personally Identifiable Information, 60 DEPAUL L. REV. 895, 895–96 (2011).

[26] Klinefelter, *supra* note 25, at 6; *see also* Eloise Gratton, *If Personal Information Is Privacy's Gatekeeper, then Risk of Harm is the Key: A Proposed Method for Determining What Counts as Personal Information*, 24 ALB. L.J. SCI. & TECH. 105, 122 (2014) (citing Lisa J. Sotto & Melinda L. McLellan, *Online Behavioral Advertising: A User's Guide*, IP LITIG., Nov.–Dec. 2012, at 1–2) ("[B]ehavioral advertising may often involve the collection of IP addresses and the processing of unique identifiers (through the use of cookies).").

[27] Klinefelter, *supra* note 25, at 8.

[28] *See* Schwartz & Solove, *supra* note 25, at 1837–40 (discussing various techniques that enable the linking of IP addresses to specific individuals and providing several examples). *See*, *e.g.*, *infra* notes 170–77. *See also* Elbert Lin, *Prioritizing Privacy: A Constitutional Response to the Internet*, 17 BERKELEY TECH. L.J. 1085, 1104 n.101 (2002) (discussing static versus dynamic IP addresses).

[29] *See* Steven C. Bennett, *Regulating Online Behavioral Advertising*, 44 J. MARSHALL L. REV. 899, 901 (2011); Alyssa Bailey, *Quiz: Are You Good in Bed?*, COSMOPOLITAN (July 2, 2014), http://www.cosmopolitan.com/sex-love/a26964/cosmo-quiz-are-you-good-in-bed/.

news and entertainment company[,]" indicated that "the site has tools in place to build individualized data profiles based on users' quiz responses—which sometimes include deeply personal information, like whether you[have] had an eating disorder or taken meds for a mental illness."[30]

Independent of IP address tracking, one of the most common ways that both first- and, particularly, third-party tracking occurs is with "cookies."[31] Cookies are small bits of text that are downloaded automatically from websites by a user's browser as one navigates the Internet.[32] They essentially identify the computer on which they are stored, carrying information about what a user does online back to the website that attached the cookie to the user's computer in the first place.[33] Websites use cookies for a variety of purposes, such as remembering a user's preferences on that site or understanding how users are actually using a site in order to improve site performance and security.[34] Cookies also track browsing activity and collect information for advertisers and data brokers.[35] Further, while cookies do not always contain

[30] BUZZFEED, *About*, http://www.buzzfeed.com/about (last visited Dec. 28, 2014); Caitlin Dewey, *The Scary, Eye-Opening Truth of Internet Tracking – on Buzzfeed Quizzes, and Everywhere Else*, WASH. POST (Jun. 26, 2014), http://www.washingtonpost.com/news/the-intersect/wp/2014/06/26/the-scary-eye-opening-truth-of-internet-tracking-on-buzzfeed-quizzes-and-everywhere-else/.

[31] *See* Joanna Geary, *Tracking the Trackers: What Are Cookies? An Introduction to Web Tracking*, THE GUARDIAN (Apr. 23, 2012, 12:08 PM), http://www.theguardian.com/technology/2012/apr/23/cookies-and-web-tracking-intro. Jay P. Kesan et al., *Information Privacy and Data Control in Cloud Computing: Consumers, Privacy Preferences, and Market Efficiency*, 70 WASH. & LEE L. REV. 341, 437 (2013) (describing cookies).

[32] *See* Kesan et al., *supra* note 31, at 437 (describing the distinction between text cookies and flash cookies).

[33] Geary, *supra* note 31; Kesan et al., *supra* note 31, at 437.

[34] Geary, *supra* note 31; *see also* Chris Jay Hoofnagle et al., *Behavioral Advertising: The Offer You Cannot Refuse*, 6 HARV. L. & POL'Y REV. 273, 276 (2012) (defining cookies as "small text files that typically contain a string of numbers that can be used to identify a computer").

[35] *See* Adam Tanner, *The Web Cookie Is Dying. Here's the Creepier Technology That Comes Next*, FORBES (June 17, 2013, 12:29 PM), http://www.forbes.com/sites/adamtanner/2013/06/17/the-web-cookie-is-dying-heres-the-creepier-technology-that-comes-next/; *see also* FTC REPORT, *supra* note 7, at v ("Data brokers rely

personally identifiable information, some companies specifically provide the service of linking cookies to users' personal information.[36]

Many websites, however, are now employing alternatives to cookies in order to track online behavior. This course change is in part a response to the so-called "Do Not Track" ("DNT") initiative.[37] In the Internet community's attempt at a Do Not Call list,[38] the DNT concept was originally envisioned as a simple way for consumers to control and limit the extent to which their online activity is tracked.[39] DNT features are often available as a setting on an individual's web browser that tells the browser to communicate to websites that the user wishes not to have his or her online activity tracked.[40] Some browsers provide a similar feature

---

on websites with registration features and cookies to find consumers online and target Internet advertisements to them based on their offline activities.").

[36] Kesan et al., *supra* note 31, at 437 (*citing* Daniel J. Solove, THE DIGITAL PERSON: TECHNOLOGY AND PRIVACY IN THE INFORMATION AGE 24–25 (2006)) ("[A] company called DoubleClick provides a service to websites, connecting cookies to personal information to enable more targeted advertising."); *see also* Dillon Reisman et al., *Cookies That Give You Away: Evaluating the Surveillance Implications of Web Tracking* 1–2 (Working Draft, Apr. 2, 2014), *available at* http://randomwalker.info/publications/cookie-surveillance.pdf (describing technical methods for identifying individuals through cookie tracking even in the absence of knowing a target's IP address); *see also* Adi Kamdar et al., *NSA Turns Cookies (And More) Into Surveillance Beacons*, ELEC. FRONTIER FOUND. (Dec. 11, 2013), https://www.eff.org/deeplinks/2013/12/nsa-turns-cookies-and-more-surveillance-beacons (describing the use of PREF cookies to uniquely identify individuals).

[37] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 42–43.

[38] Zach Miners, *How Bickering and Greed Neutered the "Do Not Track" Privacy Initiative*, PC WORLD (May 22, 2014, 6:22 AM), http://www.pcworld.com/article/2158220/do-not-track-oh-what-the-heck-go-ahead.html.

[39] WHITE HOUSE BIG DATA REPORT, *supra* note 4 at 42–43. FEDERAL TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: A PROPOSED FRAMEWORK FOR BUSINESSES AND POLICYMAKERS at vi–vii (2010), *available at* http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-bureau-consumer-protection-preliminary-ftc-staff-report-protecting-consumer/101201privacyreport.pdf [hereinafter 2010 FTC Report].

[40] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 42–43; *see also Do Not Track Test Page*, MICROSOFT, ie.microsoft.com/TEStdrive/Browser/DoNotTrack/

by offering privacy settings that allow the wholesale blocking of third-party cookies.[41] However, currently, many sites do not honor DNT requests. Even sites that do respond to DNT requests often interpret such requests in different ways due to a lack of consensus among Internet companies and service providers. Other sites simply show unwillingness on their part to acquiesce to the intent of DNT.[42] As a result, while utilizing a DNT feature on one's browser will indeed send a signal to a host website that the user wishes not to be tracked, that request will likely be ignored.[43] In what is generally a "self-regulating" industry, this phenomenon is not encouraging.[44]

This has given rise to so-called "web beacons," also known as "pixel" tracking.[45] These small bits of code embedded into a web page that are invisible to the user and track that user's activity wherever they go online, sending signals regarding his or her activity back to the beacons' hosts.[46] This is particularly relevant in

---

Default.html (last visited Aug. 22, 2014) ("This page detects whether or not your browser has a Do Not Track preference set.").

[41] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 42–43.

[42] *See* Elizabeth Dwoskin, *Yahoo Won't Honor 'Do Not Track' Requests from Users*, WALL ST. J. (May 2, 2014, 8:22 PM), blogs.wsj.com/digits/2014/05/02/yahoo-wont-honor-do-not-track-requests-from-users.

[43] Miners, *supra* note 38; *see also* Dwoskin, *supra* note 42; Fred B. Campbell, Jr., *The Slow Death of "Do Not Track,"* N.Y. TIMES, (Dec. 26, 2014), http://www.nytimes.com/2014/12/27/opinion/the-slow-death-of-do-not-track.html?_r=0.

[44] *See infra* notes 268–69 and accompanying text.

[45] Geary, *supra* note 31; Violet Blue, *Facebook Turns User Tracking "Bug" Into Data Mining "Feature" for Advertisers*, ZD NET (June 17, 2014, 12:01 PM), http://www.zdnet.com/facebook-turns-user-tracking-bug-into-data-mining-feature-for-advertisers-7000030603/; *see also* Matthew Sundquist, *Online Privacy Protection: Protecting Privacy, the Social Contract, and the Rule of Law in the Virtual World*, 25 REGENT U.L. REV. 153, 161–62 (2013) ("Lotame Solutions uses web beacons that record what a person types on a website in order to create a user profile, while Apple, Verizon, Target, and others compile information from customers' interactions with their products.").

[46] Geary, *supra* note 31; Blue, *supra* note 45; *see also Pixel Tracking in Third-Party and Custom Creatives*, GOOGLE, https://support.google.com/dfp_premium/answer/1347585?hl=en (last visited Dec. 28, 2014) ("A tracking pixel is simply code inserted into a custom or third-party creative that makes a server call and returns a transparent 1x1 image (normally a GIF file)."); *see also*

light of the changes to Facebook's privacy policy that took place in the summer of 2014. The new policy stated that Facebook would begin collecting information about users from sites they visit, apps they use, and their browsing histories, even when a user is not logged into Facebook.[47] The Facebook blog stated that, "[i]n short, your browsing habits on any site or mobile app with a Facebook like button (who doesn't have that nowadays) can also be viewed by Facebook and thus used for advertising data."[48]

A relatively new technology known as "canvas fingerprinting" is also growing in popularity as the tracking method *du jour*.[49]

Andrea M. Matwyshyn, *Of Nodes and Power Laws: A Network Theory Approach to Internet Jurisdiction Through Data Privacy*, 98 NW. U.L. REV. 493, 496 n.17 (2004) ("A web bug, also known as a 'pixel tag.' 'web beacon,' or 'clear GIF,' is a lx1 pixel image embedded in html code. If a web bug is present, instead of simply fetching an image, a server invokes a CGI program that logs information about the user's actions.").

[47] *See, e.g.*, Camila Domonoske, *Facebook Ad Targeting Will Use Even More of Your Data*, NPR (June 12, 2014. 1:14 PM), http://www.npr.org/blogs/alltechconsidered/2014/06/12/321325434/facebook-ad-targeting-will-use-even-more-of-your-data. *See also Data Use Policy*, FACEBOOK, http://www.facebook.com/about/privacy/your-info (last visited Dec. 13, 2014) (discussing "other information we receive about you"). It should be noted that the Facebook privacy terms have since been further amended; however, as one technology writer reported, quoting the Washington Post's Switch blog, "Facebook rewrites its privacy policy so that humans can understand it," and also quoting Fortune's Tech blog as stating, "Facebook's privacy policy is clearer, but no less complicated." Jennifer Abel, *Facebook Rewrites and Sort-of Updates Its Privacy Policies, Again*, CONSUMERAFFAIRS.COM (Nov. 24, 2014), http://www.consumeraffairs.com/news/facebook-rewrites-and-sort-of-updates-its-privacy-policies-again-112414.html.

[48] Shruti Dhapola, *How Facebook's New "Ad Preference" Policy is Threatening Your Privacy*, TECH2 (June 18, 2014, 9:52 AM), http://tech.firstpost.com/news-analysis/how-facebooks-new-ad-preference-policy-is-threatening-your-privacy-226028.html (quoting Facebook's blog).

[49] *See Sneaky New Tactics May Be Tracking You Online*, CBS NEWS (July 24, 2014, 5:00 AM), http://www.cbsnews.com/news/sneaky-new-tactics-may-be-tracking-you-online; *see also* Gunes Acar et al., The Web Never Forgets: Persistent Tracking Mechanisms in the Wild (Aug. 10, 2014) (unpublished manuscript), *available at* https://securehomes.esat.kuleuven.be/~gacar/persistent/the_web_never_forgets.pdf; Olga Kharif, *The Cookies You Can't Crumble*, BUSINESSWEEK (Aug.

Whereas traditional tracking methods involve the transfer of small segments of code or files, such as cookies, thus enabling them to be more easily identified by the user, canvas fingerprinting is virtually undetectable to the average Internet user.[50] This method is predicated on the idea that no two computer systems are exactly alike due to the extremely high number of possible combinations of a multitude of variables from one device to another—in terms of various operating systems, versions of those systems, browsers, graphics settings, font settings, and the like—à la "fingerprinting."[51] Utilizing this phenomenon, websites send a request to a user's browser to draw a small text image, to which the browser complies. Due to the differences in settings from one user's computer to another, this creates an identifiable "fingerprint," in that each computer and browser displays the text in a singular and uniquely identifiable manner.[52] Thus, when different websites use the same tracking methodology, they can track a single user from one site to another, generally unhindered by his or her use of anti-tracking tools or browser privacy settings.[53]

The use of these and similar technologies are commonly employed to monitor the ongoing web activity of virtually all Internet users on the planet. A 2010 Wall Street Journal investigation found that the "nation's 50 top websites on average installed 64 pieces of tracking technology onto the computers of

---

21, 2014), http://www.businessweek.com/articles/2014-08-21/facebook-google-go-beyond-cookies-to-reap-data-for-advertisers (discussing canvas fingerprinting).

[50] *See* Joseph Steinberg, *You Are Being Tracked Online By a Sneaky New Technology—Here's What You Need to Know*, FORBES (July 23, 2014, 8:30 AM), http://www.forbes.com/sites/josephsteinberg/2014/07/23/you-are-being-tracked-online-by-a-sneaky-new-technology-heres-what-you-need-to-know/.

[51] *Id.*

[52] *Id.*

[53] *Id.*; *see also* Chris Smith, *The Creepiest Internet Tracking Tool Yet Is "Virtually Impossible" to Block*, BGR (July 22, 2014, 1:00 PM), bgr.com/2014/07/22/canvas-fingerprinting-internet-tracking-tool/. *But see* Jeremy Kirk, *'Canvas Fingerprinting' Online Tracking Is Sneaky but Easy to Halt*, PC WORLD (July 25, 2014, 6:31 AM), www.pcworld.com/article/2458280/canvas-fingerprinting-tracking-is-sneaky-but-easy-to-halt.html.

visitors, usually with no warning."[54] Privacy issues quickly emerge when third parties routinely track user activity en masse across multiple websites, allowing trackers to "infer users' interests, perhaps sensitive ones, such as medical conditions, political opinions, or even sexual fetishes."[55] Furthermore, promising efforts are now underway to link consumers' mobile devices to their home computers—a difficulty that has stymied advertisers and data brokers for some time—providing yet another source of personal data to private sector trackers.[56]

---

[54] *See* Hoofnagle et al., *supra* note 34, at 275 (*quoting* Julia Angwin, *The Web's New Gold Mine: Your Secrets*, WALL ST. J. (July 30, 2010), *available at* http://www.wsj.com/articles/SB10001424052748703940904575395073512989404).

[55] Hoofnagle et al., *supra* note 34, at 276; *see also* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 45 ("[Such precise profiling] represents a powerful capacity on the part of the private sector to collect information and use that information to algorithmically profile an individual . . . . This application of big data technology, if used improperly, irresponsibly, or nefariously, could have significant ramifications for targeted individuals.").

[56] *See, e.g.*, DRAWBRIDGE, http://www.drawbrid.ge/technology (last visited Aug. 21, 2014) (describing Drawbridge's cross-linking of mobile devices to computers); *see also* SILVERPUSH, http://www.silvrpush.com/ (last visited Aug. 21, 2014); Anthony Ha, *SilverPush Says Its Using "Audio Beacons" for an Unusual Approach to Cross-Device Ad Targeting*, TECHCRUNCH (July 24, 2014), http://techcrunch.com/2014/07/24/silverpush-audio-beacons/ (describing "audio beacons"). Basically, the company says it uses "ultrasonic inaudible sounds." If you are browsing the web and encounter a SilverPush advertiser, then at the same time that they are dropping a cookie on your computer, they also play one of those sounds. You will not be able to hear it, but if you have installed any app that uses the SilverPush software development kit, it will actually be listening for that sound in the background, and when it detects an "audio beacon," it is able to identify that your desktop/laptop computer and your phone/tablet belong to the same person. *Id.*; *see also* Kurt Wagner, *Twitter to Start Tracking Which Apps Its Users Have Downloaded*, RE/CODE (Nov. 26, 2014, 9:45 AM), http://recode.net/2014/11/26/twitters-now-collecting-data-on-which-apps-you-download/ (describing how both Apple's iOS and Google's Android already allow third parties to "ping a user's device at any time and recall a list of apps that are currently running on their smartphone"); Robert McMillan, *Verizon & AT&T Are the Only Wireless Carriers Using 'Perma-Cookies'*, WIRED (Nov. 7, 2014, 6:30 AM), http://www.wired.com/2014/11/permacookie-free/ (describing Verizon's use of "perma-cookies"—small strings

To illustrate the practice of online tracking, several companies now offer products that allow Internet users to visualize online tracking as it occurs. For instance, in late 2013, Mozilla released an add-on[57] called "Lightbeam" that was designed to demonstrate the tracking phenomenon to a mainstream audience.[58] Funded by grants from the Ford Foundation and the Natural Sciences and Engineering Research Council, with assistance from students at the Emily Carr University of Art and Design, the Lightbeam add-on allows Firefox users to view a graphical representation of the tracking of their Internet browsing in real time. The tool represents the first-party sites that the user actually visits as circles and the third-party sites that monitor a user's activities as triangular icons that revolve around their first party site counterpart.[59] This allows users to watch as, on average, approximately ten to thirty third parties monitor their activities on the vast majority of first-party sites.[60]

The Internet of things has further enhanced private companies' ability to track consumers' activities, and, thus, increased the private sector's capacity to gather detailed and personal information on users. "Internet of things" is:

---

of data that come preinstalled on phones and are slipped into all users' web traffic for identification purposes).

[57] *See Add-on*, TECHTERMS, http://www.techterms.com/definition/addon (last visited Jul. 25, 2014) (defining "add-on" as "a software extension that adds extra features to a program. It may extend certain functions within the program, add new items to the program's interface, or give the program additional capabilities.").

[58] *See* Olivia Solon, *Mozilla Releases Add-on that Reveals Online Data Tracking*, WIRED (Oct. 25, 2013), http://www.wired.co.uk/news/archive/2013-10/24/lightbeam; *see also* Samuel Gibbs, *Mozilla's Lightbeam Firefox Tool Shows Who's Tracking Your Online Movements*, THE GUARDIAN (Oct. 28, 2013), http://www.theguardian.com/technology/2013/oct/28/mozilla-lightbeam-tracking-privacy-cookies.

[59] *See Add-ons for Firefox*, MOZILLA, https://addons.mozilla.org/en-US/firefox/addon/lightbeam/ (last visited July 10, 2014).

[60] These figures were generated from firsthand use of the Lightbeam add-on. Notably, the Lightbeam privacy policy states "[b]y default, data collected by Lightbeam remains in your browser and is not sent to us." *Add-ons for Firefox*, *supra* note 59.

> [A] term used to describe the ability of devices to communicate with each other using embedded sensors that are linked through wired and wireless networks. These devices could include your thermostat, your car, or a pill you swallow so the doctor can monitor the health of your digestive tract. These connected devices use the Internet to transmit, compile, and analyze data.[61]

As GPS chips ping our mobile devices and offer us alternative routes home to avoid traffic jams, and as we adjust our thermostats, home security cameras, and other household devices remotely, those devices produce data that is collected and available for analysis, creating a massive new set of tools for data production.[62] For instance, Nest, a recent $3.2 billion Google acquisition, is a company that creates systems that allow users to remotely control their thermostats.[63] While one's thermostat usage may not appear particularly compromising on its own, it does reveal highly particularized information about one's living patterns—when one rises in the morning, goes to work, comes home in the evening, and goes to bed—which, when combined

---

[61] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 2; *see also* Rolfe Winkler & Alistair Barr, *Nest to Share User Information with Google for the First Time*, WALL ST. J. (June 24, 2014, 12:16 AM), blogs.wsj.com/digits/2014/06/24/nest-to-share-user-information-with-google-for-first-time/.

[62] *See* Chowdry, *supra* note 6; *see also* Bill Wasik, *In the Programmable World, All Our Objects Will Act as One*, WIRED (May 14, 2013, 6:30 AM), http://www.wired.com/2013/05/internet-of-things-2/ ("A decade after Wi-Fi put all our computers on a wireless network—and half a decade after the smartphone revolution put a series of pocket-size devices on that network—we are seeing the dawn of an era when the most mundane items in our lives can talk wirelessly among themselves, performing tasks on command, giving us data we've never had before."); Howard Baldwin, *A Match Made Somewhere: Big Data and the Internet of Things*, FORBES (Nov. 24, 2014, 11:06 AM), http://www.forbes.com/sites/howardbaldwin/2014/11/24/a-match-made-somewhere-big-data-and-the-internet-of-things/ ("[O]nce the Internet of Things gets rolling, stand back. We're going to have data spewing at us from all directions—from appliances, from machinery, from train tracks, from shipping containers, from power stations.").

[63] Kashmir Hill, *Nest Hackers Will Offer Tool to Keep the Google-Owned Company from Getting Users' Data*, FORBES (July 16, 2014, 9:25 AM), http://www.forbes.com/sites/kashmirhill/2014/07/16/nest-hack-privacy-tool/. Press Release, Google, Google to Acquire Nest (Jan. 13, 2014), *available at* https://investor.google.com/releases/2014/0113.html.

with additional data, helps to fill out the picture of an individual.[64] Google Glass—a wearable technology with Internet connectivity and video and picture-taking capability—is another example.[65] Before Google recently discontinued the Google Glass Explorer program,[66] Google Glass was combined with facial recognition programs and apps such as NameTag.[67] NameTag allows strangers to immediately access a person's name, photos, and dating website profiles simply by looking at them.[68] These capabilities of Google Glass in conjunction with several other data-producing and privacy-reducing features have led to some degree of public backlash, as evidenced by the rise of the term "glasshole" in the social vernacular, if nothing else.[69] Furthermore, the development

---

[64] As an additional, noteworthy security risk, researchers at the University of Central Florida recently discovered that it was possible to take control of Nest and secretly siphon off data from the Nest system. Hill, *supra* note 63; s*ee also* David Perera, *Smart Grid Powers Up Privacy Worries*, POLITICO (Jan. 1, 2015, 9:00 AM), http://www.politico.com/story/2015/01/energy-electricity-data-use-113901.html.

[65] *See*, *e.g.*, Hayley Tsukayama, *Everything You Need to Know About Google Glass*, WASH. POST (Feb, 27, 2014), http://www.washingtonpost.com/blogs/the-switch/wp/2014/02/27/everything-you-need-to-know-about-google-glass/.

[66] *See*, *e.g.*, Jennifer Booton, *Why Google Glass Wasn't a Failure*, MARKETWATCH (Jan. 30, 2015, 9:14 AM) http://www.marketwatch.com/ story/no-google-glass-wasnt-a-failure-2015-01-29.

[67] *See generally* NAMETAG, http://www.nametag.ws (last visited Jan. 31, 2015).

[68] *See*, *e.g.*, Kashmir Hill, *Google Glass Facial Recognition App Draws Senator Franken's Ire*, FORBES (Feb. 5, 2015, 5:23 PM) http://www.forbes.com/sites/kashmirhill/2014/02/05/google-glass-facial-recognition-app-draws-senator-frankens-ire/.

[69] *Id.*; Jat Singh & Julia Powles, *The Internet of Things – The Next Big Challenge to Our Privacy*, GUARDIAN (July 28, 2014, 3:52 PM), http://www.theguardian.com/technology/2014/jul/28/internet-of-things-privacy; *see also* STOP THE CYBORGS, http://ww.stopthecyborgs.org (last visited Aug. 21, 2014); MICHAEL JUNGLING & PATRICK A. WOOD, MORGAN STANLEY RES., MEDICAL DEVICES AND SERVICES (2014), *available at* http://www.sensium-healthcare.com/sites/default/files/Pages%20from%20morgan_stanley_iot_april_2014.pdf (discussing Sensium's new wearable healthcare patch, data collection, and anonymization); Rachel Metz, *Google Glass is Dead; Long Live Smart Glasses*, MIT TECH. REV. (Nov. 26, 2014), http://www.technologyreview.com/featuredstory/532691/google-glass-is-dead-long-live-smart-glasses/ (describing the failures of the Glass

of ubiquitous technologies and systems capable of tracking individuals in real time through their cell phones and wearable devices as a means of maximizing potential advertising efficacy has already begun.[70]

The Internet of Things market is growing *rapidly*. A recent study by the International Data Corporation estimates that the Internet of Things market worldwide will be worth approximately $7.1 trillion by 2020.[71] Some estimates put the global number of devices connected to the Internet around 6 billion; others estimate that it will reach around 200 billion by 2020.[72] Economic incentives entice increasing numbers of consumers to use devices such as Progressive Insurance's "Snapshot" program, which offers automobile insurance rate reductions in exchange for the installation of a vehicular tracking device that monitors driving speed, time, and habits. As a result, some have argued that this "unraveling of privacy" is creating unprecedented challenges to existing privacy law, which was established primarily with an eye

---

marketing strategy, but concluding that the ubiquity of similar wearables is inevitable).

[70] *See* Stephanie Clifford & Quentin Hardy, *Attention Shoppers: Store Is Tracking Your Cell*, N.Y. TIMES (July 14, 2013), http://www.nytimes.com/2013/07/15/business/attention-shopper-stores-are-tracking-your-cell.html?pagewanted=all; Keith Wagstaff, *New York City Nixes Advertising 'Beacons' in Telephone Booths*, NBC NEWS (Oct. 6, 2014, 4:27 PM), http://www.nbcnews.com/tech/security/new-york-city-nixes-advertising-beacons-telephone-booths-n219281.

[71] *See* Press Release, Int'l Data Corp., The Internet of Things Moves Beyond the Buzz: Worldwide Market Forecast to Exceed $7 Trillion by 2020, IDC Says (June 3, 2014), *available at* http://www.businesswire.com/news/home/20140603005446/en/Internet-Moves-Buzz-Worldwide-Market-Forecast-Exceed#.VN9qEPnF91Z; Chowdry, *supra* note 6. *But see* Marco della Cava, *Privacy Integral to Future of the Internet of Things*, USA TODAY (July 11, 2014, 3:27 PM), http://www.usatoday.com/story/tech/2014/07/10/internet-of-things-privacy-summit/12496613/ (stating that the 2020 market will be worth approximately $15 billion).

[72] Stefan Ferber, *How the Internet of Things Changes Everything*, HARV. BUS. REV. (May 7, 2013), http://blogs.hbr.org/2013/05/how-the-internet-of-things-cha/; Mark Van Rijmenam, *How the Internet of Things Will Make a Smart World—Infographic*, DATAFLOQ, https://datafloq.com/read/internet-of-things-will-make-our-world-smart-infographic/302 (last visited Nov. 26, 2014).

toward preventing uninvited intrusions as opposed to willful surveillance as part of an economic transaction.[73]

Consumers should also not be surprised that the apps they routinely download to smartphones and tablets often collect huge swaths of information about them.[74] The permissions to which consumers must agree as a prerequisite for the download or usage of an app often function as a consumer's consent to have their data tracked and recorded. The Pew Research Group has identified an ever-growing list of over 126 different permissions that apps typically ask for, including a user's location, browser history and bookmarks, calendar events, contact data, cell phone bills, email accounts, mapping applications, and hardware permissions that allow an app to access or use, for instance, a device's camera or

---

[73] *See* Neil M. Richards, *Symposium: Privacy and Technology: The Dangers of Surveillance*, 126 HARV. L. REV. 1934, 1940 (2013) (describing this phenomenon as to the Progressive "MyRate" program); *see also Snapshot Privacy Statement*, PROGRESSIVE.COM, http://www.progressive.com/auto/snapshot-privacy-statement/ (last updated Mar. 11, 2014).

[74] *See, e.g.*, Kenneth Olmstead, *Mobile Apps Collect Information About Users, With Wide Range of Permissions*, PEW RESEARCH (Apr. 29, 2014), http://www.pewresearch.org/fact-tank/2014/04/29/mobile-apps-collect-information-about-users-with-wide-range-of-permissions/; *see also* Chloe Albanesius, *Android Flashlight App Shared User Data Without Permission*, PC MAG. (Dec. 5, 2013, 2:50 PM), http://www.pcmag.com/article2/0,2817,2427999,00.asp (describing the Android Flashlight App's practice of collecting user data and distributing users' personal information to third parties without users' consent, and the settlement with the Federal Trade Commission that followed); Julia Angwin & Jeff Larson, *FAQ About NSA's Interest in Angry Birds and Other 'Leaky Apps,'* PRO PUBLICA (Jan. 28, 2014, 1:30 PM), https://www.propublica.org/article/faq-about-nsas-interest-in-angry-birds-and-other-leaky-apps (describing American and British intelligence agencies' widespread practice of targeting and hacking app developers because of their veritable treasure troves of personal data on individuals); *What They Know – Mobile*, WALL ST. J., http://blogs.wsj.com/wtk-mobile/ (last visited Aug. 22, 2014) (part of an ongoing investigative series into data privacy in which the Wall Street Journal analyzes the data collected by over 100 iPhone and Android apps, and describes what each app stated to users as to the information being gathered).

microphone surreptitiously.[75] Additional personal information is voluntarily surrendered in the form of "loyalty programs" and "customer rewards cards," which companies use to track and analyze user purchases and habits.[76] There are already documented instances of healthcare systems culling information gleaned from these rewards programs and using it to proactively predict the impact of certain buying patterns on an individual's health.[77]

In the interest of fairness, and despite the unsettling overtones of such constant monitoring, there are undoubtedly aspects of this data collection that are not without their upside. After all, online advertising and marketing practices are fueled by big data analytics, and these industries effectively subsidize a tremendous amount of activity online.[78] However, while many consumers appear to be content with this "value for value" exchange, in which

---

[75] Olmstead, *supra* note 74; *see also* Chris Smith, *Facebook's Android App Wants to Do Strange Things to Your Phone*, BGR (Mar. 6, 2014, 2:21 PM), http://bgr.com/2014/03/06/facebook-android-app-permissions/.

[76] *See, e.g.*, Tom Groenfeldt, *Sears Competes on Big Data and Loyalty Programs*, FORBES (May 2, 2012, 10:20 PM), http://www.forbes.com/sites/tomgroenfeldt/2012/05/02/sears-competes-on-big-data-and-loyalty-programs/; *see also* Rajkumar Venkatesan, *Big Data Is an Opportunity to Win More Customers*, WASH. POST (Aug. 17, 2014), http://www.washingtonpost.com/business/capitalbusiness/big-data-is-an-opportunity-to-win-more-customers/2014/08/15/15a31396-2254-11e4-958c-268a320a60ce_story.html; Tom Brewster, *Facebook, Google, and Personal Data: What's Your Worth?*, BBC (May 12, 2014), https://www.bbc.com/future/story/20140509-how-much-is-your-facebook-worth; Ben Kepes, *Is This the Final Straw? Uber's Android Application—"Literally Malware,"* FORBES (Nov. 26, 2014, 6:35 PM), http://www.forbes.com/sites/benkepes/2014/11/26/is-this-the-final-straw-ubers-android-application-literally-malware/.

[77] *See* Shannon Pettypiece & Jordan Robertson, *Hospitals Are Mining Patients' Credit Card Data to Predict Who Will Get Sick*, BUS. WK. (July 3, 2014), http://www.businessweek.com/articles/2014-07-03/hospitals-are-mining-patients-credit-card-data-to-predict-who-will-get-sick; *see also* Kelly Dilworth, *Health Care Companies Turn to Big Data*, YAHOO FIN. (Aug. 14, 2014, 8:00 AM), https://finance.yahoo.com/news/health-care-companies-turn-big-120000707.html; Joseph Walker, *Data Mining to Recruit Sick People*, WALL. ST. J. (Dec. 17, 2013), http://online.wsj.com/news/article_email/SB10001424052702303721045792401405 54518458-lMyQjAxMTA0MDAwNjEwNDYyWj.

[78] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 50.

personal information is readily traded, there is merit in being cognizant of the fact that the Internet is not "free." As one commentator recently put it, "[w]e have become the product."[79]

B. *The Value of Data*

The reason for this seemingly endless collection of data is that this data is proving to be extraordinarily valuable to data brokers, advertisers, investigators, law enforcement agencies, and many others.[80] In 2003, there were already more than one thousand companies conducting data-mining activities on American consumers.[81] In 2012, data was a $300 billion per year industry employing more than three million people in the United States alone.[82] Acxiom Corporation, for instance, sometimes described as "the biggest company you've never heard of,"[83] has been said to have amassed the largest commercial database on consumers

[79] Claire Porter, *Little Privacy in the Age of Big Data*, THE GUARDIAN (June 20, 2014, 12:19 AM), http://www.theguardian.com/technology/2014/jun/20/little-privacy-in-the-age-of-big-data; s*ee also* Katherine J. Strandburg, *Free Fall: The Online Market's Consumer Preference Disconnect*, 2013 U. CHI. LEGAL F. 95, 96 (2013) (noting the failure of the data-for-services economy in part because customers rarely know the cost, thus preventing them from making educated purchase choices based on cost and desire as they would for any other product).

[80] *See*, *e.g.*, WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 41 ("Users, more often than not, do not understand the degree to which they are a commodity in each level of this marketplace."); s*ee also* CUKIER BIG DATA, *supra* note 4, at 98–122 ("The crux of data's worth is its seemingly unlimited potential for reuse."); s*ee also* Ashkan Soltani et al., *NSA Uses Google Cookies to Pinpoint Targets for Hacking*, WASH. POST (Dec. 10, 2013), http://www.washingtonpost.com/blogs/the-switch/wp/2013/12/10/nsa-uses-google-cookies-to-pinpoint-targets-for-hacking/ (describing how the NSA piggybacks into users' systems using private sector cookies).

[81] Andrew J. McClurg, *A Thousand Words Are Worth a Picture: A Privacy Tort Response to Consumer Data Profiling*, 98 NW. U.L. REV. 63, 65 (2003) (citing Robert O'Harrow, *Data Firms Getting Too Personal?*, WASH. POST, Mar. 8, 1998, at A1).

[82] Jason Morris & Ed Lavandera, *Why Big Companies Buy, Sell Your Data*, CNN (Aug. 23, 2012, 8:52 PM), http://edition.cnn.com/2012/08/23/tech/web/big-data-acxiom/ (citing information from the McKinsey Global Institute).

[83] *Id.*

currently in existence.[84] In 2012, Acxiom executives boasted that their database contained information on more than a half billion consumers—including a majority of adults in the United States—with approximately 1,500 individual data points per person.[85] Intelius, Inc. provides its customers with background check and public record information from a database containing more than twenty billion records.[86] PeekYou uses "patented technology that analyzes content from over sixty social media sites, news sources, homepages, and blog platforms to provide clients with detailed consumer profiles."[87]

In order to categorize consumers based on lifestyle, habits, and preferences, some data brokers have identified individuals that fit into certain discrete groups of their own creation, such as "Ethnic Second-City-Strugglers," "Retiring on Empty: Singles," "Tough Start: Young Single Parents," "Credit Crunched: City Families," and "Rural and Barely Making It."[88] Additional categories include "'Rural Everlasting,' which comprises single men and women over the age of 66 with 'low educational attainment and low net worths,'" as well as "Expectant Parent," "Diabetes Interest," and "Cholesterol Focus."[89]

---

[84] Natasha Singer, *Mapping, and Sharing, the Consumer Genome*, N.Y. TIMES (June 16, 2012), http://www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html?pagewanted=all&_r=0.

[85] *Id.* ("[Acxiom] peers deeper into American life than the F.B.I. or the I.R.S., or those prying digital eyes at Facebook and Google. If you are an American adult, the odds are that [Acxiom] knows things like your age, race, sex, weight, height, marital status, education level, politics, buying habits, household health worries, vacation dreams—and on and on.); s*ee also* Philip Bump, *How Facebook Plans to Become One of the Most Powerful Tools in Politics*, WASH. POST (Nov. 26, 2014), http://www.washingtonpost.com/blogs/the-fix/wp/2014/11/26/how-facebook-plans-to-become-one-of-the-most-powerful-tools-in-politics/ (describing the partnership between Facebook and Acxiom in terms of amassing and analyzing individuals' personal data for use in political campaigns).

[86] FTC REPORT, *supra* note 7, at 9; s*ee generally How We Do It*, INTELLIUS, http://corp.intelius.com (last visited Aug. 18, 2014).

[87] FTC REPORT, *supra* note 7, at 9; *see also* PEEKYOU, http://www.peekyou.com (last visited Aug. 18, 2014).

[88] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 44.

[89] FTC REPORT, *supra* note 7, at v.

The U.S. Government Accountability Office ("GAO") recently reported on one data broker that maintains detailed profiles on specific consumers, cataloging ailments ranging from cancer and diabetes to clinical depression and prostate problems.[90] A 2013 Senate report "describes another data broker that keeps 75,000 data elements about consumers in its system, including the use of yeast infection products, laxatives, and OB/GYN services, among other health-related data."[91] All of this collection takes place outside of the regulatory scope of the Health Insurance Portability and Accountability Act ("HIPAA"), which governs patient privacy and confidentiality.[92] Furthermore, the recent shift amongst hospitals to move to digital record keeping has likewise led to patient records—presumably stripped of identifying information[93]—being sold to third-party data-aggregation companies by the state agencies with which hospitals share those records.[94] Similarly, in an effort apparently aimed at skirting certain provisions of the Fair

[90] Julie Brill, Commissioner, FTC, Address at the Woodrow Wilson School of Public and International Affairs, Princeton University: Big Data and Consumer Privacy: Identifying Challenges, Finding Solutions, at 4, Feb. 20, 2014 (citing U.S. GOV'T ACCOUNTABILITY OFFICE, REPORT TO THE CHAIRMAN, COMM. ON COMMERCE, SCIENCE, AND TRANSP., U.S. SENATE, INFORMATION RESELLERS CONSUMER PRIVACY FRAMEWORK NEEDS TO REFLECT CHANGES IN TECH. AND THE MARKETPLACE 53 (2013)), *available at* http://www.ftc.gov/system/files/documents/public_statements/202151/140220princetonbigdata_0.pdf

[91] *Id.* at 4 (citing STAFF OF S. COMM. ON COMMERCE, SCIENCE, AND TRANSP., 113th CONG., A REVIEW OF THE DATA BROKER INDUSTRY: COLLECTION, USE AND SALE OF CONSUMER DATA FOR MARKETING PURPOSES 12, 14 (2013) (citing documentary submission from Equifax and listing health care-related data elements that Equifax maintains)).

[92] *Id.*

[93] *See infra* Part III.D (discussing the failures of anonymization).

[94] *See, e.g.*, Jordan Robertson, *Your Medical Records Are for Sale*, BUS. WK. (Aug. 8, 2013), http://www.businessweek.com/articles/2013-08-08/your-medical-records-are-for-sale; *see also* Jordan Robertson, *States Hospital Data for Sale Puts Privacy in Jeopardy*, BLOOMBERG (June 5, 2013, 12:01 AM), http://www.bloomberg.com/news/2013-06-05/states-hospital-data-for-sale-puts-privacy-in-jeopardy.html; Thomas Claburn, *FTC: Data Brokers Know You Better Than Your Mom Does*, INFO. WK. (May, 28, 2014), https://www.informationweek.com/mobile/mobile-business/ftc-data-brokers-know-you-better-than-mom-does/d/d-id/1269227?piddl_msgid=219930#msg_219930.

Credit Reporting Act ("FCRA"), subprime mortgage and payday lenders are using consumer profiles to identify vulnerable new potential customers.[95]

There are also a growing number of instances of companies using big data analytics to deploy differential pricing models designed to target specific consumers for higher prices based on their consumer profiles.[96] For instance, there are documented examples of consumers paying different prices based on their geographic location.[97] An analysis of the online pricing practices of Staples, Inc., by the Wall Street Journal, for instance, ironically found that areas that had a higher average income tended to be able to purchase a given item from Staples at a lower price.[98] In the same article, it was discovered that Office Depot "uses customers' browsing history and geolocation to vary the offers and products it displays to a visitor to its [web]site."[99] Moreover, this type of behavior is far from unusual. Amazon, Capital One, Discover Financial Services, Orbitz, Lowe's, and Rosetta Stone have all employed big data analytics to vary pricing based on a given consumer's data profile.[100] Differential pricing has become of particular use to companies operating in industries in which prices vary substantially and often, such as in the hotel and airline

---

[95] Nathan Newman, *The Costs of Lost Privacy: Consumer Harm and Rising Economic Inequality in the Age of Google*, 40 WM. MITCHELL L. REV. 849, 857 (2014); *see also* FTC REPORT, *supra* note 7, at i; John Lippert, *Lender Charging 390% Uses Data to Screen Out Deadbeats*, BLOOMBERG (Oct. 3, 2014, 4:49 PM), http://www.bloomberg.com/news/2014-10-01/lender-charging-390-uses-data-to-screen-out-deadbeats.html.

[96] *See, e.g.*, Adam Tanner, *Different Customers, Different Prices, Thanks to Big Data*, FORBES (Mar. 26, 2014, 6:00 AM), http://www.forbes.com/sites/adamtanner/2014/03/26/different-customers-different-prices-thanks-to-big-data/; *see also* Adam Ozimek, *Will Big Data Bring More Price Discrimination?*, FORBES (Sept. 1, 2013, 10:48 AM), http://www.forbes.com/sites/modeledbehavior/2013/09/01/will-big-data-bring-more-price-discrimination/.

[97] Jennifer Valentino-Devries et al., *Websites Vary Prices, Deals Based on Users' Information*, WALL STREET J. (Dec. 24, 2012), http://www.wsj.com/articles/SB10001424127887323777204578189391813881534.

[98] *Id.*

[99] *Id.* (internal quotation omitted).

[100] *Id.*

industries, as well as in industries in which costs and prices are somewhat shrouded, such as in the insurance industry. More and more, differential pricing is being routinely deployed against consumers in the American marketplace.[101]

Given this framework, it is not surprising that some estimates put the value of a single individual's data profile upwards of $5,000 per year.[102] It is perhaps for this reason that some pro-consumer initiatives have developed in an effort to reclaim one's personal data. For instance, the Citizenme initiative seeks to shift Internet power and economics back in the direction of consumers by providing a long-term plan to facilitate the deliberate sale of consumers' personal information directly to specific buyers, rather than having it clandestinely stripped by others.[103] Essentially, Citizenme is an app to which a user would link his or her Facebook, Twitter, and other accounts.[104] It then allows users to see what data is shared on those networks, highlights particularly alarming privacy policy provisions in red, and alerts users and permits them to vote for or against changes made to privacy policies and terms of service.[105] Similarly, DataCoup empowers

---

[101] Tanner, *supra* note 96; *see generally* Walter Baker et al., *Using Big Data to Make Better Pricing Decisions*, MCKINSEY & CO. (June 2014) http://www.mckinsey.com/insights/marketing_sales/using_big_data_to_make_better_pricing_decisions (describing the process of using data to make more profitable pricing decisions).

[102] *See, e.g.*, Newman, *supra* note 95, at 865–66 (citing Quentin Fottrell, *Who Would Pay $5,000 to Use Google? (You)*, MARKET WATCH (Jan. 25, 2012, 12:24 PM), http://blogs.marketwatch.com/realtimeadvice/2012/01/25/who-would-pay-5000-to-use-google-you/?mg=blogs-sm.

[103] *See* Klint Finley, *The App That Lets You Spy on Yourself and Sell Your Own Data*, WIRED (July 9, 2014, 1:55 PM), http://www.wired.com/2014/07/citizenme/.

[104] *Id.*

[105] *Id.*; *see also* Stilgherrian, *Big Data Is Just a Big, Distracting Bubble, Soon to Burst*, ZD NET (July 11, 2014), http://www.zdnet.com/big-data-is-just-a-big-distracting-bubble-soon-to-burst-7000031480 (describing the Respect Network, and the "Login with Respect" initiative); David Braue, *Respect Network Marries Security, Trust in Portable Cloud Data Push*, CSO (July 9, 2014, 9:05 PM), http://www.cso.com.au/article/549571/_respect_network_marries_security_trust_portable_cloud_data_push/.

willing consumers to aggregate, package, and sell their own personal data, thus cutting out the data broker as an unnecessary intermediary.[106] Whether users will ultimately be willing to sell their personal data to brokers and advertisers remains something of an open question. Regardless, brokers are already selling users' information to each other.

C. *The Negotiation of Information*

As one company gleans information, it is sold to another. Of particular interest to attorneys is the trend represented by the emergence of so-called "people search" products offered by data brokers.[107] These products offer personal information about individuals and are unique in that they are marketed for use by individuals rather than businesses, advertisers, or corporations.[108] These products are already capable of providing huge amounts of information on a targeted individual, such as a given person's

> aliases, age and date of birth, news stories, telephone number, gender, interests/affiliations, address history, education information, death records, relatives, employment history, marriage records, email address, criminal records, divorce records, civil records (including bankruptcies, liens, judgments), property ownership and sales history (including loan activity), social media information (including usernames, profile URL, friend connections), [and] neighbors.[109]

The companies that provide people search products often perform sophisticated web crawls across the Internet to gather information

---

[106] *See* DATA COUP, https://datacoup.com/ (last visited Aug. 18, 2014); *see also* Tom Simonite, *Would You Let a Startup Track Your Social Media Accounts and Credit Card Transactions in Exchange for Cash?*, MIT TECH. REV. (Feb. 12, 2014), http://www.technologyreview.com/news/524621/sell-your-personal-data-for-8-a-month/; Tom Simonite, *Datacoup Wants to Buy Your Credit Card and Facebook Data*, MIT TECH. REV. (Sept. 8, 2014), http://www.technologyreview.com/news/530486/datacoup-wants-to-buy-your-credit-card-and-facebook-data/; Sam Harnett, *How to Sell Your Private Data—If You Really Want to*, MARKETPLACE (Aug. 27, 2014), http://www.marketplace.org/topics/tech/how-sell-your-private-data-if-you-really-want.

[107] FTC REPORT, *supra* note 7, at 52; *see, e.g.*, INTELIUS, http://www.intelius.com (last visited Aug. 18, 2014).

[108] FTC REPORT, *supra* note 7, at 52.

[109] *See, e.g.*, *id.* at 53.

on a given subject from publicly available sources and then compare that data to data acquired from other data brokers to gauge the accuracy of their information.[110]

Data brokers sell information to other data brokers, governmental entities, utility and energy companies, hospitality companies, individual consumers, insurance companies, lenders and financial services firms, marketers, advertisers, pharmaceutical companies, real estate services companies, telecommunications firms, attorneys, investigators, and others.[111] Indeed, as informal discovery in civil litigation has become increasingly more productive since the advent of the Internet, some have proactively advocated for the increased use of data brokers' products in facilitating pre-trial adversarial investigation.[112] Moreover, as

---

[110] *Id.* at 56.

[111] *Id.* at 58. While the FTC report's graphic breaking down the types of products purchased by various industries indicates that currently attorneys and investigators are generally only purchasing direct marketing services, it is interesting to note that, as individual consumers frequently use people search services, it may be nearly impossible to accurately ascertain who the end users are of a given data broker product.

[112] *See, e.g.*, Todd B. Baker, *Symposium: The Internet and the Law: Informal Discovery on the Internet*, 52 THE ADVOCATE 23, 27 (2010) (advocating for the employment of Intelius' services in conducting pretrial discovery); *see also* Jayni Foley, *Are Google Searches Private? An Originalist Interpretation of the Fourth Amendment in Online Communication Cases*, 22 BERKELEY TECH. L.J. 447, 473 (2007) ("Data brokering companies now aggregate information on individuals and sell it to both government and private litigants."); Corey Ciocchetti, *The Privacy Matrix,* 12 J. TECH. L. & POL'Y 245, 249 n.10 (2007) ("For years, [data brokers] have made millions quietly selling personal information to law enforcement, corporations, attorneys, collection agencies and the news media." (quoting Jill Burcum, *Hackers' Assaults May Prod Wave of Reforms: Data-Selling Industry Comes Under Scrutiny*, MINNEAPOLIS STAR TRIBUNE, May 29, 2005, at A-1.)). "[M]any data-broker companies such as ChoicePoint and LexisNexis profit from the sale of [personally identifiable information]." *Id.*; s*ee also* Joseph T. Thai, *Symposium: The Jurisprudence of Justice Stevens: Panel I: Criminal Justice: Is Data Mining Ever a Search Under Justice Stevens' Fourth Amendment?*, 74 FORDHAM L. REV. 1731, 1751 (2006) ("Credit card companies, banks, insurers, employers, landlords, attorneys, detectives, angry spouses, and other private parties may avail themselves of the services these data brokers offer.").

attorneys turn to the Internet and, increasingly, to social media to conduct informal discovery, the practice becomes progressively more accepted within the profession from an ethical standpoint, with some even persuasively arguing that there exists an ethical obligation upon attorneys to investigate an opponent's social networking information.[113]

Given these trends, taking the next step to the widespread development of commercially available algorithms to be put to work making deductions about human behavior, lifestyle, and activities from consumers' already available digital footprints is not much of a stretch.[114] If every link clicked indicates an interest; every purchase made demonstrates a trait; and the sum total of individuals' data is being aggregated, bundled, and sold, the focus must be this: what can one realistically do with all of that information? The following section begins to answer this question.

III. THE WORLD OF BIG DATA

"*Technology is neither good nor bad; nor is it neutral.*"[115]

There is no unified definition for the phenomenon that is "big data." Contemporary writers have defined it as "the ability of society to harness information in novel ways to produce useful insights or goods and services of significant value."[116] Others have

---

[113] *See, e.g.*, Steven C. Bennett, *Ethical Limitations on Informal Discovery of Social Media Information*, 36 AM. J. TRIAL ADVOC. 473, 473–74 (2013).

[114] *See, e,g,*, WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 44 (acknowledging the "powerful capacity on the part of the private sector to collect information and use that information to algorithmically profile an individual, possibly without that individual's knowledge or consent[,]" and that, if used "nefariously, could have significant ramifications for targeted individuals").

[115] *See* James R. Hansen, *Technology and the History of Aeronautics: An Essay*, U.S. CENTENNIAL OF FLIGHT COMM'N, http://www.centennialofflight.net/essay/Evolution_of_Technology/Tech-OV1.htm (last visited Jan. 3, 2015) (quoting Melvin C. Kranzberg's First Law of the History of Technology).

[116] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 2; *see also* Sean Fahey, *The Democratization of Big Data*, 7 J. NAT'L SECURITY L. & POL'Y 325, 325 (2014) (defining big data, somewhat frustratingly, "as a collection of data that is

defined it as:

> [A] generalized, imprecise term that refers to the use of large data sets in data science and predictive analytics . . . . First, it refers to technology that maximizes computational power and algorithmic accuracy. Second, it describes types of analyses that draw on a range of tools to clean and compare data. Third, it promotes the belief that large data sets generate results with greater truth, objectivity, and accuracy.[117]

The recent White House Big Data Report correctly noted "most definitions reflect the growing technological ability to capture, aggregate, and process an ever-greater volume, velocity, and variety of data."[118] Essentially, it is a field that applies algorithmic computer processing tools and computer-assisted deductive reasoning to extremely large datasets to make predictions and draw rational conclusions from those datasets. This was not previously possible until recent technological innovations both drove down the costs of data storage and processing while increasing processing power. As with any new and powerful technology, the tools of big data may be harnessed to serve ends either noble—such as the early identification of disease outbreaks—or nefarious.

A. *Predictive Analytics & Deductive Reasoning*

Like "big data," the concept of "predictive analytics" is subject to more than one accepted definition. Some have defined it, simply enough, as "a new discipline that combines data with analysis to

---

so large that it exceeds one's capacity to process it in an acceptable amount of time with available tools").

[117] Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 96 (2014) (citations omitted).

[118] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 2; *see also* Svetlana Sicular, *Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused With the Three V's*, FORBES (Mar. 27, 2013, 8:00 AM), http://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/ (describing the so-called "three V's" and defining "big data" as "high-volume, -velocity and -variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making").

make predictions."[119] As massive amounts of data of all types can now be collected and organized efficiently, highly accurate predictions predicated on that data might now be drawn from the patterns that emerge.[120]

Not so long ago, marketers, researchers, political analysts, and others seeking to research a given problem or phenomenon would use sampling data to arrive at a conclusion.[121] For example, if someone wanted to know more about the political preferences and tendencies of a specific subcategory of the American population, one would submit surveys to a "sample" of several hundred individuals fitting the given demographic and then extrapolate those results to the remaining population.[122] This was a reasonable and manageable method of studying a population and making deductions when costs and practical difficulties prevented researchers from researching or surveying all, or even most, members of a given population. However, with the computing power and storage capacity now available, it has become a debatable issue whether sampling continues to possess its past utility in the age of big data—why analyze only *some* of the data in instances where we now possess the means and the wherewithal to analyze *all* of the data?[123] Despite considerable privacy concerns

---

[119] John O. McGinnis & Russell G. Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 FORDHAM L. REV. 3041, 3052 (2014); *see also* Crawford & Schultz, *supra* note 117, at 98 ("By combining the use of these data sets with predictive analytics, Big Data can dramatically increase the amount of related data that may be considered private.").

[120] *See* McGinnis & Pearce, *supra* note 119.

[121] *See* CUKIER BIG DATA, *supra* note 4, at 23 (describing sampling as "an outgrowth of an era of information-processing constraints").

[122] *See* Deepa Sankar, *Sampling in the Age of Big Data*, SAP (Dec. 11, 2013, 7:23 AM), http://scn.sap.com/community/business-intelligence/blog/2013/12/11/sampling-in-the-age-of-big-data (noting that the typical national polling size is somewhere between 1000 and 1500 participants, with a margin of error of +/- three percentage points); *see also* CUKIER BIG DATA, *supra* note 4, at 12–13 ("Since the nineteenth century, society has depended on using samples when faced with large numbers.").

[123] *See, e.g.*, CUKIER BIG DATA, *supra* note 4, at 13 ("[T]he need for sampling is an artifact of a period of information scarcity, a product of the natural

and potential pitfalls for litigants, the newfound technological capability to draw meaning from information that was practically useless only a few decades ago has led to an astounding array of practical applications, some of which have undeniable societal value and utility.

Google Flu Trends, for example, "uses aggregated Google search data to estimate flu activity" for specific geographic areas and regions.[124] Historically, the Centers for Disease Control and Prevention ("CDC") have monitored flu pandemics in order for public health agencies to properly respond as infections develop and spread.[125] However, a lag time of several weeks or more existed between an epidemic's development in an area and the time when the CDC would receive reports of that epidemic from healthcare professionals and hospitals, delaying the CDC's ability to mount a timely response.[126] Google, on the receiving end of more than three billion search queries per day,[127] found itself in a unique position to speed the flu-recognition process. Given the massive trove of data at its disposal, Google engineers wondered if they could anticipate flu outbreaks and track them in real time by analyzing Google search queries. In a 2009 paper published in the scientific journal *Nature*, Google engineers reported that, by comparing historical search terms with historical flu outbreak information provided by the CDC, they could identify correlations between a combination of forty-five specific search terms and flu

---

constraints on interacting with information in an analog era."); *see also* Steven Swoyer, *Big Data Analytics and the End of Sampling as We Know It*, COMPUTER WEEKLY (Aug. 2012), http://www.computerweekly.com/feature/Big-data-analytics-and-the-end-of-sampling-as-we-know-it.

[124] *See Flu Trends*, GOOGLE, http://www.google.org/flutrends/ (last visited Aug. 22, 2014); *see also* Schwartz & Solove, *supra* note 25, at 1868.

[125] *See generally Seasonal Influenza*, CENT. FOR DISEASE CONTROL, http://www.cdc.gov/flu (last visited Jul. 29, 2014) (providing a hub of resources to help organizations combat flu outbreaks) .

[126] *See* CUKIER BIG DATA, *supra* note 4, at 1–2.

[127] Dan Farber, *Google Search Scratches Its Brain 500 Million Times a Day*, CNET (May 13, 2013, 6:16 PM), http://www.cnet.com/news/google-search-scratches-its-brain-500-million-times-a-day/.

outbreaks in discrete geographical regions.[128] Thus, by analyzing massive amounts of seemingly random data and comparing that data against historically recorded phenomena, accurate predictions could be made about disease trends in real time.

Consider also the possible utility of analyzing aggregated locational data. Global Positioning System ("GPS") technology was opened to non-military uses in the 1980s and, coupled with the steadily decreasing cost of producing GPS modules, has ultimately led to the inclusion of GPS systems in everything from cell phones and computers to the majority of new automobiles.[129] Putting this information to practical use and employing so-called "population analytics," a company called AirSage has a website that boasts, "As long as a mobile phone is active on the cellular network, AirSage receives wireless signals and uses them to anonymously determine location. With AirSage's carrier and partner relationships, we have nationwide coverage—more than any other location-based services (LBS) provider."[130] This information can then be used to identify traffic congestion patterns, groups of migrating protesters, or consumer shopping patterns, based on the number of devices reporting in a given area.[131]

---

[128] Jeremy Ginsberg et al., *Detecting Influenza Epidemics Using Search Engine Query Data*, 457 NATURE 1012, 1014 (Feb. 2009); *see also* CUKIER BIG DATA, *supra* note 4, at 2; Schwartz & Solove, *supra* note 25, at 1868.

[129] *See* CUKIER BIG DATA, *supra* note 4, at 88–89; *see also* Jaclyn Trop, *The Next Data Privacy Battle May Be Waged Inside Your Car*, N.Y. TIMES (Jan. 10, 2014), http://www.nytimes.com/2014/01/11/business/the-next-privacy-battle-may-be-waged-inside-your-car.html (describing so-called "black boxes," which "collect[] information like direction, speed and seatbelt use in a continuous loop. It is in nearly every car today, and in September, it is set to become mandatory.").

[130] AIRSAGE, *How it Works*, http://www.airsage.com/Technology/How-it-works/ (last visited Jan 3, 2015); *see also* Anton Troianovsky, *Phone Firms Sell Data on Customers*, WALL ST. J., (May 21, 2013), http://online.wsj.com/news/articles/SB10001424127887323463704578497153556847658 ("Big phone companies have begun to sell the vast troves of data they gather about their subscribers' locations, travels and web-browsing habits.").

[131] AIRSAGE, *What We Do: See How People Move Through the Day*, http://www.airsage.com/Technology/What-we-do/ (last visited May 20, 2014); *see also* CUKIER BIG DATA, *supra* note 4, at 90–91.

Despite the obvious ingenuity behind such an application of GPS technology, there are also clear privacy implications.[132] As Justice Sotomayor recently stated in her concurring opinion in *United States v. Jones*, "GPS monitoring generates a precise, comprehensive record of a person's public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations."[133] Though such locational information, if perfectly and impenetrably anonymized, may well pose only a small privacy risk to the individual, contemporary research indicates with increasing consistency that true anonymization is likely unattainable.[134]

Law enforcement agencies around the world have also begun employing predictive analytic solutions to tailor specific crime prevention strategies. For instance, Predpol (short for "predictive policing") claims to offer targeted, real-time crime prediction designed for and successfully tested by officers in the field.[135] By forecasting likely future criminal activity in real-time, and basing its calculations on the "times and locations of previous crimes, combined with sociological information about criminal behavior and patterns," the PredPol program recently resulted in a 19% reduction in burglaries in the Santa Cruz, California area; at the time of its introduction, the city was facing a 30% increase in

---

[132] *See generally infra* Part III.D.

[133] United States v. Jones, 132 S. Ct. 945, 955 (2012) (Sotomayor, J., concurring) ("Disclosed in [GPS] data . . . will be trips the indisputably private nature of which takes little imagination to conjure: trips to the psychiatrist, the plastic surgeon, the abortion clinic, the AIDS treatment center, the strip club, the criminal defense attorney, the by-the-hour motel, the union meeting, the mosque, synagogue or church, the gay bar and on and on." (citing People v. Weaver, 909 N.E. 2d 1195, 1199 (2009)).

[134] *See infra* notes 150–61 and accompanying text.

[135] PREDPOL.COM, http://www.predpol.com/ (last visited June 19, 2014); *see also How Predpol Works*, PREDPOL.COM, http://www.predpol.com/how-predpol-works/ (last visited Jan. 3, 2013) ("Using only three data points—past type, place and time of crime and a unique algorithm based on criminal behavior patterns, PredPol's powerful software provides each law enforcement agency with customized crime predictions for the places and times that crimes are most likely to occur.").

crime against a 20% decrease in police staff.[136] IBM also has been developing predictive policing software for several years now, "using databases of past crimes and information like timing and weather to identify trends and map out predictions."[137]

B. *The Power of Correlation*

In order to understand how big data becomes meaningful information through predictive analytics, it is first necessary to discuss the distinction between correlation and causation. *Big Data*, the recent collaboration by Viktor Mayer-Schonberger, Professor of Internet Governance and Regulation at Oxford University, and Kenneth Cukier's, data editor for the Economist, discusses the power and utility of correlative information, given a large enough sample size, even in the absence of causative understanding.[138] The Google Flu program, for instance, was predicated on correlative data—the search query analysis did not *cause* the flu patterns, or vice versa, but the presence of one gave rise to a reasonable probability of the existence of the other.[139]

As an example of this phenomenon, Mayer-Schonberger and Cukier recount the experiences of researchers at the University of Ontario Institute of Technology. Working in conjunction with IBM, the researchers used software to capture patient vital signs—heart rate, blood oxygen levels, and others—in real time. Ultimately, they collected over a thousand data points per second to detect and record subtle changes in the condition of premature babies to detect the onset of health complications and infections in instances where a physician would be incapable of making such a deduction.[140] The information reveals predictable commonalities among infant patients, which occur just prior to the deterioration of a patient's condition. As Mayer-Schonberger and Cukier readily point out, this method is not diagnostic and thus does not illustrate

---

[136] Heather Kelly, *Police Embracing Tech That Predicts Crimes*, CNN (May 26, 2014, 7:08 PM), http://www.cnn.com/2012/07/09/tech/innovation/police-tech/.
[137] *Id.*
[138] *See generally* CUKIER BIG DATA, *supra* note 4, at 90–91.
[139] *Id.* at 53.
[140] *Id.* at 59–60.

why the infant patients are headed for trouble, only that they *are*.[141] Thus, by using computers to detect biological signals common to ailing infants, healthcare workers are able to timely allocate personnel and resources to monitor a patient that the data indicates is at risk, even in the absence of a complete understanding as to why he or she is at risk.[142]

This is the nature of correlative study: the use of thousands of data points, often studied against the backdrop of actual past events, to create predictive models of high probabilities.[143] The aforementioned examples of using massive quantities of data for such diverse purposes as monitoring real-time traffic patterns and predicting health failures through the analysis of thousands upon thousands of data points pertaining to an individual's vital signs offer a brief glimpse of the power of deduction when one possesses enough information on a given subject. Now, given the preceding examples, one must consider the uses to which data brokers, investigators, law enforcement, attorneys, and others could put the massive datasets that result from the wide scale data collection efforts discussed in Part II. The question then becomes whether and to what extent datasets can be linked to a particular person and used to make deductions about that person's traits, habits, medical conditions, political opinions, finances, sexual orientation, psychological conditions, and on and on. The power of aggregated data to identify specific individuals and to identify specific characteristics about them is discussed below.[144]

---

[141] *Id.* at 60; *see also* Brill, *supra* note 90, at 1.

[142] *See also Achieving Small Miracles from Big Data*, IBM, https://www.ibm.com/smarterplanet/global/files/ca__en_us__healthcare__smarter_healthcare_data_baby.pdf (last visited Aug. 18, 2014) (describing the Artemis project).

[143] *See* CUKIER BIG DATA, *supra* note 4, at 68 ("Big data turbocharges non-causal analyses, often replacing causal investigations."). *But see* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 7.

[144] *See, e.g.*, Michael Barbaro & Tom Zeller, Jr., *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?pagewanted=all&_r=0.

C. *The Power of Deduction: Identification, Behavior, and Propensities*

Part of the basic thesis herein—and, indeed, part of the nature of predictive analytics insofar as studying human behavior and characteristics are involved—is that the more information one has on a given individual, the more varied, accurate, and detailed predictions and deductions one can make about that individual.[145] In what is now a widely reported example of big data's ubiquity and potential for invasiveness, in 2002 a statistician employed by Target was at work at his desk when two colleagues stopped by and proposed a question: "If we wanted to figure out if a customer is pregnant, even if she didn't want us to know, can you do that?"[146] The desire to ascertain this information was simple—newly pregnant moms are looked upon as "holy grails" to marketers.[147] This is because research indicates that individuals develop buying habits over time, and those habits are only probable to change upon the occurrence of certain discrete life events, one of the most significant of which is the birth of a child.[148]

Target researchers had discovered that expectant mothers exhibit a number of regular, predictable buying habits. For instance, while lotion is a common purchase among consumers, expectant mothers by and large purchase unscented lotions, and in great quantities, generally around the beginning of the second trimester.[149] A careful retroactive analysis of the company's baby shower registry further showed specific and predictable times at which expectant mothers purchased zinc, calcium and magnesium supplements, hand sanitizer, and an array of other products.[150] Armed with this information, based upon a woman's buying patterns, the statistician was able to create a formula through which

---

[145] *See, e.g.*, *infra* notes 159–61 and accompanying text.

[146] Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES (Feb. 16, 2012), http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=1&_r=1&hp.

[147] *Id.* at 1 ("Their [newly pregnant moms'] shopping patterns and brand loyalties are up for grabs.").

[148] *Id.*

[149] *Id.* at 6.

[150] *Id.*

Target was able to deduce with shocking reliability whether a woman was pregnant, as well as the date of conception, and her due date, in order to timely target the woman with key advertisements as her pregnancy progressed.[151] This practice ultimately received some press attention when it culminated in an irate father storming into a Minneapolis-area Target and demanding an explanation as to why his unmarried teenage daughter was receiving mailers for baby clothes and cribs, only to ultimately apologize to the store manager after returning home to learn of his daughter's unplanned pregnancy.[152] Target knew that the man's daughter was pregnant before he did.

Returning for a moment to the application of predictive analytical models to crime prevention, the previous section touched upon predictive policing, generally, in terms of identifying where and when crimes are likely to transpire based on a historical analysis of the data. However, law enforcement agencies are also applying big data analytics to identify specific *individuals* whom the data indicates warrant additional scrutiny.[153] For instance, the city of Chicago recently used predictive analytics to develop a list of several hundred individuals who fit a demonstrated "profile" for having a propensity for violent criminality.[154] By shifting the focus from geography to identity, and by identifying large numbers of variables that are consistent amongst violent criminals, law enforcement officers are identifying persons for whom they have a "heightened awareness" based on "factors beyond charges and convictions."[155] While it is presently unclear to what extent these

---

[151] *Id.*

[152] *Id.* at 7.

[153] Although the Fourth Amendment implications of making surveillance and investigation decisions based on the development of data profiles are obviously tremendous, this issue is beyond the scope of the instant article. For further reading, see WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 28–31.

[154] *See id.* at 31 (citing Andrew G. Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. (forthcoming 2015), *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2394683; *Predictive Policing Res.*, NAT'L INST. OF JUSTICE (Jan. 13, 2014), http://www.nij.gov/topics/law-enforcement/strategies/predictive-policing/Pages/research.aspx).

[155] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 37.

techniques will be used moving forward, let alone sanctioned by the courts from a Fourth Amendment perspective, the *Minority Report* overtones have not gone unnoticed by either technology writers or privacy advocates.[156]

Programs employing this technology have already been deployed in communities around the country. For instance, a majority of state parole boards now use predictions grounded in data analysis as a factor in determining whether an inmate should be paroled.[157] Similarly, the Department of Homeland Security's Future Attribute Screening Technology ("FAST") project analyzes vital signs, physiological patterns, and the like to identify those who are about to commit crimes.[158] Intelius, one of the nation's largest data brokers with access to over 600 million criminal case records and over 40 million defendant records, has already produced programs that use thousands of criminal records and combine that information with everything from gender, eye, and skin color, to traffic ticket histories and seemingly insignificant data such as an individual's tattoos to create algorithms used to arrive at a probability of an individual's engagement in criminal activity.[159]

---

[156] *See* Yaniv Mor, *Big Data and Law Enforcement: Was "Minority Report" Right?*, WIRED (Mar. 5, 2014, 12:25 PM), http://www.wired.com/2014/03/big-data-law-enforcement-minority-report-right/; *see also* CUKIER BIG DATA, *supra* note 4, at 157–58. *See Minority Report* (Amblin Entm't 2002) (telling the story in which a Washington, D.C., police department develops a "PreCrime" system wherein criminals are clairvoyantly identified, apprehended, sentenced, and jailed prior to having broken the law).

[157] CUKIER BIG DATA, *supra* note 4, at 158; *see also Prison Breakthrough: Big Data Can Help States Decide Whom to Release from Prison*, THE ECONOMIST (Apr. 19, 2014), http://www.economist.com/news/united-states/21601009-big-data-can-help-states-decide-whom-release-prison-prison-breakthrough; Attorney General Eric Holder's Speech at the Nat'l Ass'n of Criminal Defense Lawyers 57th Annual Meeting and 13th State Criminal Justice Network Conference (Aug. 1, 2014), *available at* http://www.justice.gov/iso/opa/ag/speeches/2014/ag-speech-140801.html (discussing the potential pitfalls of employing big data analytics at criminal sentencings).

[158] Mor, *supra* note 156.

[159] Jordan Robertson, *How Big Data Could Help Identify the Next Felon—Or Blame the Wrong Guy*, BLOOMBERG (Aug. 15, 2013, 12:01 AM), http://www.

> The accuracy of the software depends on the number of false positives one is willing to tolerate, a range that [Jim] Adler[, former chief privacy officer at Intelius] calls the "anarchy to tyranny" spectrum. At its most aggressive, his program can correctly identify all 51,246 felons [in his sample set] while misidentifying 2,220 non-felons, numbers an iron-fisted ruler could live with. At a more lenient setting, it can correctly identify 37,842 felons while misidentifying 152 non-felons[.][160]

Similarly, the Department of Homeland Security ("DHS") recently developed the first department-wide big data capability: the dual pilot programs Neptune and Cerberus.[161] Neptune serves as a massive "data lake" into which information from an array of sources flows and is retained.[162] As unclassified data is fed into Neptune, the data is tagged and sorted before being fed into Cerberus, which adds classified information to the mix.[163] These programs provide the ability for investigators to, among other things, "perform person and characteristic searches while investigating a crime."[164]

The deductions that are possible are limited only by the amount of data that is available and the creativity of those mining it. For instance, recent scholarship demonstrates that an analysis of a user's Facebook "likes" "can be used to automatically and accurately predict a range of highly sensitive personal attributes including: sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive

---

bloomberg.com/news/2013-08-14/how-big-data-could-help-identify-the-next-felon-or-blame-the-wrong-guy.html.

[160] *Id.*

[161] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 27.

[162] *Id.* (citing U.S. DEP'T OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR THE NEPTUNE PILOT (Sept. 25, 2013), *available at* http://www.dhs.gov/sites/default/files/publications/privacy-pia-dhs-wide-neptune-09252013.pdf; U.S. DEP'T OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR THE CERBERUS PILOT (Nov. 22, 2013), *available at* http://www.dhs.gov/sites/default/files/publications/privacy-pia-dhs-cerberus-nov2013.pdf).

[163] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 27 n.69.

[164] *Id.* at 28.

substances, parental separation, age, and gender."[165] While many in the private sector have emphasized there is no reason for concern regarding private sector data collection because of policies which dictate that collected data is to be "anonymized," recent research has brought that claim into question.

D. *The Myth of Anonymization*

While the privacy policies of many websites and Internet services state that they will only share non-personally identifiable information[166]—data which cannot be used to indicate an individual's identity—the processes of "de-identification" of aggregated data are becoming less and less effective as re-identification strategies prove to be more and more successful.[167] A recent White House report, for instance, stated:

> As techniques like data fusion make big data analytics more powerful, the challenges to current expectations of privacy grow more serious. When data is initially linked to an individual or device, some privacy-protective technology seeks to remove this linkage, or "de-identify" personally identifiable information—*but equally effective techniques exist to pull the pieces back together through "re-identification."* Similarly, integrating diverse data can lead to what some analysts call the "mosaic effect,"

---

[165] Michael Kosinski et al., *Private Traits and Attributes Are Predictable From Digital Records of Human Behavior*, PROCEEDINGS OF THE NAT'L ACAD. OF SCIENCES (Feb. 12, 2013), *available at* http://www.pnas.org/content/110/15/5802.full.pdf.

[166] *See, e.g.*, *Privacy Policy*, GOOGLE, https://www.google.com/intl/en/policies/privacy/ (last modified Mar. 31, 2014) (stating that it may share "aggregated, non-personally identifiable information publicly and with [its] partners—like publishers, advertisers, or connected sites"); *see also Privacy & Terms: Key Terms*, GOOGLE, https://www.google.com/intl/en/policies/privacy/key-terms/#toc-terms-info (last visited May 29, 2014) (defining "non-personally identifiable information" as "information that is recorded about users so that it no longer reflects or references an individually identifiable user").

[167] *See* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 8 (citing HARVARD LAW PETRIE-FLOM CTR., ONLINE SYMPOSIUM ON THE LAW, ETHICS, & SCIENCE OF RE-IDENTIFICATION DEMONSTRATIONS, PCAST REPORT, BIG DATA AND PRIVACY (2013), *available at* http://blogs.law.harvard.edu/billofhealth/2013/05/13/online-symposium-on-the-law-ethics-science-of-re-identification-demonstrations/) "Many technologists are of the view that de-identification of data as a means of protecting individual privacy is, at best, a limited proposition." *Id.*

> whereby personally identifiable information can be derived or inferred from datasets that do not even include personal identifiers, bringing into focus a picture of who an individual is and what he or she likes.[168]

Mayer-Schonberger and Cukier reached a similar conclusion, ultimately finding that "[g]iven enough data, perfect anonymization is impossible no matter how hard one tries."[169]

In 2006, AOL intentionally released the search queries of 658,000 subscribers to the public for research purposes.[170] Although no names or user IDs were released, AOL assigned individual accounts unique "identification numbers," not dissimilar from the IP addresses that identify each unique Internet connection or the identification numbers assigned to many first- and third-party cookies.[171] However, it was apparent almost immediately that even a novice researcher could deduce extremely intimate details from such information, including a specific user's identity, in short order.[172]

Two New York Times reporters at the time took it upon themselves to attempt to ascertain an individual's identity from his or her search queries alone. User number 4417749 conducted several

---

[168] *Id.* (emphasis added).

[169] CUKIER BIG DATA, *supra* note 4, at 155 ("Researchers have recently shown that not only conventional data but also the social graph—people's connections with one another—is vulnerable to de-anonymization."). *Compare* Ann Cavoukian & Daniel Castro, *Big Data and Innovation, Setting the Record Straight: De-Identification* Does *Work*, PRIVACY BY DESIGN (June 16, 2014), http://www.privacybydesign.ca/content/uploads/2014/06/pbd-de-identification_ITIF1.pdf, *with* Arvin Narayanan & Edward W. Felten, *No Silver Bullet: De-Identification Still Doesn't Work*, RANDOM WALKER (July 9, 2014), http://randomwalker.info/ publications/no-silver-bullet-de-identification.pdf.

[170] Dawn Kawamoto & Elinor Mills, *AOL Apologizes for Release of User Search Data*, CNET (Aug. 7, 2006, 2:30 PM), http://news.cnet.com/AOL-apologizes-for-release-of-user-search-data/2100-1030_3-6102793.html.

[171] *Id.*

[172] *See id.* (illustrating the personal nature of something as seemingly benign as a given person's search history, one search log included queries for "how to tell your family you're a victim of incest;" "casey middle school;" "surgical help for depression;" "can you adopt after a suicide attempt;" "Fishman David Dr. – 2.6 miles NE – 160 E 34th St, New York 10016 – (212) 731-5345;" and "gynecology oncologists in new york city," among others.).

hundred searches over the three-month period for which data was available, for topics such as "numb fingers," "60 single men," and "dog that urinates on everything."[173] Unaided by sophisticated algorithms or computer-assisted analytical tools, the reporters quickly found that as more and more pieces of information were analyzed, the easier it became to establish the user's identity.[174] Additional searches were conducted for "landscapers in Lilburn, Ga," as well as searches for several people with the last name Arnold and "homes sold in shadow lake subdivision gwinnett county Georgia [sic]."[175] This data trail led quickly to Thelma Arnold, a 62-year-old widow who makes her home in Lilburn, Georgia.[176] Indeed, the personal nature of one's casual Internet activity is not to be underestimated. "Foods to avoid when breast feeding," "calorie counting," "how to kill oneself by natural gas," "child porno," "termites," "the best season to visit Italy," "fear that spouse contemplates cheating," and "depression and medical leave" are just a very few examples of the many more hundreds of thousands of search queries detailed in the Times article.[177]

---

[173] Barbaro & Zeller, *supra* note 144.

[174] *See also* CUKIER BIG DATA, *supra* note 4, at 157 ("In order to fully investigate an individual, analysts need to look at the widest possible penumbra of data that surrounds the person—not just whom they know, but whom those people know too, and so on.").

[175] Barbaro & Zeller, *supra* note 144.

[176] *Id.*

[177] *Id.* There is already empirical evidence that the pervasive tracking of Internet activity, by both private entities as well as government bodies such as the National Security Administration (NSA), is having a worldwide speech chilling effect. A recent paper by the Massachusetts Institute of Technology's Catherine Tucker and Alex Matthews, entitled *Government Surveillance and Internet Search Behavior*, details the changes in Internet activity across populations globally in response to the June 2013 revelations that the NSA has been cooperating with major tech companies such as Microsoft, Google, and Yahoo! to obtain real-time data content on individual users. *See* Alex Matthews & Catherine Tucker, *Government Surveillance and Internet Search Behavior* (Mar. 24, 2014), *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2412564. By using Google Trends to investigate 282 search terms—the search terms used in the study derived from a list of search terms that DHS tracks on social media sites, Google's top 50 search terms for 2013, and other potentially embarrassing search terms—the researchers discovered a measurable decrease in

Data brokers already sell data profiles on individuals to consumers of all types. They already comb the Internet for information, and purchase information from first- and third-party websites, and from one another. Given the ability of private companies to collect personal information ranging from what type of cologne a person bought their grandfather last Christmas to whether or not that person has herpes, as well as the emergence of the developing field of predictive analytics, the ability to create increasingly detailed data profiles on individuals grows by the day. As the products offered by data brokers become more sophisticated, accurate, and invasive moving forward, the value of this information to attorneys and their clients should not be underestimated. In the current, largely unregulated, environment, where personal information is readily sold as a commodity, the risk that such personal information will someday be used against a person increases with each day that passes, each transfer of personal data between parties, and each click and keystroke. As such, the only true protection a person has is to limit the information that he or she volunteers to the world, to the best that they are able. Furthermore, and perhaps most troubling, as discussed below, the only true restraints on the industry are self-imposed.

IV. THE CURRENT LEGAL AND REGULATORY LANDSCAPE

As new technologies have emerged and forced the law to adapt, privacy law in the United States has historically been greatly influenced by public opinion. For instance, in the Supreme Court's

---

the frequency of searches for both search terms which were either potentially embarrassing or likely to be flagged by the NSA. *Id.* at 3. These search terms run the gamut of everything from "abortion" and "Accutane" to "flu" and "dirty bomb." *Id.* at 33–37. While the First Amendment implications are immediately apparent given the presence of empirical data demonstrating self-censorship in response to governmental surveillance, while also acknowledging that the issue of government partnership with, and hacking of, private companies to facilitate mass data collection further complicates the matter, such considerations are beyond the scope of this Article.

1928 decision in *Olmstead v. United States*,[178] the Court held that the practice of wiretapping phone lines did not infringe upon an individual's Fourth Amendment rights.[179] *Olmstead* essentially permitted law enforcement officers to listen in on citizens' telephone calls with impunity. Then, in 1967, the Supreme Court set forth the reasonable expectation of privacy test in *Katz v. United States*,[180] and found that "one who occupies it, [a telephone booth] shuts the door behind him, and pays the toll that permits him to place a call is surely entitled to assume that his conversation is not being intercepted."[181] This was the result of changes in public attitudes about privacy, as well as technological developments in the intervening years between *Olmstead* and *Katz* that helped to bring about a change in public sentiment, which together effected a change in the constitutionally-protected status of Americans' telephone conversations.[182] These decisions are noteworthy in that they together demonstrate the phenomenon of slow-moving sea changes in the way a society and, accordingly the law, views complex and emerging issues that collide at the intersection of law and technology.

There may never be a *Katz* moment for the Internet, wherein the High Court sweeps down to protect the rights and privacy of all Internet users. Indeed, attorneys charged with safeguarding the interests of their clients must assume that there never will be. Moreover, issues of user consent, difficulties in determining data ownership, the international nature of the Internet, implicit constitutional questions about individual rights, and additional complications not yet thought of may ultimately prove too problematic for a comprehensive piece of legislation or a single Supreme Court decision to address all attendant privacy

---

[178] 277 U.S. 438 (1928).

[179] *Id.* at 466.

[180] 389 U.S. 347 (1967).

[181] *Id.* at 361 (internal quotation omitted).

[182] *See* Ken Gormley, *One Hundred Years of Privacy*, 1992 WIS. L. REV. 1335, 1362–70 (1992) (describing the years following *Katz*, technological advances in surveillance capabilities, and changing social and political ideas about privacy).

concerns.[183] Nevertheless, the attorney's duty to provide competent representation remains, regardless of whether the law is able to keep pace with technology.

The protection of one's digital data privacy in the United States is grounded in principles of contract and tort law and subject to very little regulation.[184] Actions sounding in privacy or tort, however, have enjoyed little successful application to online data collection, in no small part because the privacy policies deployed by most websites and digital services currently operate as blanket customer consent forms to use an individual's personal data as the holders of that data see fit.[185] Thus, tort and contract remedies, as well as actions under the few applicable federal statutes, have had little practical success for parties aggrieved by private sector data collection practices.

A. *Privacy Policies & The Problem of Consent*

The privacy policies of virtually all websites and Internet services describe, generally in the vaguest possible terms, what data is collected and what uses are made of the data that the user consents to share by using that website.[186] This "notice and

---

[183] *See, e.g.*, WHITE HOUSE BIG DATA REPORT *supra* note 4 n.9 ("Harvard Professor of Science & Technology Studies Sheila Jasanoff argues that framing the policy implications of big data is difficult precisely because it manifests in multiple contexts that each call up different operative concerns, including big data as property (who owns it); big data as common pool resources (who manages it and on what principles); and big data as identity (it is us ourselves, and thus its management raises constitutional questions about rights).").

[184] *See* Lori Chiu, *Drawing the Line Between Competing Interests: Strengthening Online Data Privacy Protection in an Increasingly Networked World*, 14 SAN DIEGO INT'L L. J. 281, 282–83 (2013) (citing Carolyn Hoang, *In the Middle: Creating a Middle Road Between U.S. and EU Data Protection Policies*, 32 J. NAT'L ASS'N ADMIN. L. JUD. 810, 818 (2012)); *see also* Klinefelter, *supra* note 25, at 19.

[185] *See, e.g.*, Deering v. CenturyTel, Inc., No. cv-10-63-BLG-RFC, 2011 U.S. Dist. LEXIS 51930 (D. Mont. May 16, 2011); In re Google, Inc., Privacy Policy Litig., No. C-12-1382-PSG, 2013 U.S. Dist. LEXIS 171124, at *39–44, *49–51 (N.D. Cal. Dec. 3, 2013); *see generally infra* Part IV.A.

[186] *See, e.g.*, *Privacy Policy*, FOX NEWS, http://www.foxnews.com/about/privacy-policy (effective as of Jul. 1, 2013) ("By using Fox News Services, you

consent" model has been a central tenet of modern privacy law that has permitted individuals to determine the manner and circumstances in which their personal information may be shared.[187] Not surprisingly, defenses predicated on users' consent have already been successfully deployed to claims mounted against data collectors under both tort theories and violations of the Electronic Communications Privacy Act ("ECPA").[188]

---

agree to the terms and conditions of this Privacy Policy.") The policy further states that Fox News and its service providers collect registration information, public information and posts, information from social media, and activity information, including, for instance, IP address, browser type, geolocation data, and other information. *Id.* It also notes that Fox News does "not respond to browser 'Do Not Track' signals, as we await the work of interested stakeholders and others to develop standards for how such signals should be interpreted." *Id.*; *see also Full Privacy Policy*, NBC UNIVERSAL, http://www.nbcuni.com/privacy/full-privacy-policy/#what_information_do_we_collect_and_how_is_it_used (last updated May 30, 2014) ("By using the online services, you expressly consent to our collection, use, disclosure, and retention of your personal information as described in this Privacy Policy.") NBC's policy further describes the collection of information such as name, home address, age, gender, phone number, email address, payment information, photos or videos of users, information about one's Internet connection, transaction information, "pages that you visit within the online services, gameplay data or other information collected through Cookies and Tracking Technologies[,]" and information collected from social networks and other publicly available data. *Id.* The NBC Universal policy goes on to state that they may "from time to time transfer your personal information to other countries and make it accessible to any of our affiliates and third-party service providers internationally." *Id.*; *see also Privacy Policy*, BUZZFEED, http://www.buzzfeed.com/about/privacy (last visited Jul. 30, 2014) (describing its data collection practices, and then stating that, "[i]n some cases, we may choose to buy or sell assets. In these types of transactions, user information, including Personal Information, is typically one of the transferred business assets").

[187] *See, e.g.*, WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 49 (describing the notice and consent model as "the core tenet of modern privacy protection . . . that has been in wide use since the 1970s); CUKIER BIG DATA, *supra* note 4, at 173 ("For decades an essential principle of privacy laws around the world has been to put individuals in control by letting them decide whether, how, and by whom their personal information may be processed.").

[188] *See, e.g.*, *Deering*, 2011 U.S. Dist. LEXIS 51930, at *1 (presenting a case in which the plaintiff sued after the defendant-Internet service provider's collected and diverted its customers' Internet communications to third parties,

A related problem is the *scope* of consent. Privacy policies are broadly written so that they may be broadly construed, in large part because most companies in the business of collecting data today have no idea to what use that data could be put in the future. A great deal of the value offered by huge datasets comes as a consequence of secondary uses, sometimes only discovered months or years after the data was first collected. As a result, providing adequate notice to consumers becomes less and less realistic.[189] Indeed, many are coming to what should have been the obvious conclusion that the focus of most websites' privacy policies is on protecting data collection practices rather than the privacy of users and visitors.[190] Moreover, the lingering question remains insofar as to what meaningful consent actually exists when recent research indicates that at least one Americans in every two erroneously believes that a privacy policy "ensures that the company keeps confidential all the information it collects on users."[191]

A related problem to the scope of consent is the frequency with which companies amend their terms and conditions and privacy policies regarding data collection. The privacy policies of many major websites and services are amended so often that other website services have sprung up for the sole purpose of monitoring changing website terms and conditions.[192] In December 2014, for

---

and the court granted the defendant's motion to dismiss the plaintiff's claims under the ECPA, as well as the claims for invasion of privacy, based on the plaintiff's "consent"); *see also infra* Part IV.B.

[189] CUKIER BIG DATA, *supra* note 4, at 173.

[190] *See, e.g.*, Jose Pagliery, *What You Really Agree to When You Click 'Accept*,*'* CNN MONEY (May 19, 2014, 9:15 AM), http://money.cnn.com/2014/05/19/technology/security/privacy-policy/index.html (describing most website privacy policies as "unintelligible"); *see generally Terms and Conditions May Apply* (Hyrax Films 2013).

[191] Aaron Smith, *Half of Online Americans Don't Know What a Privacy Policy Is*, PEW RESEARCH CTR. (Dec. 4, 2014), http://www.pewresearch.org/fact-tank/2014/12/04/half-of-americans-dont-know-what-a-privacy-policy-is/.

[192] *See* Pagliery*, supra* note 190; *see, e.g.*, TOSBACK, https://tosback.org/ (last visited Aug. 23, 2014); TERMS OF SERVICE; DIDN'T READ, http://www.tosdr.org (last visited Aug. 23, 2014).

instance, the website tosback.org, a terms of service tracking collaboration between the Electronic Frontier Foundation,[193] the Internet Society,[194] and ToS;DR (which is itself a tongue-in-cheek shorthand for "Terms of Service; Didn't Read"),[195] reported over forty significant changes in the terms of service or privacy policies of major websites, including Google, Gmail, Yahoo, LinkedIn, Youtube, and Flickr, among others.[196] While, arguably, these constant alterations should diminish the legal efficacy of employing user consent as a defense to suits brought by consumers seeking to prevent companies' data collection practices, there is little evidence to support that such arguments are having any success.

Facebook's data use policy, for instance states, "We receive data about you whenever you use or are running Facebook[.]"[197] It further states, "We receive data whenever you visit a game, application, or website that uses Facebook Platform or visit a site with a Facebook feature . . . ."[198] It goes on to state that "an advertiser may tell us information about you," and "[w]hen we get your GPS location, we put it together with other location information we have about you . . . ."[199] Similarly, Twitter's privacy policy states,

> When you use our Services, we may receive information ("Log Data") such as your IP address, browser type, operating system, the referring web page, pages visited, location, your mobile carrier, device information (including device and application IDs), search terms, and cookie information . . . . We may revise this Privacy Policy from time to time . . . . If we make a change to this policy that, in our sole discretion, is material, we will notify you . . . .[200]

---

[193] ELEC. FRONTIER FOUND., https://www.eff.org/ (last visited May 22, 2014).
[194] INTERNET SOCIETY, http://www.internetsociety.org/ (last visited May 22, 2014).
[195] TERMS OF SERVICE; DIDN'T READ, http://tosdr.org/ (last visited Aug. 23, 2014).
[196] *See* TOSBACK, https://tosback.org (last visited Aug. 23, 2014).
[197] *Data Use Policy → Information We Receive about You*, FACEBOOK, http://www.facebook.com/about/privacy/your-info (last visited Jan. 4, 2015).
[198] *Id.*
[199] *Id.*
[200] *Twitter Privacy Policy*, TWITTER, twitter.com/privacy (last visited Jan. 4, 2015).

Not surprisingly, Google's Terms of Service states that by using Google's services, you agree to its terms.[201] Google's privacy policy then states that Google "may share aggregated, non-personally identifiable information publicly and with [its] partners[;]" that Google "will share personal information with companies, organizations or individuals outside of Google if [it has] a good-faith belief that access, use, preservation or disclosure of the information is reasonably necessary to" comply with any legal process or enforceable governmental request; and that "[i]f your Google Account is managed for you by a domain administrator (for example, for Google Apps users) then your domain administrator and resellers who provide user support . . . will have access to your Google Account information (including your email and other data)."[202] In the face of such user agreements, it is difficult to determine exactly which information, if any, users have *not* consented to disclosing. From the consumer's perspective, the lack of an available common law cause of action capable of deterring the collection of one's data is compounded by the somewhat vacuous state of legislation currently in place at the federal level.[203]

B. *Federal Law & Data Privacy*

At present, information privacy law in the United States is governed by a random assortment of federal and state statutes which focus on very specific areas—such as healthcare, credit reporting, and video rental records, among others—rather than by

[201] *Terms of Service*, GOOGLE, https://www.google.com/intl/en/policies/terms/?fg=1 (last modified Apr. 14, 2014).

[202] *Privacy Policy*, GOOGLE, https://www.google.com/intl/en/policies/privacy/?fg=1 (last modified Dec. 19, 2014).

[203] It should be noted that, while there have been some efforts by states to craft effective legislation in this area, it has focused primarily on beefing up consumer protections in the event of data breaches. *See, e.g.*, Florida Information Protection Act of 2014, Fla. S.B. 1524 (2014), *available at* https://www.flsenate.gov/Session/Bill/2014/1524/BillText/er/PDF. *But see* Cal. Assembly Bill No. 2306 (Sept. 30, 2014), *available at* http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201320140AB2306 (seeking to expand the scope of California's invasion of privacy statute to permit broader consumer protections).

any uniform legislative act or regulatory body.[204] Generally speaking, the Electronic Communications Privacy Act ("ECPA") of 1986 is still the primary piece of federal legislation affecting data privacy.[205] However, the ECPA was passed at a time when it was not uncommon for a person to pull over one's car to use a payphone after one's pager went off.[206] In short, the statute is woefully inadequate and antiquated.

---

[204] *See* Solove, *supra* note 24, at 1440–44 (describing the FCRA, the Privacy Act of 1974, The Family Educational Rights and Privacy Act of 1974 (FERPA), the Cable Communications Policy Act (CCPA) of 1984, the Electronic Communications Privacy Act (ECPA), the Video Privacy Protection Act of 1988, the Telephone Consumer Protection Act of 1991, the Driver's Privacy Protection Act of 1994, the Health Insurance Portability and Accountability Act (HIPAA) of 1996, and the Children's Online Privacy Protection Act (COPPA) of 1998, and concluding that, rather than following, for instance, the footsteps of the European Union in adopting large-scale privacy protections, "Congress has passed a series of statutes narrowly tailored to specific privacy problems"); Devin W. Ness, *Note: Information Overload: Why Omnipresent Technology and the Rise of Big Data Shouldn't Spell the End for Privacy as We Know It*, 31 CARDOZO ARTS & ENTM'T L.J. 925, 944 (2013). (citing Fair Credit Reporting Act, 15 U.S.C. §§ 1681–81(u); ECPA of 1986, 18 U.S.C. §§ 2510–2522, 2701–2712, 3121–3127; Privacy Act of 1974, 5 U.S.C. § 552a; Family Educational Rights and Privacy Act of 1974, 20 U.S.C. § 1232g; Right to Financial Privacy Act of 1978, 12 U.S.C. §§ 3401–3422; Cable Communications Policy Act of 1984, 47 U.S.C. 521–573; and Video Privacy Protection Act of 1988, 18 U.S.C. § 2710 (additional citations omitted)) (discussing Congress' historically piecemeal and seemingly arbitrary approach to privacy legislation).

[205] *See* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 33 (stating that the ECPA protects stored electronic communications); *see also* Erica M. Scott, *Protecting Consumer Data While Allowing the Web to Develop Self-Sustaining Architecture: Is a Trans-Atlantic Browser-Based Opt-In for Behavioral Tracking the Right Solution?*, 26 PAC. MCGEORGE GLOBAL BUS. & DEV. L.J. 285, 298 (2013) (citing Declan McCullagh, *Google, Facebook Go Retro in Push to Update 1986 Privacy Law*, CNET (Oct. 21, 2011, 8:50 AM), http://news.cnet.com/8301-1009_3-20004071-83.html) (stating that the ECPA, "promulgated in 1986, before the Internet reached beyond university campuses, is still the primary piece of legislation that affects data privacy on the Internet").

[206] *See* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 49 (noting that, at the time of the ECPA's passage, most important documents were kept in hard copies in the home).

The ECPA includes the Stored Communications Act ("SCA"),[207] which addresses private data collection more closely than any other federal statute,[208] and the Wiretap Act, which litigants have also attempted to use as a vehicle to challenge private sector data collection.[209] The Wiretap Act provides that "any person whose wire, oral, or electronic communication is intercepted, disclosed, or intentionally used in violation of this chapter may in a civil action recover from the person or entity" responsible.[210] However, litigants have had little success utilizing this statute, in part, because the statutory language only prohibits interception of the "contents" of a message, and courts have held that automatically generated data, such as geolocation information that is perpetually sent to service providers does not constitute "content" under the statute.[211] This approach appears to be attaining increasing support among a majority of federal courts.[212] This is troubling because, as previously discussed, such information can be subjected to

---

[207] 18 U.S.C. §§ 2701–2710.

[208] *See* Kesan et al., *supra* note 31, at 399–400 (discussing the three components of the ECPA—the Wiretap Act, the SCA, and the Pen Register statute); *see also* Orin S. Kerr, *The Next Generation Communications Privacy Act*, 162 U. PA. L. REV. 373, 383 (2014) (describing the SCA as "by far the most important" section of the ECPA).

[209] *See* 18 U.S.C. §§ 2510–2522, 2701–2712, 3121–3127; *see, e.g.*, *In re* Doubleclick Privacy Litig., 154 F. Supp. 2d 497, 519 (S.D.N.Y. 2001).

[210] 18 U.S.C. § 2520(a); *see also* In re iPhone App. Litig., 844 F. Supp. 2d 1040, 1055, 1061–63 (N.D. Cal. 2012).

[211] *See In re* iPhone App. Litig., 844 F. Supp. 2d at 1062 (citing 18 U.S.C. § 2510(5)) (additional citations omitted).

[212] *See, e.g.*, In re Google, Inc., Cookie Placement Consumer Privacy Litig., 988 F. Supp. 2d 434, 443–44 (D. Del. 2013) (citing *In re* iPhone App. Litig., 844 F. Supp. 2d at 1062) ("[P]ersonally identifiable information that is automatically generated by the communication" is not "contents" for the purposes of the Wiretap Act); Sams v. Yahoo!, Inc., No. 10-5897, 2011 U.S. Dist. LEXIS 53202, at *6–7 (N.D. Cal. May 18, 2011) (holding that records identifying persons using Yahoo ID and email address, IP addresses, and login times were not content-based); *In re* § 2703(d) Order, 787 F. Supp. 2d 430, 435–36 (E.D. Va. 2011) (holding that the Wiretap Act did not cover unique Internet Protocol ('IP') number, Twitter subscriber, user, and screen names, addresses (including e-mail addresses), telephone or instrument number or other subscriber number or identity, and temporarily assigned network address").

algorithmic analyses to draw highly intimate conclusions about individuals. Other courts have refused to find liability on the part of third-party data collectors under the Wiretap Act, because the consumer-plaintiffs consented to the data collection in the first place through acquiescence to website terms and conditions, and consent constitutes a statutory exception to liability.[213]

As to the SCA: without focusing too much herein on what has become a relatively antiquated statutory distinction between the providers of electronic communications services ("ECS") versus providers of remote computing services ("RCS"),[214] the SCA addresses both the circumstances in which the government may compel providers to disclose information about consumers as well as the circumstances in which the providers may voluntarily disclose such information to third parties.[215] Essentially, the SCA prohibits providers from voluntarily disclosing the *contents* of a communication subject to seven exceptions, one of which is when the disclosure occurs "with the lawful *consent* of the originator or an addressee."[216] The SCA also permits providers to divulge "a *record* or other information pertaining to a subscriber to or customer of such a service . . . with the lawful consent of the customer or subscriber."[217] Further, there is also a provision of the SCA that goes so far as to state that service providers may disclose user "records" (but not communications' "contents") "to any

---

[213] *See, e.g.*, *In re* Doubleclick Privacy Litig., 154 F. Supp. 2d 497, 514 (S.D.N.Y. 2001); 18 U.S.C. § 2511(2)(d).

[214] A "remote computing service" is statutorily defined as "the provision to the public of computer storage or processing services by means of an electronic communications system." 18 U.S.C. § 2711(2). An "electronic communication service" is defined as "any service which provides to users thereof the ability to send or receive wire or electronic communications." 18 U.S.C. § 2510(15); *see also* Kerr, *supra* note 208, at 395–97 (describing this distinction as "obsolete," while also concluding that "ECPA likely offers no protection for access to stored search queries . . . because it does not fit the 1986 dichotomies codified by the statute").

[215] 18 U.S.C. §§ 2702–2703; *see also* Kerr, *supra* note 208, at 384.

[216] 18 U.S.C. § 2702(b) (emphasis added).

[217] 18 U.S.C. § 2702(c)(2); *see also* Neil M. Richards & Jonathan H. King, *Big Data Ethics*, 49 WAKE FOREST L. REV. 393, 417 (2014) (discussing the limited protections available to personal metadata).

person other than a governmental entity."[218] Given the broad scope of most online privacy policies, it is not difficult to see why consumers have had little success in challenging private sector data collection practices. Moreover, while the SCA also covers the instances in which service providers shall turn over the content of users' communications as well as the records of those communications to the government,[219] one can see little remaining value in any protections provided to online activities from governmental intrusion when there are documented instances of providers such as Google voluntarily scanning users' email accounts in search of evidence of criminality and then turning the evidence over to law enforcement.[220]

Recent attempts to amend the ECPA to keep pace with developing technologies have stalled in Congress and, in any event, have focused more on updating the warrant requirement for law enforcement access to emails rather than addressing private sector data collection.[221] Notably, in February of 2014, Senators

---

[218] 18 U.S.C. § 2702(c)(6).

[219] As to governmental acquisition of either the "contents" of an electronic communication, or the "records" of communications, and illustrating the antiquated nature of the SCA, the SCA prohibits service providers from granting a governmental entity access to the contents of an electronic communication, such as an email, without a warrant, *unless* the message is more than 180 days old. 18 U.S.C. § 2703(a); *see also* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 60. However, after 180 days, the government may obtain the contents of any electronic communications pursuant to either a simple administrative subpoena or a court order to an ECS provider; consumer "records" may be obtained with as little as an administrative subpoena, or where the subscriber has consented to the disclosure. *See* 18 U.S.C. § 2703(c). Court orders shall issue upon a showing that there "are reasonable grounds to believe that the contents of a wire or electronic communication, or the records or other information sought, are relevant and material to an ongoing criminal investigation." 18 U.S.C. § 2703(d).

[220] *See, e.g.*, Conor Dougherty, *Google Gives Child Pornography Email Evidence to Police*, N.Y. TIMES, (Aug. 4, 2014), http://bits.blogs.nytimes.com/2014/08/04/google-gives-child-pornography-email-evidence-to-police/; *see also* 18 U.S.C. § 2703(e) (providing insulation from liability for providers in cooperating with law enforcement).

[221] *See* S. 607, 113th Cong., 1st Session, *available at* http://thomas.loc.gov/cgi-bin/query/z?c113:S.607: (known as the "Leahy-Lee ECPA Amendments

John D. Rockefeller and Edward Markey introduced the Data Broker Accountability and Transparency Act of 2014, which aims to provide consumers with the right to access to their personal files held by data brokers, correct inaccuracies therein, and decide for themselves whether they want to permit their information to be sold.[222] However, this legislation has received little in the way of press attention or legislative traction since its debut.[223] This has resulted in a situation wherein consumers wishing to prevent the widespread collection and analysis of their data are essentially without a remedy.

For instance, in January 2000, a class action lawsuit was filed against DoubleClick, Inc., a Delaware corporation that was the largest provider of Internet advertising services in the world at the

---

Act," this bill was unanimously reported out of the Senate Judiciary Committee in 2012, but has not advanced, largely due to pressure from the Securities and Exchange Commission, as well as other governmental agencies, which lacks warrant authority and wishes to continue to be able to obtain emails by subpoena); *see also* Cameron F. Kerry, *Microsoft Challenges the Government: Litigating Extraterritoriality in a Virtual World*, BROOKINGS (Jul. 31, 2014, 7:30 AM), http://www.brookings.edu/blogs/techtank/posts/2014/07/31-microsoft-ireland-lawsuit-kerry; Kate Tummarello, *Obama's 'Big Data' Report Calls for New Privacy Laws*, THE HILL (May 1, 2014, 2:28 PM), http://thehill.com/policy/technology/204961-white-house-big-data-report-calls-for-new-privacy-laws; Erin Mershon, *USA Freedom Is Out, Now What?—Buzz: Markey, Hatch to Introduce Bill on Student Privacy—Rockefeller Ready to Crackdown on Cramming*, POLITICO (Jul. 30, 2014, 10:01 AM), http://www.politico.com/morningtech/0714/morningtech14827.html; Julian Hattem, *Tech's Bad Year in Washington*, THE HILL (Jan. 3, 2015, 6:11 AM), http://thehill.com/policy/technology/227863-techs-bad-year (describing the Email Privacy Act's failure to even get out of committee).

[222] S. 2025, 113th Cong., 2D Session (2014) (staff working draft), *available at* http://www.commerce.senate.gov/public/?a=Files.Serve&File_id=13d141a3-76b8-4191-810b-ebbfd5125759; Press Release, Democratic Press Office, Rockefeller, Markey, Introduce Data Broker Bill to Ensure Accuracy, Accountability for Consumers (Feb. 12, 2014), *available at* http://www.commerce.senate.gov/public/index.cfm?p=PressReleases&ContentRecord_id=5b5b0622-1c5b-4584-91c2-769f9b009778.

[223] *See, e.g.*, David Lazarus, *Adjusting for Life in the World of Big Data*, L.A. TIMES (Aug. 25, 2014, 1:37 PM), http://www.latimes.com/business/la-fi-lazarus-20140826-column.html (stating that the Act "has gone nowhere in the Senate").

time.[224] The plaintiffs filed suit seeking injunctive and monetary relief for DoubleClick's online data collection practices following DoubleClick's 1999 acquisition of Abacus Direct Corporation, which possessed a "database of names, addresses, telephone numbers, retail purchasing habits and other personal information on approximately ninety percent of American households, which it sold to direct marketing companies."[225] The suit essentially sought to prevent DoubleClick from combining its database of online profiles with Abacus' database of offline profiles "in order to create a super-database capable of matching users' online activities with their names and addresses."[226]

Emblematic of the lack of legal tools with which the plaintiffs could prevent these data collection practices, the court held that the plaintiffs could not succeed under the SCA because they could not show that Doubleclick's placement of cookies on users' computers and subsequent collection of data was unauthorized.[227] The court likewise found that the plaintiffs had no claim under the Federal Wiretap Act,[228] because the "DoubleClick-affiliated Web sites [were] 'parties' to the plaintiffs' intercepted communications under the Wiretap Act and . . . they consented to DoubleClick's interceptions."[229] Further, the plaintiffs' claims under the Computer Fraud and Abuse Act[230] ("CFAA") failed, as the court held that the plaintiffs

---

[224] *In re* Doubleclick Privacy Litig., 154 F. Supp. 2d 497, 500 (S.D.N.Y. 2001).

[225] *Id.* at 505. This lawsuit also closely followed an FTC investigation that had concluded that DoubleClick had violated no U.S. laws. *Id.* at 506.

[226] *Id.* at 505.

[227] *Id.* at 507, 513–14.

[228] *See* Wiretap Act, 18 U.S.C. §§ 2510–2522; *see also* 18 U.S.C. § 2511 ("[A]ny person who intentionally intercepts, endeavors to intercept, or procures any other person to intercept or endeavor to intercept, any wire, oral, or electronic communication . . . shall be punished . . . .").

[229] *Doubleclick*, 154 F. Supp. 2d at 519. *See* 18 U.S.C. § 2511(2)(d) ("It shall not be unlawful . . . for a person not acting under color of law to intercept a wire, oral, or electronic communication where such person is a party to the communication or where one of the parties to the communication has given prior consent to such interception . . . .").

[230] 18 U.S.C. § 1030.

could not possibly allege the statutory damages threshold with regard to any particular computer.[231]

This type of failure has been the rule rather than the exception in lawsuits filed by consumers attempting to challenge the current state of online data collection. In *Kirch v. Embarq Management Company*,[232] the Tenth Circuit affirmed the district court's grant of summary judgment to the defendant-Internet service provider ("ISP"), finding that the ISP did not "intercept" the plaintiffs communications under the Wiretap Act in part because the ISP's funneling of the plaintiffs' information to an advertiser was carried out in the "ordinary course of its business," thus placing the activity outside of the statutory definition of "interception"—another exception.[233] The district court therein had also found that, in any event, the plaintiffs would have been unable to recover because they had consented to the data collection via the Terms of Service agreement with the ISP.[234]

In *In re iPhone Application Litigation*,[235] the court, relying in part on the *Doubleclick* decision, dismissed the plaintiffs' claims under the SCA on four independent grounds, and dismissed the plaintiffs' claims under the Wiretap Act, finding that geolocation and other data being collected did not constitute the "contents" of communications under the statute.[236] The court then dismissed the plaintiffs' claims for invasion of privacy stating, "the information allegedly disclosed to third parties included the unique device identifier number, personal data, and geolocation information from Plaintiffs' iDevices. Even assuming this information was transmitted without Plaintiffs' knowledge and consent, a fact disputed by Defendants, such disclosure [did] not constitute an

---

[231] *Doubleclick*, 154 F. Supp. 2d at 526.

[232] 702 F.3d 1245 (10th Cir. 2012).

[233] *Id.*, aff'g Kirch v. Embarq Mgmt. Co., 10-2047-JAR, 2011 U.S. Dist. LEXIS 92701, at *26 (D. Kan. Aug. 19, 2011).

[234] *Kirch*, 2011 U.S. Dist. LEXIS 92701, at *24–29.

[235] In re iPhone App. Litig., 844 F. Supp. 2d 1040 (N.D. Cal. 2012).

[236] *Id.* at 1056–62.

egregious breach of social norms" and thus did not satisfy the third prong for invasion of privacy under California law.[237]

The *In re Google, Incorporated, Privacy Policy Litigation*[238] plaintiffs sued challenging Google's 2012 privacy policy changes wherein Google began combining user information across all of its services.[239] The court began by noting that, while the plaintiffs' loss of their personally identifiable information was not sufficient to establish injury-in-fact for standing purposes, their allegations of economic and statutory injuries did establish standing.[240] Nevertheless, citing *Kirch*, the court dismissed the plaintiffs' claims under the Wiretap Act because Google had transmitted the plaintiffs' information in the ordinary course of business.[241] Then noting that "[t]he SCA is not a catchall statute designed to protect the privacy of stored Internet

---

[237] *Id.* at 1063. Subsequently, the plaintiffs' class action claims for violations of California's Consumer Legal Remedies Act, CAL. CIV. CODE §§ 1750–1784, and California's Unfair Competition Law, CAL. BUS. & PROF. CODE §§ 17200–17209, were dismissed for lack of standing because the court found that the plaintiffs had failed to show causation, although the court stopped short of stating that Apple's privacy misrepresentations did not constitute an injury in fact. *See generally* In re iPhone Application Litig., No. 11-MD-2250-LHK, 2013 U.S. Dist. LEXIS 169220, at *1, 74 (N.D. Cal. Nov. 25, 2013).

[238] *In re* Google, Inc., Privacy Policy Litig., No. C-12-1382-PSG, 2013 U.S. Dist. LEXIS 171124 (N.D. Cal. Dec. 3, 2013).

[239] *Id.* at *5–6.

[240] *Id.* at *28. Several courts have also been willing to find that standing existed in instances where plaintiffs sought damages for the collection of personal data when those claims were made in conjunction with a data breach on the part of the data-holder. *See, e.g.*, In re Sony Gaming Networks and Customer Data Security Breach Litig., 996 F. Supp. 2d 942, 962 (S.D. Cal. 2014) (finding standing where the plaintiffs had "plausibly alleged a 'credible threat' of impending harm based on the disclosure of their Personal Information following the intrusion").

[241] *Id.* at *30–37. *But see* In re Google, Inc., Gmail Litig., No. 13-MD-2430-LHK, 2013 U.S. Dist. LEXIS 172784, at *1, 58 (N.D. Cal. Sept. 26, 2013) (presenting a case in which plaintiffs' state and federal wiretapping claims against Google, based on Google's practice of scanning emails and using customer data to create user profiles and using that data in a manner unrelated to its providing of its services, *survived* a motion to dismiss in part because the ordinary course of business exception to the Wiretap Act did not apply when Google violated its own privacy policy).

communications,"[242] the court dismissed the plaintiffs' claims under the SCA because: (1) Sec. 2701(c) of the SCA exempts "conduct authorized by the person or entity providing a wire or electronic communications service" from criminal punishment; and (2) the plaintiffs had equivocated in their allegations under Sec. 2702 regarding whether or not Google had actually disclosed the plaintiffs' information to third parties, thus failing to state a claim.[243] The court then dismissed the plaintiffs' claim for misappropriation of likeness and intrusion upon seclusion based on the users' consent, and similarly dismissed the plaintiffs' breach of contract claims.[244]

The case law indicates that users wishing to either enjoin private sector data collection practices or recover monetary damages resulting from those practices have been largely unsuccessful. Tort and contract claims quickly encounter insurmountable consent issues insofar as pleading a cognizable claim is concerned. Claims under the ECPA routinely fail both because of the aforementioned consent issue, and because the ECPA is riddled with exceptions.[245] Lacking comprehensive congressional action, or the emergence of radically creative judicial applications of the ECPA, consumers are left without a remedy. However, though certainly not a systemic solution, the FTC's recent efforts at curbing unfair and deceptive practices in this regard are nevertheless noteworthy.

C. *Federal Trade Commission Involvement*

The FTC recently called upon Congress to "enact[] legislation that would enable consumers to learn of the existence and activities of data brokers and provide consumers with reasonable access to

---

[242] *In re* Google, Inc., Privacy Policy Litig., 2013 U.S. Dist. LEXIS 171124, at *38 (quoting Orin S. Kerr, *A User's Guide to the Stored Communications Act, and a Legislator's Guide to Amending It,* 72 GEO. WASH. L. REV. 1208, 1214 (2004)).

[243] *Id.* at *37–39.

[244] *Id.* at *39–44, 49–51. *See also In re* Zynga Privacy Litig., 750 F.3d 1098, 1109 (9th Cir. 2014) (dismissing the plaintiff's claims under both the SCA and the Wiretap Act).

[245] *See, e.g.*, 18 U.S.C. § 2702(c).

information about them held by these entities."[246] It also recommended providing consumers with the ability to "opt out" of having their information shared.[247] Of particular note, given the previous discussion regarding predictive analytics as applied to the individual, the FTC report also stated:

> [T]o further enhance transparency, the Commission recommends that Congress consider legislation requiring data brokers to clearly disclose to consumers (e.g., on their website) that they not only use the raw data that they obtain from their sources, such as a person's name, address, age, and income range, *but that they also derive certain inferences from the data*.[248]

Despite the absence of effective legislation, FTC enforcement has achieved some degree of success in its efforts to improve transparency and limit some of the more egregious collection practices used by data brokers.[249] In 2009, for example, the FTC settled a complaint lodged against Sears Holdings Management Corporation that charged that Sears failed to disclose the scope of its tracking and collection of consumers' personal information.[250] This stemmed from Sears's invitation to consumers to become members of the "My SHC Community" in which consumers were enticed to participate by Sears' offer of ten dollars to each participant.[251] While participation in the "My SHC Community" program asked consumers to download "research" software that was billed as confidentially tracking online browsing, the FTC alleged that consumers were not informed that the software would also collect information from consumers' online shopping carts, online bank statements, drug prescription records, email histories,

---

[246] FTC REPORT, *supra* note 7, at vii.

[247] *Id.* at viii; *see generally* CUKIER BIG DATA, *supra* note 4, at 156 (describing "opting out" as one of three generally recognized privacy strategies, alongside notice and consent, and anonymization).

[248] FTC REPORT, *supra* note 7, at 52 (emphasis added).

[249] *See* Federal Trade Comm'n Act, 15 U.S.C. § 45(a),(n).

[250] *See* Press Release, FTC, Sears Settles FTC Charges Regarding Tracking Software (Jun. 4, 2009), *available at* http://www.ftc.gov/news-events/press-releases/2009/06/sears-settles-ftc-charges-regarding-tracking-software) [hereinafter *Sears* Press Release].

[251] *Id.*

and the like.[252] The final settlement required Sears to make specific disclosures to consumers as to precisely what data was being collected and the purpose for which it would be used. It also required Sears to delete all data that was collected during the program.[253]

In 2011, Facebook settled claims brought by the FTC on charges that Facebook's privacy practices were "unfair and deceptive."[254] Among the listed violations of Facebook's privacy promises to customers set forth in the FTC complaint were allegations that Facebook: (1) had promised users it would not share personal information with advertisers, although it did; (2) had a "Verified Apps" program in which it represented that Facebook had certified the security of participating apps when it had not; and (3) "claimed that when users deactivated or deleted their accounts, their photos and videos would be inaccessible. But Facebook allowed access to the content, even after users had deactivated or deleted their accounts."[255]

Similarly, in 2012, Google settled with the FTC, agreeing to pay a record $22.5 million penalty to settle charges that it "misrepresented to users of Apple Inc.'s Safari Internet browser that it would not place tracking 'cookies' or serve targeted ads to those users, violating [the] earlier privacy settlement between the company and the FTC."[256] Likewise, in May of 2014, the FTC reached a settlement agreement with Snapchat, a popular app that

---

[252] *Id.*

[253] *Id.*; *see also* G.S. Hans, *Privacy Policies, Terms of Service, and FTC Enforcement: Broadening Unfairness Regulation for a New Era*, 19 MICH. TELECOMM. TECH. L. REV. 163, 173–75 (2012) (discussing the FTC enforcement action against Sears).

[254] Press Release, FTC, Facebook Settles FTC Charges That It Deceived Consumers by Failing to Keep Privacy Promises (Nov. 29, 2011), *available at* http://www.ftc.gov/news-events/press-releases/2011/11/facebook-settles-ftc-charges-it-deceived-consumers-failing-keep.

[255] *Id.*

[256] Press Release, FTC, Google Will Pay $22.5 Million to Settle FTC Charges It Misrepresented Privacy Assurances to Users of Apple's Safari Internet Browser (Aug. 9, 2012), *available at* http://www.ftc.gov/news-events/press-releases/2012/08/google-will-pay-225-million-settle-ftc-charges-it-misrepresented.

"billed itself as a way of sending messages—snaps—which would self-destruct within a set timeframe after being viewed by the recipient, over violations of Snapchat's promises regarding the " 'ephemeral' nature of 'snaps.' "[257] Snapchat's violations received additional attention in November 2014, when a massive hack dubbed the "Snappening" resulted in the release of "a database containing over 100,000 images and videos sent across Snapchat leaked online for the titillation of the masses."[258]

In addition to the successes of the FTC in combating alleged deceptive and unfair practices within the big data industry, the FTC has been among the most vocal advocates for comprehensive Congressional action to increase transparency within the data broker industry.[259] Despite these successes, the overwhelming consensus is that current privacy laws are woefully inadequate to deal with the phenomenon that massive data collection—let alone the practice of predictive analytics—represents.

D. *Executive Involvement and the Consumer Privacy Bill of Rights*

For its part, the White House appears to support the efforts of the FTC, at least publicly. In 2012, the Obama administration released a report that set forth the so-called Consumer Privacy Bill of Rights ("CPBR").[260] Based largely on the Fair Information

---

[257] *See* Press Release, Snapchat Settles FTC Charges That Promises of Disappearing Messages Were False, FED. TRADE COMM'N (May 8, 2014), *available at* http://www.ftc.gov/news-events/press-releases/2014/05/snapchat-settles-ftc-charges-promises-disappearing-messages-were; *see also* Charlie Osborne, *FTC Finalizes Charges Against Snapchat Over User Privacy*, ZDNET (Jan. 2, 2015), http://www.zdnet.com/article/ftc-finalizes-charges-against-snapchat-over-user-privacy/.

[258] Charlie Osborne, *supra* note 257.

[259] *See, e.g.*, FTC REPORT, *supra* note 7, at vii.

[260] *See generally* EXECUTIVE OFFICE OF THE PRESIDENT, CONSUMER DATA PRIVACY IN A NETWORKED WORLD: A FRAMEWORK FOR PROTECTING PRIVACY AND PROMOTING INNOVATION IN THE GLOBAL DIGITAL ECONOMY (Feb. 2012), *available at* http://www.whitehouse.gov/sites/default/files/privacy-final.pdf [hereinafter CONSUMER DATA PRIVACY REPORT].

Practice Principles,[261] this proposed legislative initiative focuses on several key areas, such as increased individual control for consumers in terms of what personal data is collected from them by private companies, and increased transparency so that customers are provided with easily understandable information regarding privacy and security policies.[262] The CPBR also seeks to establish a consumer right to ensure that companies handle their data in a secure manner, while also addressing several other areas, including an increased focus on providing consumers with the ability to access and correct personal data "in a manner that is appropriate to the sensitivity of the data and the risk of adverse consequences to consumers if the data is inaccurate."[263]

Although at the time of its release the report stated that "[s]trengthening consumer data privacy protections in the United States is an important Administration priority[,]"[264] and the subsequent 2014 White House Report on Big Data similarly recommended amending the ECPA,[265] as of this writing there has been little in the way of serious legislative attempts at implementation. However, while it is noteworthy that the U.S. Department of Commerce's National Telecommunications and Information Administration recently requested public "comment on 'big data' developments and how they impact the Consumer Privacy Bill of Rights," it was unclear at the conclusion of the public comment period what, if any, efforts at implementation would likely follow.[266] Unsurprisingly,

---

[261] Commonly known as the "FIPPs," the Fair Information Practice Principles emerged from recommendations made by the U.S. Department of Health, Education and Welfare in its 1973 report entitled *Records, Computers, and the Rights of Citizens*. *See* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 17.

[262] CONSUMER DATA PRIVACY REPORT, *supra* note 260, at 1.

[263] *Id.* It is noteworthy, and somewhat ironic, that one of the key goals of what is intended to be a privacy-enhancing proposal—empowering consumers to correct false information held by data brokers—requires consumers to give data brokers more information about themselves.

[264] *Id.* at 5.

[265] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 60.

[266] Big Data and Consumer Privacy in the Internet Economy, 79 Fed. Reg. 32714 (June 6, 2014), *available at* http://www.ntia.doc.gov/files/ntia/publications/big_data_rfc.pdf; *see also* Alexei Alexis, *NTIA Leads Privacy Bill of Rights*

the data broker industry opposes the implementation of the CPBR.[267]

In the meantime, the data broker industry has largely been subject to self-regulation.[268] In the absence of effective legislation governing Internet privacy policies, the recent White House Report on Big Data described the "self-regulatory" nature of the current regime, stating that "companies agree to a set of principles when engaged in 'behavioral' or multisite advertising where they collect information about user activities over time and across different websites in order to infer user preferences."[269] However, self-regulation is problematic as the primary protection offered to consumers. The industry has already racked up a less-than-laudable track record of deceptive privacy policies, violations of existing privacy policies, and settlements with the FTC for unfair or deceptive business practices. Additionally, the Do Not Track initiative has largely failed because of websites' unwillingness to honor consumer requests. The prudent consumer, or attorney, is one who assumes that self-regulation is ultimately a failing proposition.

The same White House Report also notes the "privacy fatigue" that is commonly experienced as a symptom of having to wade through a seemingly endless barrage of legalese to use a given service.[270] As much of the data collection industry's power stems

---

Review in Light of 'Big Data' Trend, BNA (June 4, 2014), http://www.bna.com/ntia-leads-privacy-n17179891045/.

[267] *See, e.g.*, Jennifer Glasgow, *Does the US Need the Privacy Bill of Rights Enacted Into Law?*, ACXIOM (Aug. 18, 2014), http://www.acxiom.com/us-need-privacy-bill-rights-enacted-law/.

[268] *See, e.g.*, Perry Simpson, *White House Supports Self-Regulation in Data-Driven Marketing*, DIRECT MARKETING NEWS (May 2, 2014), http://www.dmnews.com/white-house-supports-self-regulation-in-data-driven-marketing/article/345388/; *see also* Glasgow, *supra* note 267.

[269] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 41.

[270] *Id.* at 42 (citing Sarah Kidner, *Privacy Fatigue Hits Facebook: Have You Updated Your Settings?,* WHICH? CONVERSATION, (Oct. 18, 2011), http://conversation.which.co.uk/technology/facebook-privacy-settings-privacy-fatigue/; Aleecia McDonald & Lorrie Cranor, *The Cost of Reading Privacy Policies*, 4 INFORMATION SOCIETY: A JOURNAL OF LAW & POLICY FOR THE INFORMATION

from the consent garnered from consumers as a result of their unwillingness to read, or inability to comprehend, such policies, this is further reason for skepticism as to the long-term success of a system of self-regulation. To be fair, however, this result—generalized malaise, rather than mass public outrage—is likely at least partially related to shifting societal norms regarding what is or is not "creepy" in terms of privacy and data collection and dissemination in the digital age.[271]

E. *Additional Considerations: A Symptom of the Disease—Permanent Retention and Creative Discovery Practices*

A byproduct of the universal realization of the enduring value of data is the widespread and prolonged *retention* of that data, which in some cases may represent a veritable treasure trove of information, discoverable or otherwise, that may never be deleted.[272] As companies like Google, Facebook, and countless others have developed policies wherein they save practically everything that users do online,[273] it should be noted that this trend

---

SOCIETY, 545, 545, 564 (2008)); *see also* Ian Ayres & Alan Schwartz, *The No-Reading Problem in Consumer Contract Law*, 66 STAN. L. REV. 545, 546 (2014) (describing the ubiquitous phenomenon of customers' routine failure to read form contracts offered by providers of goods and services, leading to problems of consumer consent and a lack of competitive pressure on business to improve the contractual terms offered to customers).

[271] *See* Omer Tene & Jules Polonetsky, *A Theory of Creepy: Technology, Privacy and Shifting Social Norms*, 16 YALE J.L. & TECH. 59, 72–73 (2014) (describing changing privacy norms and stating that the advent of caller ID in the 1980s was widely regarded as "creepy" and resulted in some attempts at regulation at the state level as critics considered it a privacy violation to see who was calling in to, for instance, an HIV help line, or a drug or alcohol clinic. This concern, however yielded in the face of changing norms, and "[t]oday, many users would not answer the phone if the number were not listed. What was initially considered a privacy violation is now considered a privacy-enhancing technology.").

[272] *See* FTC REPORT, *supra* note 7, at 22.

[273] *See also* CUKIER BIG DATA, *supra* note 4, at 174 (advocating for limited time frames for data retention and reuse, stating that such an approach would "banish[] the specter of 'permanent memory'—the risk that one can never escape one's past because the digital records can always be dredged up"); *see, e.g.*, *Privacy & Terms*, GOOGLE, https://www.google.com/intl/en/policies/

has not gone unnoticed in either criminal or civil litigation. Individual users' search queries have already been the subject of governmental, private, and international discovery.[274] "Government agencies, courts and parties in civil litigation regularly ask technology and communications companies for information about how a person has used the companies' services."[275] Google's most recent transparency report notes that government requests for user information has increased 120% since Google first began publishing such numbers in 2009.[276]

Equally troubling is the persuasive suggestion by some scholars that the current legal landscape makes it unclear the extent to which the SCA prevents a party in civil litigation from obtaining communications maintained by cloud services providers as part of civil discovery. Even communications deleted by users may be recoverable. The SCA may allow discovery of such communications despite the fact that the SCA does not on its face authorize service providers to make such disclosures for the purposes of civil discovery.[277] In *Flagg v. City of Detroit*,[278] the defendant-City argued that the SCA precluded the production in civil litigation of electronic communications that had been previously deleted by the users but had nevertheless been stored by a non-party service

---

privacy/?fg=1 (last modified Dec. 19, 2014) (describing what is collected, and also stating that "after you delete information from our services, we may not immediately delete residual copies from our active servers and may not remove information from our backup systems").

[274] *See* Ron A. Dolin, *Search Query Privacy: The Problem of Anonymization*, 2 HASTINGS SCI. & TECH. L.J. 137, 137 (2010); *see also* ACLU v. Gonzales, 478 F. Supp. 2d 775 (E.D. Pa. 2007); Gonzales v. Google, Inc., 234 F.R.D. 674 (N.D. Cal. 2006); *see generally* Foley, *supra* note 112, at 451–54 (discussing the use of subpoenas to acquire records of one's Internet activities in criminal litigation).

[275] *Google Transparency Report—Requests for User Information: Legal Process*, GOOGLE, https://www.google.com/transparencyreport/userdatarequests/legalprocess/ (last visited June 26, 2014).

[276] *Access to Information*, GOOGLE, https://www.google.com/transparencyreport/ (last visited June 26, 2014).

[277] *See* Kesan et al., *supra* note 31, at 415.

[278] 252 F.R.D. 346 (E.D. Mich. 2008).

provider.[279] After a thorough analysis of the applicable provisions of the SCA, the court concluded that the information held by the third party could be produced via a Rule 34 [280] request for production directed to the defendant-City *itself*, rather than the third party.[281] The court reasoned that the defendant maintained "control" over that information because the issuance of its consent to the third party would allow the third party to disclose the "deleted" communications under the SCA.[282] Similarly, in *Thayer v. Chiczewski*,[283] the court first acknowledged "most courts have concluded that third parties cannot be compelled to disclose electronic communications pursuant to a civil—as opposed to criminal—discovery subpoena . . . ." Nevertheless, the court, applying reasoning similar to that found in *Flagg*, ordered third-party America Online ("AOL") to turn over its records of the plaintiff's previously deleted emails, because the subpoena in question sought "documents that [the plaintiff] would be required to produce if he had not deleted them from his email accounts."[284]

---

[279] *Id.* at 347.

[280] See FED. R. CIV. P. 34(a) which states:

> A party may serve on any other party a request within the scope of Rule 26(b):
>
> (1) to produce and permit the requesting party or its representative to inspect, copy, test, or sample the following items in the responding party's possession, custody, or control:
>
> (A) any designated documents or electronically stored information . . . stored in any medium . . . .

Id.

[281] *See Flagg*, 252 F.R.D. at 353.

[282] *See id.* at 352–53 (citations omitted) (stating that Rule 34 requests for production "may properly extend to items that are in that party's 'control.'" Then, after finding that electronic information may be said to be within a party's control when it is "maintained by a third party of [a] company's behalf[,]" the court found that such information may be obtained in discovery in civil litigation). *But see* Bower v. Mirvat El-Nady Bower, 808 F. Supp. 2d 348, 349 (D. Mass. 2011) (denying the plaintiff's attempts to produce the defendant's emails, lacking the defendant's consent, pursuant to Rule 45 of the Federal Rules of Civil Procedure).

[283] No. 07-C-1290, 2009 U.S. Dist. LEXIS 84176 (N.D. Ill. Sept. 11, 2009).

[284] *Id.* at *15; *see also* Negro v. Superior Court, 230 Cal. App. 4th 879, 883 (Cal. Ct. App. 2014) (finding that the lower court's imputation of consent on the

These cases stand for the proposition that, even if the SCA prevents entities like Facebook, Google, or Yahoo! from opening up their data vaults in response to Rule 45 subpoenas without a user's consent, courts are increasingly willing to order litigants to request copies of electronic records and communications from providers themselves pursuant to Rule 34. This practice sidesteps the consent requirement and, for all intents and purposes, the few remaining privacy protections of the SCA altogether. This is significant because, if courts interpret electronic information stored in the cloud—search query histories, websites visited, emails previously deleted by users, etc.—as perpetually within the custody or control of the consumer due to the consumer's ability to consent to disclosure by third parties, and third parties retain user information forever, then courts can simply order litigants to request this information from service providers pursuant to a Rule 34 subpoena.[285] Such an interpretation potentially places an individual's entire Internet history into play in future civil litigation, which may contain discoverable information itself, be used to lead to discoverable information, or be used as fuel for predictive analytical machinery.

---

part of the petitioner was improper, but nevertheless finding that consent existed where petitioner had been ordered by a Florida court to give his express consent to a third-party service provider—in this case Google—and he had done so, thus, the SCA did not protect petitioner's emails); Jake Vandelist, *Status Update: Adapting the Stored Communications Act to a Modern World*, 98 MINN. L. REV. 1536, 1547–48 (2014) (collecting cases, and discussing the use of discovery requests served on users rather than providers as an end-run around the SCA).

[285] *See* FTC v. Sterling Precious Metals, L.L.C., No. 12-80597-CIV, 2013 U.S. Dist. LEXIS 50976 (S.D. Fla. Apr. 9, 2013); *see also* Doe v. City of San Diego, No. 12-cv-689-MMA, 2013 U.S. Dist. LEXIS 74802, at *12–13 (S.D. Cal. May 28, 2013) (finding that, while the SCA prohibited Verizon from disclosing customer records to a subpoena issued by the City of San Diego, "the City [was] permitted to seek Plaintiff's cell phone records by serving a request for production of documents pursuant to Federal Rule of Civil Procedure 34").

V. AN ATTORNEY'S DUTY OF COMPETENCE—NEW ETHICAL OBLIGATIONS ARISE IN THE WAKE OF RAPID TECHNOLOGICAL DEVELOPMENTS

Just as Intelius's predictive models could accurately predict individual criminality,[286] so too can data brokers use similar processes to make other predictions and deductions about individuals that possess practical utility in conducting opposition research for litigants. Data brokers already sell "people search" products containing personal information to whomever wishes to purchase them, which one could reasonably argue demonstrates a certain industry-wide comfort with selling sensitive personal information at retail. Add to this state of affairs the aforementioned mass storage of online activities due to increases in both storage capacity and the value of data, and a conclusion begins to emerge that technological advances are leading to the rise of an affirmative ethical and professional duty for attorneys to instruct and educate their clients on responsible, preferably anonymous, use of the Internet.

It seems the only thing standing in the way of the widespread commercial availability of publicly accessible profiles detailing everything from a litigant's proclivity for criminality or existing mood or mental disorders to the health of his or her spouse or the financial circumstances in a given household is the will to create it. There is no reason why algorithms could not be tasked with processing existing data to assess whether someone is rich or poor, faithful or lecherous, healthy or sick, trustworthy or false, politically active or apathetic, sober or otherwise, or engaged in any manner of malfeasance or criminality.[287] The problem with relying on a self-regulatory regime is that, under such a regime, the only true check on data brokers' activities is public outrage.

---

[286] *See supra* notes 159–60 and accompanying text.

[287] *See, e.g.*, Jennifer Golbeck, *Smart People Prefer Curly Fries*, SLATE (Oct. 7, 2014, 7:48 AM), http://www.slate.com/articles/technology/future_tense/2014/10/youarewhatyoulike_find_out_what_algorithms_can_tell_about_you_based_on_your.html (discussing the use of algorithmic analyses to make intimate and personal deductions from innocuous bits of social media data); *see also* Robertson, *supra* note 159; Kosinski, *supra* note 165.

However, if the past two decades have been any indication, the public's capacity for outrage and indignation at digital privacy intrusions may be waning.[288] If the public's comfort with intrusive private sector surveillance is increasing, while the principal check on data broker activities is public sentiment, it is reasonable to conclude that the likelihood of "people search" products becoming progressively more detailed and invasive in the years ahead represents an unacceptably high risk to litigants. When one considers that information, once disseminated, remains in the digital world forever, it will be too late to avoid falling victim to such investigative tactics in the future unless individuals stop the information seepage today.

The Internet's relatively recent assumption of its role as a tool central to the practice of law has brought with it increasingly complex ethical questions for attorneys as they attempt to navigate new and uncertain issues in the digital marketplace.[289] For instance, Google's practice of digitally scanning the contents of its users' emails in order to deliver more accurately targeted ads initially led to difficult questions in terms of third-party disclosures and the possible waiver of attorney-client confidentiality.[290] Further, some

---

[288] *See supra* note 270 and accompanying text.

[289] *See, e.g.*, Am. Honda Motor Co., Inc. v. Motorcycle Information Network, Inc., No. 5:04-cv-12-oc-10GRJ, 2008 WL 906739, at *7 (M.D. Fla. Apr. 2, 2008) (noting that, "[i]ndeed, the failure to use computerized legal research may be a basis for a claim of malpractice in some instances.").

[290] *See, e.g.*, Shellie Stephens, *Going Google: Your Practice, The Cloud, and the ABA Commission on Ethics 20/20*, 2011 U. ILL. J.L. TECH. & POL'Y 237, 239 (2011) (describing cloud computing's impact on an attorney's obligations under Rule 1.6); *see also* N.Y. St. Bar Ass'n, Op. 820 (2008), *available at* http://www.nysba.org/CustomTemplates/Content.aspx?id=5222; Google Terms of Service, GOOGLE, https://www.google.com/intl/en/policies/terms/ (last visited June 16, 2014) ("Our automated systems analyze your content (including emails) to provide you personally relevant product features, such as customized search results, tailored advertising, and spam and malware detection. This analysis occurs as the content is sent, received, and when it is stored."); *see generally* Timothy Peterson, *Cloudy With a Chance of Waiver: How Cloud Computing Complicates the Attorney-Client Privilege*, 46 J. MARSHALL L. REV. 383 (2012) (discussing waiver of privilege through use of cloud-based computer services).

scholars have persuasively argued that the phenomenon of the online tracking of attorneys as they conduct online legal research could be said to "produce a limited or general disclosure that constitutes a waiver of both attorney-client and work product and violates the attorney's ethical commitment to confidentiality."[291]

Generally, the prudent attorney should advise against a client's needless dissemination of vast amounts personal information that could potentially be used by another to the client's detriment. Advocating for individuals to protect personal information because it could at some point in the future be used against them is not in and of itself a novel suggestion nor is it a novel interpretation of the attorneys' duty of competence to suggest that this duty places an obligation upon lawyers to advise their clients regarding the potential pitfalls of recklessly circulating intimate details of their personal lives in the digital realm.[292] For instance, a recent New York County Lawyer's Association ethics opinion's topic addressed "[w]hat advice is appropriate to give a client with respect to existing or proposed postings on social media sites."[293] Noting that personal injury defendants, rather than hiring private investigators, have with increasing frequency turned to YouTube, Facebook, and other social media websites in order to research the activities of their opponents in litigation, the opinion stated that an attorney's obligation to competently represent clients under Rule 1.1 of the Rules of Professional Conduct could give rise to an obligation to advise clients in terms of how their position might be adversely affected by their use of social media.[294] Most of the bar associations that have addressed this issue thus far concur.[295] Thus,

---

[291] Klinefelter, *supra* note 25, at 22.

[292] *See, e.g.*, NYCLA Opinion, *supra* note 2; NYSBA Opinion, *supra* note 2; PBA Opinion, *supra* note 2; Penn. Opinion, *supra* note 2; NCB Opinion, *supra* note 2.

[293] NYCLA Opinion, *supra* note 2, at 1.

[294] *Id.* at 3. This consideration is of course tempered by legal duties to refrain from suppressing or concealing evidence, as well as issues of spoliation, under applicable law.

[295] *See, e.g.*, NYSBA Opinion, *supra* note 2; *see also* PBA Opinion, *supra* note 2 (concurring with the conclusions of the New York State Bar Association in stating that attorneys may advise clients regarding the removal of potentially

an attorney's obligation to prevent a client from needlessly exposing personal details of his life to the world, which could potentially be used against him by his adversaries in the course of litigation, is not new. It is an obligation that stems directly from the attorney's duty of competence. For attorneys, however, it is a constantly evolving obligation and fast moving technological advancements significantly complicate the attorney's task.

Few, if any, standards of attorney competence currently exist, however, with regard to advising clients as to how to safely browse the Internet, as well as how to generally conduct personal matters and business affairs online in the era of big data, and in an environment in which every click and keystroke is recorded, personal Internet activity data is being commoditized and sold, and the science of predictive analytics is developing with increasing rapidity.[296] Given the rate at which personal information is being

---

damaging information from social media subject to the obligations to preserve evidence).

[296] *See* WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 37 ("Controversially, predictive analytics can now be applied to analyze a person's individual propensity to criminal activity," and describing how police forces in the United States are "shift[ing] the focus of predictive policing from geographical factors to identity"); Mandi Woodruff, *The Secret Way Companies Are Using Big Data to Score You*, YAHOO FINANCE (Apr. 2, 2014, 10:13 AM), https://finance.yahoo.com/news/the-secret-way-companies-are-using-big-data-to-score-you-135018683.html (describing the commoditization of personal information culled by an individual's internet activity to create credit "scores" not subject to the FCRA); Brad Howarth, *Big Data: How Predictive Analytics is Taking Over the Public Sector*, THE GUARDIAN (Jun. 13, 2014, 7:55 PM), http://www.theguardian.com/technology/2014/jun/13/big-data-how-predictive-analytics-is-taking-over-the-public-sector (describing the Australian government's use of big data analytics to, among other things, predict emergency room admissions on a given day with 93% accuracy; and identify tax havens and businesses failing to meet compliance obligations); Rebecca Merrett, *NSW Police Force Sees Opportunities in Predictive Analytics*, CIO (May, 30, 2014, 1:20 PM), https://www.cio.com.au/article/546402/nsw_police_force_sees_opportunities_predictive_analytics/ (describing the implementation of predictive analytic models by law enforcement to deploy resources proactively); Ger Daly, *Embracing the Police Force of the Future*, CNN (Sept. 19, 2013, 10:55 AM), http://www.cnn.com/2013/09/18/tech/innovation/police-future-technology/index.html?iref=allsearch (describing the use of data mining and predictive analytics to "predict patterns of future criminal

collected, analyzed, and sold as well as the potential value of such personal information in the context of civil litigation, attorneys may be falling short of their ethical obligations if they do not at the very least avail clients of current best practices and options available to them insofar as anonymous Internet browsing and safeguarding their personal data is concerned. The most recent addition to Rule 1.1 is a step in that direction.

Rule 1.1 of the Model Rules of Professional Conduct addresses attorney competency.[297] The rule states "[a] lawyer shall provide competent representation to a client. Competent representation requires the legal knowledge, skill, thoroughness, and preparation reasonably necessary for the representation."[298] The difficulty for attorneys is that what constitutes "preparation reasonably necessary for the representation" of a given client is quickly changing in the wake of swift technological advancements. Comment 8 to Rule 1.1 has recently been added to provide that:

> [In order t]o maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, *including the benefits and risks associated with relevant technology*, engage in continuing study and education and comply with all continuing legal education requirements to which the lawyer is subject.[299]

With this obligation in mind, as well as the foregoing discussion regarding the ubiquity of privatized data collection and the developing field of predictive analytics, the question is this: Does an attorney have an ethical obligation insofar as advising clients regarding their day-to-day Internet usage? If an obligation exists to

---

behavior"); Edith Ramirez, *The Secret Eyes Watching You Shop*, CNN (May 30, 2014, 10:35 AM), http://www.cnn.com/2014/05/30/opinion/ramirez-data-brokers-ftc/index.html?iref=allsearch (describing collection and analysis by data brokers, and stating that "[d]ata brokers scoop up the digital breadcrumbs we leave as we shop in stores and online, and apply 'big data' analytical tools to predict where we're going, what we'll buy, and what we'll do next—sometimes even before we know ourselves what we'll buy next.").

[297] *See generally* ABA MODEL RULES OF PROF'L CONDUCT R. 1.1.

[298] *Id.*

[299] *Id.* at R. 1.1 cmt. 8 (emphasis added); *see also* id. at R 1.0(h) ("'Reasonable' or 'reasonably' when used in relation to conduct by a lawyer denotes the conduct of a reasonably prudent and competent lawyer.").

discuss with clients what Facebook posts a client should think twice about, and an attorney perceives a likelihood (or even a chance) that the totality of a client's accrued daily Internet activity over time could ultimately be used to yield far more information than any single post to social media, should an obligation not also exist to advise the client in terms of how to prevent irresponsible web browsing?

Bar associations are rushing to keep pace with evolving ethical obligations in the face of rapid technological advances, and the tendency is to advise attorneys to err on the side of caution. For instance, in 2011, grounding its analysis in rules 1.6(a)[300] and 1.1[301] of the Model Rules of Professional Conduct, the ABA released an ethics opinion that concluded that unencrypted email communications between attorney and client were likely permissible under Rule 1.6 because there was "a reasonable expectation of privacy from a technological and legal standpoint."[302] Nevertheless, since the current legal protections afforded to, for instance, emails sent from an employee's workplace computer are in flux and many such emails have been held to be admissible in court proceedings despite attorney-client privilege, "a lawyer typically should instruct the employee-client to avoid using a workplace device or system for sensitive or substantive communications, and perhaps for any attorney-client communications[.]"[303] This is "because even seemingly ministerial communications involving matters such as scheduling can have substantive ramifications."[304]

A parallel may be drawn herein: As is the case with regards to the uncertain and fluctuating legal protections pertaining to workplace emails, attorneys are now similarly faced with an unregulated data collection industry to which the applicable laws,

---

[300] *See id.* at R. 1.6(a) (stating that a lawyer must safeguard "information relating to the representation of a client unless the client gives informed consent").

[301] *See id.* at R. 1.1.

[302] ABA Comm. on Ethics & Prof'l Responsibility, Formal Op. 11-459 (2011) (citing ABA Comm. on Ethics & Prof'l Responsibility, Formal Op. 99-413 (1999)).

[303] ABA Formal Op. 11-459.

[304] *Id.*

to the extent there are any, are uncertain and in tremendous flux. The 2011 opinion errs on the side of caution, advising clients to avoid using workplace email accounts because the risk of information contained in such accounts being used to the advantage of an adversary is unacceptably high. The same logic applies with equal force to the personal data that most consumers freely surrender into the digital sphere each day.

Attorneys are hardly failing to recognize the potential value of information that can be accessed through online investigation and analysis, including the utilization of major data brokers. Practitioners have published articles offering advice and insight as to the effective use of search engines, social networking sites—in both civil and criminal proceedings—blogs, online court records, and personal data brokers, such as Intelius, as part of a comprehensive informal discovery strategy.[305] Bar associations have begun offering CLE programs with titles such as "Cybersleuth's Guide to the Internet," in which one of the covered topics was "the advantages (and limitations) of [using] fee-based data broker databases to create dossiers about your subject."[306] It has also been suggested that there is an affirmative obligation for attorneys to inquire into social networking information that may hold potential relevance in a given matter.[307] Already on the cutting edge of providing advanced data-mining technologies to litigators, LexisNexis offers products that promise to aid litigants to "get to the right

---

[305] *See generally* Todd B. Baker, *The Internet and the Law: Informal Discovery on the Internet*, 52 THE ADVOCATE 23 (2010); *see also* Steven C. Bennett, *Symposium: Ethical Limitations on Informal Discovery of Social Media Information*, 36 AM. J. TRIAL ADVOC. 473, 500-02 (2013); Craig Ball, *Cybersleuthing for People Who Still Can't Program Their VCRs*, 20 GPSOLO 40, 45 (2003) ("Literally hundreds of data brokers sell their services online, ranging from law-abiding corporate behemoths like Choice-Point and Experian to fly-by-night outfits on both sides of the law.").

[306] *See Continuing Legal Education: Cybersleuth's Guide to the Internet*, KING CNTY. BAR ASS'N, https://www.kcba.org/cle/pdf/212-Brochure.pdf (last visited Jan. 31, 2015).

[307] Steven C. Bennett, *Ethical Limitations on Informal Discovery of Social Media Information*, 36 AM. J. TRIAL ADVOC. 473, 478 (2013).

decisions sooner with in-depth judge profiles,"[308] "sharpen ongoing case strategy and manage client expectations informed by the comprehensive collection of data on experts, judges and attorneys[,]"[309] and "[u]tilize the largest, most comprehensive collection of jury verdicts and settlements available online . . . to evaluate risk and opportunity, gain insight into potential outcomes and determine an initial course of action."[310] In short, the tools of computer-assisted data mining and analysis are already being put to work in the legal world, as are people search products and other data broker services.

The ABA, for its part, thus far seems to generally sanction the practice of delving into one's digital footprint in the course of litigation. A recently released ethics opinion addressed the ethical issues that arise when attempting to investigate a juror's, or potential juror's "internet presence."[311] The opinion concludes that, subject to an attorney's obligations under Rule 3.5(b),[312] and Rule 8.4(a),[313] it is permissible to "passively" use websites and social media to access publicly available information on jurors, if no direct communication between the attorney and the juror takes place.[314] Thus, the general trend, from an ethical standpoint, seems to be one where research on both jurors and litigants in the digital sphere is expected, accepted, and, sometimes, obligatory.

---

[308] *The LexisNexis Litigation Research Portfolio*, LEXISNEXIS, http://www.lexisnexis.com/en-us/legal-solutions/litigation-portfolio.page (last visited Aug. 4, 2014).

[309] *LexisNexis Litigation Profile Suite*, LEXISNEXIS, http://www.lexisnexis.com/en-us/products/lexisnexis-profile-suite.page (last visited Aug. 4, 2014).

[310] *LexisNexis Verdict & Settlement Analyzer*, LEXISNEXIS, http://www.lexisnexis.com/en-us/products/verdict-and-settlement-analyzer.page (last visited Aug. 4, 2014).

[311] ABA Comm. on Ethics & Prof'l Responsibility, Formal Op. 14-466 (2014).

[312] *See* MODEL RULES OF PROF. CONDUCT R. 3.5(b) ("A lawyer shall not . . . communicate ex parte with [a judge, juror, prospective juror or other official] unless authorized to do so by law or court order.").

[313] MODEL RULES OF PROF. CONDUCT R. 8.4(a) ("It is professional misconduct for a lawyer to violate or attempt to violate the Rules of Professional Conduct, knowingly assist or induce another to do so, or do so through the acts of another.").

[314] ABA Comm. on Ethics & Prof'l Responsibility, Formal Op. 14-466 (2014).

Even if, lacking systemic protections for personal data, one were to simply trust that the intentions and motives of each and every member of the data broker industry are ethical and benevolent in nature, it must be remembered that these companies are still vulnerable to hacking and other scams just like everyone else.[315] For instance, in October 2013, the Justice Department brought criminal charges against a Vietnamese man who took part in the purchase of over 500,000 consumer records from data broker Experian, which he then sold to third parties for the purposes of identity theft.[316] This incident was reminiscent of a similar occurrence at ChoicePoint in 2005, wherein ChoicePoint discovered that it had been routinely selling personal data dossiers to criminal enterprises posing as legitimate businesses. This incident ultimately resulted in an FTC settlement wherein ChoicePoint was fined $10 million in

---

[315] *See*, *e.g.*, Mike Isaac, *Data Breaches in New York Hit Record High in 2013, State Attorney General Says*, N.Y. TIMES (July 15, 2014, 12:01 AM), http://bits.blogs.nytimes.com/2014/07/15/attorney-general-says-new-york-had-more-than-900-data-breaches-in-2013/; Alastair Jamieson & Erin McClaim, *Millions of Target Customers' Credit, Debit Card Accounts May Be Hit by Data Breach*, NBC NEWS (Dec. 19, 2013), http://www.nbcnews.com/business/consumer/millions-target-customers-credit-debit-card-accounts-may-be-hit-f2D11775203 (describing massive data breaches at Target; the 2007 data breach of more than 45 million customers from T.J. Maxx, Marshalls, and others; the 2011 hack on over 100 million Sony Playstation user accounts); Charles Riley, *Data Breach in UPS Stores in 24 States*, CNN (Aug. 21, 2014, 9:39 PM), http://money.cnn.com/2014/08/21/technology/security/ups-store-data-hack/index.html (citing a CNN Money analysis which found that fifty percent of Americans were the victims of data breaches in a recent twelve month period); Alexis Tsotsis, *Employee Data Breach the Worst Part of Sony Hack*, TECHCRUNCH (Dec. 16, 2014), http://techcrunch.com/2014/12/16/hack-sony-twice-shame-on-sony/.

[316] *See* Press Release, Dep't of Justice, Off. of Public Affairs, Vietnamese National Charged in Widespread International Scheme to Steal and Sell Hundreds of Thousands of U.S. Persons' Personally Identifiable Information (Oct. 18, 2013), *available at* http://www.justice.gov/opa/pr/2013/October/13-crm-1116.html; *see also* Gil Aegerter, *Credit Giant Experian Tangled in ID Theft Case*, CNBC NEWS (Oct, 24, 2013), http://www.cnbc.com/id/101143539#.; Letter from Sen. John D. Rockefeller IV, Chairman of Senate Comm. on Commerce, Science and Transp., to Don Robert, CEO of Experian (Oct. 23, 2013), *available at* http://msnbcmedia.msn.com/i/MSNBC/Sections/NEWS/A_U.S.%20news/US-news-PDFs/experian-commerce-letter.pdf.

civil penalties and $5 million in consumer redress.[317] Given the ubiquity of high profile hacks and data breaches, not to mention old-fashioned scams and fraud, the inevitable conclusion is that personal data amassed by third-party brokers is never truly safe, either from the legal or illegal sale of that data.

If it is permissible and advisable to consult with clients regarding their use of social media due to the potential for such uses to adversely affect a client's position in litigation, it must follow that it is equally permissible and advisable to consult with clients regarding their daily Internet activities due to the potential for such activities to adversely affect a client's position in litigation. Consider the following: there is currently a lack of ethical prohibitions on attorneys in terms of conducting research on litigants, witnesses, and jurors. Personal information has become highly profitable in the context of civil litigation, and the data market is essentially unregulated, largely unpredictable, and once a keystroke is struck it is stored forever.[318] Attorneys should view these trends in concert with developments in predictive analytic techniques and technologies and the deductions that are now possible as a result. In so doing, attorneys must recognize that an ethical obligation is arising to both instruct clients as to the ramifications of irresponsible Internet usage, as well as to provide, at the very least, the resources necessary for clients to prevent the dissemination of their information into the digital universe for collection, to the extent possible.

VI. PROTECTING ONE'S DIGITAL FOOTPRINT—THE BASICS OF AVOIDING ONLINE TRACKING

No single article can serve to provide an all-encompassing strategy capable of evading tracking of all activities across all servers, platforms, and technologies. Technology today simply

---

[317] *See* Bob Sullivan, *Database Giant Gives Access to Fake Firms*, NBC NEWS (Feb. 14, 2005), http://www.nbcnews.com/id/6969799/ns/technology_and_science-security/t/database-giant-gives-access-fake-firms/#.U-DDyVPLc6A.

[318] *See* FTC REPORT, *supra* note 7, at 22 (stating that some of the largest data brokers in the United States "store all data indefinitely").

develops and comes to market too quickly. However, if nothing else, this Article is meant to serve predominantly as a wake-up call to attorneys that their ethical obligations demand that they stay abreast of current technological trends in the area of private digital surveillance as well as changes in the law to that effect in order to properly advise clients. This will require a concerted effort on the part of bar associations across the country to provide relevant CLEs that address these issues, as well as an individual effort on the part of attorneys to remain up-to-date on technological trends and developments to avoid their clients being blindsided in litigation. With that caveat, however, there are several technologies and practices whose use has been shown to dilute, diminish, or disrupt the creation of a person's digital footprint. Several of these are discussed below.

A. *Tor*

Initially developed by the United States Navy's Naval Research Laboratory, first and foremost as a means of protecting government communications, Tor (shorthand for "the onion router") is a freely available software and open network that is used primarily as a means to mask an Internet user's identity.[319] At the time of The Tor Project's launching in 2002, the focus had shifted somewhat from using Tor solely to protect government communications to protecting individual users' web activity from the prying eyes of private corporations.[320] Today, the Tor Project is a 501(c)(3) non-profit organization, based in Cambridge, Massachusetts, which receives funding from a variety of sources, including Google, Human Rights Watch, the Department of Defense, and the National Science Foundation.[321]

---

[319] *Tor: Overview*, THE TOR PROJECT, https://www.torproject.org/about/overview.html (last visited June 24, 2014); *see also* Stuart Dredge, *What is Tor? A Beginner's Guide to the Privacy Tool*, THE GUARDIAN (Nov. 5, 2013, 7:47 AM), http://www.theguardian.com/technology/2013/nov/05/tor-beginners-guide-nsa-browser.

[320] Dredge, *supra* note 319.

[321] *Tor People*, THE TOR PROJECT, https://www.torproject.org/about/corepeople.html (last visited June 24, 2014); Brian Fung, *The Feds Pay for 60 Percent of Tor's*

Essentially, the Tor software masks a user's location—and, consequently, that user's identity—by distributing a user's web traffic over several locations across the Internet, funneling one's web activity through multiple relays, thus inhibiting tracking.[322] "Instead of taking a direct route from source to destination, data packets on the Tor network take a random pathway through several relays that cover your tracks so no observer at any single point can tell where the data came from or where it's going."[323] These relays are maintained by series of computers on the Tor network that are selected from Tor's own volunteer-operated network in order to disguise the origin and location of information as it is routed through the Internet.[324] Since Tor disguises a user's IP address, making one's online activity appear to have originated from the Tor network itself, a Tor user is, subject to some exceptions,[325] able to operate on the Internet without being tracked.[326] This freedom

---

Development. Can Users Trust It?, WASH. POST (Sept. 6, 2013), http://www.washingtonpost.com/blogs/the-switch/wp/2013/09/06/the-feds-pays-for-60-percent-of-tors-development-can-users-trust-it/; *see also* Keith D. Watson, *Note: The Tor Network: A Global Inquiry Into the Legal Status of Anonymity Networks*, 11 WASH. U. GLOBAL STUD. L. REV. 715, 718 (2012); *see also* THE TOR PROJECT, INC. AND AFFILIATE CONSOLIDATED FINANCIAL STATEMENTS AND REPORTS REQUIRED FOR AUDITS IN ACCORDANCE WITH GOVERNMENT AUDITING STANDARDS AND OMB CIRCULAR A-133 (Dec. 31, 2013), *available at* https://www.torproject.org/about/findoc/2013-TorProject-FinancialStatements.pdf.

[322] *Tor: Overview*, *supra* note 319.

[323] *Id.*

[324] Warwick Ashford, *Growing Call for Anonymity Online, Says Cambridge Researcher*, COMPUTER WEEKLY (June 20, 2014, 3:04 PM), http://www.computerweekly.com/news/2240223087/Growing-call-for-anonymity-online-says-Cambridge-researcher.

[325] *See, e.g.*, *SSD Project EFF: "Surveillance 'Self-Defence Guide' to 'Survive and Defend' Your Civil Liberties' Online,"* ACE NEWS GROUP (Jan. 23, 2014), http://acenewsservices.com/2014/01/23/ssd-project-eff-surveillance-self-defence-guide-to-survive-and-defend-your-civil-liberties-on-line/ (noting that Tor alone will not defend against malware, and may not be completely successful against extremely "resourceful and determined opponents" who have the means and the wherewithal to monitor one's activities at multiple places simultaneously).

[326] *See* Stephanie K. Pell, *Jonesing for a Privacy Mandate, Getting a Technology Fix—Doctrine to Follow*, 14 N.C. J.L. & TECH. 489, 525–26 (2013) ("Tor is an 'onion routing' technology which hides a user's IP address, making it appear to

applies to both the contents of a message itself as well as a user's metadata.[327] There are a select few ways for law enforcement agencies and other dedicated and sophisticated digital trackers to unmask a user's identity, particularly if that user is less than scrupulous in terms of maintaining a certain degree of what might be termed "anonymity discipline" with regard to the use of certain unmasking programs and applications such as Flash player.[328] Even so, the consensus at present seems to be that, even in the face of governmental surveillance, let alone private sector tracking, Tor successfully protects anonymity for most people most of the time.[329]

---

originate from a Tor server rather than the actual address from which the user is connecting to the Internet"). As with all technologies, however, the efficacy of the security and anonymity offered by Tor must be regularly monitored, as both governments at home and abroad, as well as researchers and scholars, are constantly attempting to penetrate the anonymity network. *See* James Ball, et al., *NSA and GCHQ Target Tor Network That Protects Anonymity of Web Users*, THE GUARDIAN (Oct. 4, 2013, 10:50 AM), http://www.theguardian.com/world/2013/oct/04/nsa-gchq-attack-tor-network-encryption. *See also* Ilya Khrennikov, *Putin Sets $110,000 Bounty for Cracking Tor as Anonymous Internet Usage in Russia Surges*, BLOOMBERG (July 29, 2014, 11:37 AM), http://www.bloomberg.com/news/2014-07-29/putin-sets-110-000-bounty-for-cracking-tor-as-anonymous-internet-usage-in-russia-surges.html. Furthermore, two researchers at Carnegie Mellon University Software Engineering Institute—funded primarily by the U.S. Department of Defense—recently backed out of a talk they were set to give at the Black Hat security conference. The researchers had claimed that they had figured out how to hack Tor to ascertain users' identities. It remains an open question whether the security holes have since been patched. *See* Joseph Menn & Jim Finkle, *Internet Privacy Service Tor Warns Users It was Attacked*, REUTERS (July 30, 2014, 6:52 PM), http://www.reuters.com/article/2014/07/30/us-privacy-software-attack-idUSKBN0FZ1RZ20140730; *see also Tor Security Advisory: "Relay Early" Traffic Confirmation Attack*, THE TOR PROJECT (July 30, 2014), https://blog.torproject.org/blog/tor-security-advisory-relay-early-traffic-confirmation-attack.

[327] Pell, *supra* note 326, at 526–27.

[328] *See, e.g.*, Kevin Poulsen, *The FBI Used the Web's Favorite Hacking Tool to Unmask Tor Users*, WIRED (Dec. 16, 2014, 7:00 AM), http://www.wired.com/2014/12/fbi-metasploit-tor/.

[329] *See, e.g.*, Allistair Charlton, *Snowden Files Reveal NSA Had "Major Problems" Tracking Tor Dark Web Users and Cracking Encryption*, INT'L BUS. TIMES (Dec. 29, 2014, 2:38 PM), http://www.ibtimes.co.uk/snowden-files-reveal-nsa-had-major-

Anonymous browsing has myriad applications. The Tor website, for instance, divides its user base into "family and friends," "businesses," "activists," "media," and "military and law enforcement."[330] Although anonymous browsing obviously draws a criminal element seeking to mask online illegal activities, Tor has also found multitudes of users across the globe in countries like Turkey, Egypt, Russia, and China due to increased censorship and surveillance.[331] Similarly, Tor recently rose in use among Iraqis in the wake of the Nouri al-Maliki administration's order to ISPs to block social media and news website access within the country.[332] Domestically, in addition to being employed as a tool to prevent either private or public sector tracking, it has also been used by victims of cyberstalking as a means to evade being tracked.[333]

As a practical matter, technical wizardry is not required to achieve anonymous browsing. While a user may not be invulnerable to certain sophisticated attacks if the full might of a governmental agency has been devoted to tracking an individual, anonymous browsing software should nevertheless be sufficient to prevent the perpetual accumulation of personal data described herein in most instances. Access to the Tor network, for instance, is most easily achieved through the downloading and use of the Tor Browser Bundle, available at the Tor Project's website.[334] The

---

problems-tracking-tor-dark-web-users-cracking-encryption-1481225; *see also supra* note 326 and accompanying text.

[330] *Tor*, THE TOR PROJECT, http://www.torproject.com (last visited July 16, 2014).

[331] *See, e.g.*, Jeff Stone, *Iraqis Download Tor Anonymity Software to Subvert Failed Internet Blockade, Browse Social Media*, INT'L BUS. TIMES (July 16, 2014, 4:35 PM), http://www.ibtimes.com/iraqis-download-tor-anonymity-software-subvert-failed-internet-blockade-browse-social-1613076.

[332] *Id.*

[333] Meghan Neal, *Tor Is Being Used as a Safe Haven for Victims of Cyberstalking*, MOTHERBOARD (May 9, 2014, 4:00 AM), http://www.motherboard.vice.com/read/tor-is-being-used-as-a-safe-haven-for-victims-of-cyberstalking.

[334] *See* Timothy B. Lee, *Five Ways to Stop the NSA From Spying on You*, WASH. POST (June 10, 2013), http://www.washingtonpost.com/blogs/wonkblog/wp/2013/06/10/five-ways-to-stop-the-nsa-from-spying-on-you/ (describing Tor as an essential component of any anti-tracking strategy, as well as describing the ease of installing and using the Tor Browser Bundle); *What is the Tor Browser?*,

browser itself, a modified version of the Firefox browser, is no more difficult to use than other browsers, such as Chrome, Explorer, or Safari, to which most Internet users have become accustomed.[335]

Another method, recently developed by a company called Pogoplug, involves the use of a product called Safeplug, which allows users to directly access the benefits of Tor by plugging the Safeplug into their home router.[336] Safeplug acts as a proxy server, and allows users to use their preferred browsers and still take advantage of the anonymity benefits of the Tor network.[337] There is also an application available for mobile phone browsing for Android operating systems called Orbot that permits users to browse anonymously from their mobile phones.[338]

The downside, however, is that currently the Tor network is likely slower than the browsing speed to which most American consumers have become accustomed.[339] This is a result of the repeated relaying of information through multiple points around the Internet before arriving at its destination.[340] The upside is that the more users that get on the Tor network and volunteer to relay

---

THE TOR PROJECT, https://www.torproject.org/projects/torbrowser.html (last visited Jun. 25, 2014).

[335] *What is the Tor Browser?*, *supra* note 334; *see also* Jason Kennedy, *How to Use Tor, and Is It Actually Safe and Anonymous*, EXTREME TECH (Oct. 26, 2011, 11:29 AM), http://www.extremetech.com/computing/101633-how-to-use-tor-and-is-it-actually-safe-and-anonymous (describing the ease of using Tor).

[336] Olivia Solon, *Safeplug Makes It Super-Easy to Harness Tor's Anonymity at Home*, WIRED (Nov. 22, 2013), http://www.wired.co.uk/news/archive/2013-11/22/safeplug-tor.

[337] *Id.*; s*ee also* Lucas Mearian, *Tiny Anonabox to Offer Online Anonymity Through Tor*, COMPUTERWORLD (Oct. 13, 2014), http://www.computerworld.com/article/2825065/tiny-anonabox-to-offer-online-anonymity-through-tor.html (describing Anonabox, a similar device that further encrypts a user's Internet traffic).

[338] *See Tor on Android*, THE TOR PROJECT, https://www.torproject.org/docs/android.html.en (last visited Jun. 25, 2014) (describing Orbot); *see also* DRAWBRIDGE *supra* note 56 and accompanying text.

[339] *See Tor FAQ: Why Is Tor So Slow?*, THE TOR PROJECT, https://www.torproject.org/docs/faq (last visited Jun. 25, 2014).

[340] *Id.*

traffic for others, the faster the network becomes.[341] Perhaps not surprisingly, in the year following Edward Snowden's leaks regarding the National Security Agency's online spying programs, reports have indicated that the Tor software has been downloaded approximately 120 million times, and this increase in users could help limit the speed issues currently facing the Tor network.[342]

B. *Non-Tracking Search Engines*

In a largely unregulated industry in which personal data collection, sale, and analysis are profitable, the best protection is to prevent information—any information—about one's self from being collected in the first place. Regular Tor usage is a big step in that direction. Another such step is use of a search engine whose own policies do not permit the logging and recording of search queries and user information.

Most major search engines constantly collect and store user information, including search queries, IP addresses, device information, and the like during usage.[343] However, in recent years privacy advocates have suggested the usage of so-called "non-tracking" search engines as part of an overall privacy strategy.[344] For instance,

---

[341] *Id.*

[342] Patrick Howell O'Neill, *Tor Internet Privacy Tool Sees Downloads Jump to 120 Million*, DAILY DOT (June 2, 2014), http://www.dailydot.com/technology/tor-downloads-120-million-snowden-nsa/; *see also What Is Tor?*, ELEC. FRONTIER FOUND., https://www.eff.org/torchallenge/what-is-tor.html (last visited Aug. 19, 2014) (stating that "[t]he more Tor relays we have running, the faster, more robust, and more secure the Tor network will be").

[343] *See, e.g.*, *Privacy & Terms: Privacy Policy*, GOOGLE, https://www.google.com/intl/en/policies/privacy/?fg=1 (last modified Dec. 19, 2014) ("When you use our services or view content provided by Google, we may automatically collect and store certain information in server logs. This includes" search queries, telephone log information, IP address, device information, and "cookies that may uniquely identify your browser or your Google Account."); *see also Yahoo! Privacy Center*, YAHOO, https://info.yahoo.com/privacy/us/yahoo/ (last updated Sept. 25, 2014) (stating that "Yahoo automatically receives and records information from your computer and browser, including your IP address Yahoo cookie information, software and hardware attributes, and the page you request").

[344] *See, e.g.*, Kate Murphy, *How to Muddy Your Tracks on the Internet*, N.Y. TIMES (May 2, 2012), http://www.nytimes.com/2012/05/03/technology/personaltech/

non-tracking search engine DuckDuckGo.com explicitly states that it "does not collect or share personal information. That is our privacy policy in a nutshell."[345] By conducting all searches using a non-tracking search engine, users can greatly reduce the amount of information available for collection, aggregation, sale to data brokers or other third parties, subpoena, or potential loss due to hacking attacks, security holes or technical incompetence. In the wake of increased public awareness of both corporate and governmental tracking of online activities, DuckDuckGo has experienced continuous and steady growth since its inception, and as of January 2014 was averaging upwards of four million queries per day.[346]

C. *Do Not Track & Private Browser Settings*

The DNT concept began gaining traction in late 2010 when the FTC issued recommendations for the creation and implementation of a mechanism somewhat akin to a "do not call" list for the Internet.[347] Initially, DNT was conceived as a means to empower users to control the degree to which first- and third-party websites may monitor their online activity through the use of easy-to-use browser settings that, when enabled, were capable of either blocking third-party cookies by default or sending a signal to websites that the user prefers not to be tracked.[348] For instance, when a user activates the DNT feature in Firefox, Firefox then

---

how-to-muddy-your-tracks-on-the-internet.html?_r=0; *see also 5 Alternative Search Engines that Respect Your Privacy*, HOW TO GEEK, http://www.howtogeek.com/113513/5-alternative-search-engines-that-respect-your-privacy/ (last visited Aug. 19, 2014).

[345] *Privacy Policy*, DUCKDUCKGO.COM, https://duckduckgo.com/privacy#2s (last visited July 16, 2014); *see also ixquick.com Privacy Policy*, IXQUICK, https://www.ixquick.com/eng/privacy-policy.html? (last visited Jan. 10, 2015).

[346] James Vincent, *DuckDuckGo Hits 1bn Annual Searches: Non-Tracking Search Engine Boosted by Privacy Fears*, THE INDEPENDENT (Jan. 10, 2014), http://www.independent.co.uk/life-style/gadgets-and-tech/duckduckgo-hits-1bn-annual-searches-nontracking-search-engine-boosted-by-privacy-fears-9051081.html.

[347] *See* 2010 FTC REPORT, *supra* note 39, at 10–11.

[348] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 43.

conveys this message to every website visited, including advertisers.[349]

As previously touched upon above, however, this initiative has met with limited success primarily due to an unwillingness by many to abide by users' Do Not Track requests or companies' inclusions of provisions within their privacy policies that still permit partial tracking in spite of such requests.[350] Mozilla, for example, candidly notes that whether websites honor these requests or not is voluntary.[351] Moreover, the Digital Advertising Alliance ("DAA"), a "self-regulatory body that develops industry best practices and effective solutions for consumer choice in online behavioral advertising,"[352] recently withdrew its support of the DNT initiative, leaving this practice's continued utility further in question.[353]

While there has been some legislative push to make respecting a consumer's DNT request a mandatory requirement, this effort has received little traction in Congress.[354] Absent widespread respect for consumers' DNT requests, consumers may choose to utilize any one of several widely available extensions such as Ghostery, AdBlock, and Disconnect, which permit users to exercise some degree of control over which particular entity is tracking them on a given website.[355] Although certainly not as

---

[349] *Mozilla Support—How Do I Turn on the Do Not Track Feature?*, MOZILLA, https://support.mozilla.org/en-US/kb/how-do-i-turn-do-not-track-feature (last visited Aug. 25, 2014).

[350] WHITE HOUSE BIG DATA REPORT, *supra* note 4, at 43.

[351] *Mozilla Support*, *supra* note 349.

[352] *Digital Advertising Alliance Announces First 100 Companies Participating in Self-Regulatory Program for Online Behavioral Advertising*, AAAA (June 7, 2011), http://www.aaaa.org/news/press/Pages/060711_alliance_first100.aspx.

[353] *See* Katy Bachman, *Digital Advertising Alliance Exits Do Not Track Group, Development Could Renew Calls for Privacy Laws*, AD WEEK (Sept. 17, 2013), http://www.adweek.com/news/technology/digital-advertising-alliance-exists-do-not-track-group-152475.

[354] *See*, *e.g.*, Do-Not-Track Online Act of 2011, S. 913, 112th Cong. (2011).

[355] *See generally* Rick Broida, *Six Browser Plug-Ins that Protect Your Privacy*, COMPUTERWORLD (Oct. 17, 2014, 3:30 AM), http://www.computerworld.com/article/2692560/six-browser-plug-ins-that-protect-your-privacy.html.

comprehensive a solution as, for instance, regular Tor usage, these add-ons can be added to a user's web browser to limit ad tracking and block online ads. Ghostery, for instance, states that it is functionally "different than opting-out or blocking cookies because those strategies still allow the browser to communicate with the web service . . . . When blocking is enabled, Ghostery never allows the communication in the first place."[356]

D. *Non-Scanning Email Services*

As some are now aware, it is not uncommon for some of the largest email providers to process the contents of email communications in order to more effectively tailor advertisements directed at specific users.[357] For instance, Google's terms of service states, "Our automated systems analyze your content (including emails) to provide you personally relevant product features, such as customized search results, tailored advertising, and spam and malware detection. This analysis occurs as the content is sent, received, and when it is stored."[358] This practice has already been the subject of ongoing litigation, as well as no small amount of ethical head-scratching by attorneys, although the general consensus now seems to be that these services are acceptable, at least from a confidentiality standpoint, provided that no human beings are actually reviewing emails.[359] However, now that Google has recently opened the door to scanning users' email accounts for evidence of criminal activity and turning these findings over to law enforcement, an additional, more direct incentive exists to close

---

[356] *Frequently Asked Questions*, GHOSTERY, https://www.ghostery.com/en/faq (last visited Jan. 10, 2014).

[357] *See, e.g.*, In re Google, Inc., Gmail Litig., No. 13-MD-2430-LHK, 2014 U.S. Dist. LEXIS 36957 (N.D. Cal. Mar. 18, 2014) (quoting Google privacy policy and discussing email scanning).

[358] *Google Terms of Service*, GOOGLE, https://www.google.com/intl/en/policies/terms/ (last modified Apr. 14, 2014).

[359] *See supra* note 290 and accompanying text; *see generally* Kevin Raudebaugh, *Trusting the Machines: New York State Bar Ethics Opinion Allows Attorneys to Use Gmail*, 6 WASH. J.L. TECH. & ARTS 83 (2010).

down any email accounts with companies whose privacy policies permit the regular scanning of emails.[360]

Erring on the side of caution, presuming both that it cannot be said with any certainty to what uses one's stored data will be put in the future, and that, as a result, the ultimate goal is to limit the total amount of one's personal data being collected and stored by the data collection industry, it is advisable to forgo using email services which employ scanning protocols as a matter of course. In place of such services, lesser-known free services such as HushMail,[361] RiseUp,[362] and Zoho[363] have been promoted by some privacy advocates.[364] RiseUp's privacy policy, for instance, states, "Our commitment is to keep as little data on you as we can. Unlike corporate providers, we do not log internet addresses of anyone using riseup.net services, including email."[365] Another option is the registration of a unique domain with an associated email address through services such as Hover or BlueHost.[366] A soon to be available addition to the growing list of privacy oriented email providers is Dark Mail, offered by the Dark Mail Technical Alliance.[367] Although billed primarily as a means to evade government snooping in light of recent disclosures regarding

---

[360] *See*, *e.g.*, Hayley Tsukayama, *How Closely Is Google Really Reading Your Email?*, WASH. POST (Aug. 4, 2014), http://www.washingtonpost.com/blogs/the-switch/wp/2014/08/04/how-closely-is-google-really-reading-your-e-mail/.

[361] *See* HUSHMAIL, http://www.hushmail.com (last visited Aug. 25, 2014) (describing itself as "secure email with built-in encryption, [and] no advertising").

[362] *See* RISE UP, https://help.riseup.net (last visited Aug. 25, 2014).

[363] *See* ZOHO, http://www.zoho.com/mail (last visited Aug. 25, 2014) ("We do not display ads, even in our free plans. Your email exchanges are never scanned for keywords.").

[364] *See* Murphy, *supra* note 344.

[365] *Riseup Email*, RISE UP, https://help.riseup.net/en/email#what-is-special-about-riseupnet-email (last visited July 17, 2014).

[366] Murphy, *supra* note 344; *see also* HOVER, http://www.hover.com/email (last visited Aug. 25, 2014); BLUEHOST, http://www.bluehost.com (last visited Aug. 25, 2014).

[367] *See generally* DARK MAIL, http://www.darkmail.info (last visited Aug. 25, 2014); *see also* Joseph Pagliery, *Dark Mail: Email that Hides from the NSA*, CNN (July 22, 2014, 7:31 PM), http://money.cnn.com/2014/07/22/technology/security/dark-mail/.

ongoing domestic NSA surveillance, Dark Mail offers standard encryption to email content, while also taking the unusual step of encrypting an email's metadata.[368]

E. *Smartphones*

Approximately ninety percent of American adults own a cellular phone as of January 2014.[369] Of those, two-thirds use their phones to access the Internet, and the total number has doubled since 2009.[370] January 2014 also marked the first time that mobile devices have accounted for a majority of the total Internet usage in the United States.[371] Given this consumer climate, it should not be surprising that companies have formed alliances that are dedicated to aiding private sector entities in the linking of consumers' computers to their mobile devices to better facilitate data collection and targeted advertising.[372] This phenomenon, combined with recent increases in some consumers' privacy sensitivity due to revelations regarding both government domestic spying programs and private sector data collection have given rise to new, market-driven technological innovations in cellular technology. For example, in 2014, Apple and Google developed iOS and Android operating system versions, respectively, with encryption that does not permit the companies to unlock the smartphones at the behest of law enforcement, even upon the receipt of a court

[368] *Dark Mail: Email that Hides from the NSA*, *supra* note 367; s*ee also* Lee Hutchinson, *Lavabit Founder Wants to Make "Dark" E-Mail Secure by Default*, ARS TECHNICA (Jan. 6, 2015, 8:00 PM), http://arstechnica.com/security/2015/01/lavabit-founder-wants-to-make-dark-e-mail-secure-by-default/.

[369] *Mobile Technology Fact Sheet*, PEW RES. INTERNET PROJECT, http:/www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/ (last visited Aug. 7, 2014).

[370] *Cell Internet Use 2013*, PEW RES. INTERNET PROJECT, http://www.pewinternet.org/2013/09/16/cell-internet-use-2013/ (last visited Aug. 7, 2014).

[371] *See* James O'Toole, *Mobile Apps Overtake PC Internet Usage in U.S.*, CNN (Feb. 28, 2014, 11:00 AM), http://money.cnn.com/2014/02/28/technology/mobile/mobile-apps-internet/.

[372] *See supra* note 56.

order to do so.[373] However, in the realm of cellular privacy, Blackphone reigns supreme, at least for the time being.

Self-described as "the product of the best privacy minds in the industry," the Blackphone utilizes PrivatOS, an Android based operating system combined with a number of unique security measures designed to prevent data collection and government snooping.[374] Because the device employs no Google services due to Google's failure to endorse PrivatOS, Blackphone will certainly feel somewhat foreign initially. However, early reviews have been generally positive due to unique privacy features, which, for instance, permit users to select specific permissions for downloaded apps and keep such permissions turned off by default.[375] The phone also has remote wiping functions and secure search, browsing, voice, video, and text functions.[376] Although it currently comes with a hefty price tag north of six hundred dollars, and thus may only be attractive in the immediate future to those in sensitive corporate or government positions or those who are uniquely privacy-oriented, the general consensus thus far seems to be that Blackphone lives up to the hype. As one tech writer recently put it, "If data has value, so, apparently, does protecting it."[377]

---

[373] *See, e.g.*, Craig Timberg et al., *FBI Blasts Apple, Google for Locking Police Out of Phones*, WASH. POST, http://www.washingtonpost.com/business/technology/2014/09/25/68c4e08e-4344-11e4-9a15-137aa0153527_story.html.

[374] BLACK PHONE, https://www.blackphone.ch (last visited Aug. 7, 2014).

[375] *See* Molly Wood, *A Smartphone for Consumers Who Want Privacy*, N.Y. TIMES (July 16, 2014), http://www.nytimes.com/2014/07/17/technology/personaltech/review-blackphone-trades-some-convenience-for-security.html?_r=0; Nate Lanxon, *Don't Call It 'NSA-Proof': Blackphone is Here (Hands-on)*, WIRED (Feb. 24, 2014), http://www.wired.co.uk/news/archive/2014-02/24/blackphone-hands-on.

[376] BLACK PHONE, https://www.blackphone.ch (last visited Aug. 7, 2014); *see also* Natasha Lomas, *Blackphone Confirms Privacy-Focused App Store and Device Sandboxes Incoming*, TECHCRUNCH (Dec. 9, 2014), http://techcrunch.com/2014/12/09/blackphone-confirms-privacy-focused-app-store-and-device-sandboxes-incoming/.

[377] Hugh Langley, *Hands On: BlackPhone Review*, TECH RADAR (Feb. 26, 2014), http://www.techradar.com/us/reviews/phones/mobile-phones/blackphone-1228305/review.

VII. CONCLUSION

In the late 1990s and early 2000s, prior to Facebook's meteoric rise and ensuing cultural ubiquity, had the legal profession been able to foresee that every tagged drunken spring break photo, 2:00 AM status update, and furious wall post would one day be vulnerable to potential exposure in the cold, unforgiving light of civil and criminal litigation, attorneys would have been well-advised to discuss the ramifications of such actions, statements, and disclosures with their clients. This Article posits that another similar phenomenon is looming in the form of data collection, aggregation, analysis, and sale,[378] and that it will be the prudent attorney who competently advises his clients to stay ahead of the curve.

This Article is meant to take no position on the obvious Fourth Amendment implications of felon-identifying programs,[379] nor is it meant to be a thorough analytical critique of the shortcomings of the Electronic Communications Privacy Act, or the notice and consent model. Nor is it meant to be an indictment of the data broker industry generally. No one should be surprised when an industry engages in practices that are profitable, innovative, and legal. This Article does, however, suggest that, given the foregoing, the most prudent course of action is one that is farsighted and includes the development of comprehensive strategies designed to maximize privacy and limit the amount of clients' personal data that flows out into the world. This is due to two simple realities: (1) the more one person knows about another, the easier it is for the information holder to manipulate that person and demand their obedience, and (2) everyone's information is for sale.

---

[378] *See, e.g.*, Martha Neil, *Selfies Posted Online Are Being Mass-Scanned for Marketing Insights*, ABA JOURNAL (Oct. 9, 2014), http://www.abajournal.com/news/article/selfies_posted_online_are_being_mass_scanned_for_marketing_insights.

[379] *See, e.g.*, Robertson, *supra* notes 159 and accompanying text.