

RETHINKING PRIVACY REGULATION

STEVEN M. BELLOVIN[†]

ABSTRACT

Today, all privacy regulations around the world are based on the 50-year-old paradigm of notice and consent. It no longer works. The systems we deal with—web pages with their multiple levels of advertising, the Internet of Things, and more—are too complex; consumers have no idea what sites they are contacting nor what their privacy policies are. Privacy harms are not well-defined, especially under U.S. law. Furthermore, privacy policies are ambiguous and confusing. Use controls—the ability for users to control how their data is used, rather than who can collect it—are more promising but pose their own challenges. We recommend research on a new privacy paradigm and give suggestions on interim changes to today’s privacy regulations until there is something new.

[†] Steven M. Bellovin is the Percy K. and Vida L.W. Hudson professor of computer science and affiliate law faculty at Columbia University. Parts of this article were submitted to an NTIA comment process. Thank you to Wendy Grossman, who aided in research from London. I would also like to thank Joseph Lorenzo Hall, Susan Landau, David Vladeck, and the late Joel Reidenberg for their many helpful comments on an earlier version of this article. The opinions expressed here, however, are mine.

TABLE OF CONTENTS

I.	THE PROBLEM WITH NOTICE AND CONSENT.....	3
II.	DATA SECURITY AND HARM	14
III.	ANALYSIS.....	16
IV.	SOME WAYS FORWARD.....	18
	A. Use Restrictions.....	19
	B. Differential Privacy	22
V.	RECOMMENDATIONS.....	23

I. THE PROBLEM WITH NOTICE AND CONSENT

Although privacy is an ancient concern—Jewish writings on the subject go back at least 1500 years¹—modern approaches to privacy date to the 1960s. The best-known work from that era is Professor Alan Westin’s classic 1967 book,² which largely reflects the work of the Committee on Science and Law of the Association of the Bar of the City of New York.³ He noted the importance of consent:⁴

A central aspect of privacy is that individuals and organizations can determine for themselves which matters they want to keep private and which they are willing—or need—to reveal.

He also noted that it was crucial to realize that consent is limited to a particular situation, and it should not be seen as blanket permission to disseminate information:⁵

Finally, it should be recognized that consent to reveal information to a particular person or agency, for a particular purpose, is not consent for that information to be circulate to all or used for other purposes. The individual may consent to tell things to his teacher or professor that ought not be circulated as part of student records without the student’s consent. Information given to life-insurance companies, credit agencies, survey researcher, or government regulatory and welfare agencies ought not to be shared, in ways that identify the particular individual, without notice of the additional use and consent to it. Unless this principle of consent is well understood and accepted as the controlling principle for information flow in a data-stream society, we will be in for serious problems of privacy in the future.

Other scholars opined as well. Donald Michael worried about the effects of having too much private data available:⁶

Private information about a person may exist which is ethically or legally restricted to those who have a legitimate right to it. Such information, about a great portion of our population, exists in business, medical, government, and political files, and in the form of psychological tests, private and government job application histories, federal and state income tax records, draft records, security and loyalty investigations, school records, bank records, credit histories, criminal records, and diaries. Each day more of these records are translated from paper to punchcards and magnetic tapes. In this way they are made more compact, accessible, sometimes more private, and, very importantly, more centralized, integrated, and articulated. The results are more complete records on each individual and a potential for more complete cross-correlations. The would-be invader who knows about these centralized or clustered inventories need not search for sources, and therefore he may be much more inclined to examine the records than if a major search for the sources of information were necessary.

He also expressed concern about people surrendering their data too easily:⁷

[W]e can expect a great deal of information about the social, personal, and economic

¹ *Bava Batra 60a The William Davidson Talmud (Koren – Steinsaltz)*, SEFARIA, https://www.sefaria.org/Bava_Batra.60a.5?lang=bi&with=all&lang2=en [https://perma.cc/T3ED-73LV].

² ALAN F. WESTIN, *PRIVACY AND FREEDOM* (1967).

³ *Id.* at ix.

⁴ *Id.* at 373.

⁵ *Id.* at 375.

⁶ Donald N. Michael, *Speculations on the Relation of the Computer to Individual Freedom and the Right to Privacy*, 33 *GEO. WASH. L. REV.* 270, 274 (1964).

⁷ *Id.* at 275.

characteristics of individuals to be supplied voluntarily—often eagerly—in order that, wherever they are, they may have access to the benefits of the economy and the government.

Today, of course, a lot of information is collected by the private sector, not just the government, and it is sold and resold by data brokers.⁸

Prof. Arthur R. Miller wrote extensively on the legal aspects of privacy,⁹ and testified on it before a Senate subcommittee.¹⁰ He spoke explicitly of the need for openness and correctness:¹¹

To insure [sic] the accuracy of the Center's files, an individual should have access to any stored information concerning him and an opportunity to challenge its accuracy. Perhaps a print-out of a person's record can be sent to him once a year. This suggestion obviously is vulnerable to a number of criticisms. It is expensive, some federal agencies will argue that the value of certain information will be lost if it is disclosed that the government has it, and the suggestion *might* produce a flow of squabbles, many of them petty, with the Data Center, that would entail costly and debilitating administrative proceedings. Nonetheless, the right of citizen to be protected against governmental dissemination of misinformation is so important, some price must be paid preserve it. The monetary cost of informing the public could be reduced by forwarding the print-out with one of the numerous governmental communications that are sent individuals every year. Alternatively, citizens could be given access to their own files on request, perhaps through a network of remote terminals in government buildings. Legitimate governmental secrecy could be preserved and disputes over file content could be reduced if the information in the Center, and access to it, were arranged hierarchically according to content and an individual received only that part of the file that is accessible to anyone outside the agency that collected it.

However, he warned about relying too much on consent as a way to protect privacy:¹²

A final note on access and dissemination. Excessive reliance should not be placed on what too often is viewed as a universal solvent—the concept of consent. How much attention is the average citizen going to pay to a governmental form requesting consent to record or transmit information? It is extremely unlikely that the full ramifications of the consent will be spelled out in the form; if they were, the document probably would be so complex that the average citizen would find it incomprehensible. Moreover, in many cases the consent will be coerced, not necessarily by threatening a heavy fine or imprisonment, but more subtly by requiring consent as a prerequisite to application for a federal job, contract, or subsidy.

This warning was prescient.

A few years later, the groundbreaking works of these and other pioneers were the foundations for a very important publication, a report by an advisory committee to the then-extant Department of Health, Education, and Welfare.¹³ This report sets forth what have become known

⁸ See, e.g., FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY (2014) [hereinafter DATA BROKER REPORT], <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf> [<https://perma.cc/8GV7-RBYX>].

⁹ See, e.g., Arthur R. Miller, *Personal Privacy in the Computer Age: The Challenge of a New Technology in an Information-Oriented Society*, 67 MICH. L. REV. 1089 (1969).

¹⁰ *Computer Privacy: Hearing on S. Res. 25 Before the Subcomm. on Admin. Prac. and Proc. of the Comm. on the Judiciary*, 90th Cong. 66 (1967) (statement of Arthur R. Miller, Professor, Univ. of Mich. Sch. of L.).

¹¹ *Id.* at 77.

¹² *Id.* at 78.

¹³ See generally U.S. DEP'T OF EDUC. & WELFARE, SEC'Y'S ADVISORY COMM. ON AUTOMATED PERSONAL DATA SYS., RECORDS, COMPUTERS, AND THE RIGHTS OF CITIZENS (1973) [hereinafter HEW REPORT].

as the Fair Information Practice Principles (“FIPPs”):¹⁴

- There must be no personal data record-keeping systems whose very existence is secret.
- There must be a way for an individual to find out what information about him is in a record and how it is used.
- There must be a way for an individual to prevent information about him that was obtained for one purpose from being used or made available for other purposes without his consent.
- There must be a way for an individual to correct or amend a record of identifiable information about him.
- Any organization creating, maintaining, using, or disseminating records of identifiable personal data must assure the reliability of the data for their intended use and must take precautions to prevent misuse of the data.

These five principles have formed the basis for modern privacy regulation, including the U.S. Privacy Act of 1974¹⁵ and the European Union General Data Protection Regulation.¹⁶

Note the third bullet: consent for each use. In technical fora, consent, in fact, is part of the very definition of privacy:¹⁷

1. The right of an entity (normally a person), acting in its own behalf, to determine the degree to which it will interact with its environment, including the degree to which the entity is willing to share its personal information with others. (See: HIPAA, personal information, Privacy Act of 1974. Compare: anonymity, data confidentiality.)
2. “The right of individuals to control or influence what information related to them may be collected and stored and by whom and to whom that information may be disclosed.”

In fact, these definitions draw on Professor Westin’s: “Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.”¹⁸ The problem, though, is that in today’s world, consent no longer works. There is far too much data, the collectors are opaque to the ordinary citizens, and new technologies render the very concept of “purpose” meaningless. Our “data shadow”¹⁹ is too large; it is not possible to control it.

The first problem is the collection of data. All but unknown to most people, data brokers²⁰ collect vast amounts of data from individuals.²¹ Their activities are not secret; indeed, Federal Trade Commission (“FTC”) members have warned about their data collection practices and

¹⁴ *Id.* at xx–xxi.

¹⁵ Privacy Act of 1974, 5 U.S.C. § 552a.

¹⁶ Council Regulation 2016/679 of May 4, 2016, On the Protection of Natural Persons With Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L119) 1.

¹⁷ See R. SHIREY, INTERNET SECURITY GLOSSARY, VERSION 2, RFC 4949 232 (THE IETF TRUST, 2007) (citations omitted).

¹⁸ WESTIN, *supra* note 2, at 7.

¹⁹ Steven M. Bellovin, *Who Coined the Phrase “Data Shadow”?*, 20 OHIO ST. TECH. L.J. 317, 319 (2024).

²⁰ Data brokers are distinct from consumer reporting agencies in several ways. Most crucially, the latter are regulated by the Fair Credit Reporting Act, 15 U.S.C. §1681 (1970).

²¹ Natasha Singer, *Axiom, the Quiet Giant of Consumer Database Marketing*, N.Y. TIMES (June 16, 2012), <https://www.nytimes.com/2012/06/17/technology/axiom-the-quiet-giant-of-consumer-database-marketing.html> [https://perma.cc/58CA-7U7L]; see also DATA BROKER REPORT, *supra* note 8, for a wealth of information on data brokers, what they do, and what the risks are.

volume: “[Acxiom’s] databases contain information about 700 million consumers worldwide with over 3000 data segments for nearly every U.S. consumer.”²² Despite this, most people are unaware of these companies or their practices: “Much of this activity takes place without consumers’ knowledge.”²³

People are also unaware of their ability to see the data collected about them:²⁴

In the past Acxiom has allowed consumers to see the part of their dossier gathered from public documents, but the request process is onerous. Anyone interested has to send in their Social Security number, date of birth, driver’s license number, current address, phone number and email address, as well as a \$5 check. Few have cleared this hurdle. Between 2009 and mid 2012 when they sent information about this process to a Congressional panel, between 77 and 342 people had asked to see their files every year, with just two to 16 annually providing enough information to get access to their file.

It is hard to see this as informed consent within the spirit of the FIPPs.

Acxiom planned to ease the process but “had hoped to start letting individuals see their consumer files by mid summer [2013] but has run into delays. ‘It’s enormously difficult to do this,’ said [Tim] Suther, who has overseen the company’s global marketing, strategy and business development activities.”²⁵

The second problem today is that there are many more ways to collect data about people. Search engines know which links you click on. In part, they use that to improve their results—“for query X, more people preferred the third answer”—but the information is also used to build behavior and interest profiles on individuals.²⁶ Websites not only track you directly when you visit their sites, but they also purchase other data about you²⁷ and use technical means to track you

²² DATA BROKER REPORT, *supra* note 8, at 8.

²³ *Id.* at 49.

²⁴ Adam Tanner, *Finally You’ll Get to See the Secret Consumer Dossier They Have on You*, FORBES (June 25, 2013), <https://www.forbes.com/sites/adamtanner/2013/06/25/finally-youll-get-to-see-the-secret-consumer-dossier-they-have-on-you/> [<https://perma.cc/SV56-2KCD>].

²⁵ See *id.*

²⁶ It is easy to determine by inspection that Google knows which link you clicked on. Consider, for example, a search for “Supreme Court”. The search result page properly displays <https://www.supremecourt.gov>. If, however, you right-click on the hyperlink and select “Copy link”, you will get something like https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.supremecourt.gov/&ved=2ahUKEwi_IPOA5OuJAXVwKlkFHRRebBZYQFnoECBwQAQ&usg=AOvVaw0upQwL243vKbw9dvAKKirC. Obviously, that link takes you to Google. While some of the fields in the URL are obscure and may in fact be encrypted, is also clear that Google is being told what site it should redirect you to: the Supreme Court’s actual website. Google’s webpages are quite silent on just how this information is used; however, the Firefox browser gives you an option to copy the link “without tracking”. In other words, the developers of Firefox believe that tracking is one purpose of that indirection. The quality of results from a search engine have a long been a concern of Google’s. In fact, one of the earliest papers describing how Google works describes the PageRank algorithm at the heart of their original system for producing high-quality answers; see Sergey Brin & Larry Page, *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, 30 COMPUT. NETWORKS & ISDN SYS. 107 (1998) (“These maps allow rapid calculation of a web page’s ‘PageRank’, an objective measure of its citation importance that corresponds well with people’s subjective idea of importance. Because of this correspondence, PageRank is an excellent way to prioritize the results of web keyword searches.”). Maintaining search result quality is a continuing problem for Google; see, e.g., Kevin Purdy, *Google Stops Letting Sites Like Forbes Rule Search for ‘Best CBD Gummies’*, ARS TECHNICA (Nov. 20, 2024), <https://arstechnica.com/gadgets/2024/11/google-cracks-down-on-parasite-seo-punishing-established-publishers/> [<https://perma.cc/PA7V-P3GP>]; see also Steven M. Bellovin et al., *It’s Too Complicated: How the Internet Upends Katz, Smith, and Electronic Surveillance Law*, 30 HARV. J.L. & TECH. 1, 66 (2016).

²⁷ See e.g., *The Dow Jones Privacy Notice*, DOW JONES (Oct. 1, 2024), <https://www.dowjones.com/privacy-policy/> [<https://perma.cc/Q53H-9ERE>] (“We may receive personal data about you in connection with the Dow Jones Services from publicly and commercially available sources, Dow Jones Affiliates, . . . business partners, . . . and, if applicable to you, the third party provider of your subscription. . .”).

elsewhere.²⁸

All of this information is, of course, disclosed in privacy policies; however, few people actually read them. In fact, according to a study by Professors Aleecia McDonald and Lorrie Cranor, the time and opportunity cost to do so are prohibitive.²⁹

[U]sing the point estimate of 244 hours per year to read privacy policies per person means an average of 40 minutes a day. This is slightly more than half of the estimated 72 minutes a day people spend using the Internet.

They estimate the opportunity cost of this activity at over \$3,500 per year.³⁰ It is perhaps the supreme irony that Chief Justice John Roberts himself does not pay attention to this fine print.³¹ A report to President Obama said it well: “[o]nly in some fantasy world do users actually read these notices and understand their implications before clicking to indicate their consent.”³²

Furthermore, privacy policies are often unhelpful. For example, Professor Reidenberg et al. pointed to vague statements, e.g., “We may collect personal information and other information about you from business partners, contractors and other third parties.”³³ Users do not know if information will be collected, from whom, or what that information might be. Some of the issue appears to be lack of regulatory oversight; companies whose privacy policies are governed by regulation are significantly less vague.³⁴ Furthermore, even experts can misunderstand what is actually said.³⁵

A staff report from the Federal Trade Commission noted that mobile devices are particularly problematic.³⁶

One theme was that consumers do not know or understand current information collection and use practices occurring on mobile devices. According to one participant, because consumers are unaware that many of these practices are taking place, they do not look for options providing them with control over such practices. Another participant noted that when made aware of these practices, consumers typically are surprised and view the practices as underhanded. Participants noted that when disclosures are made, consumers often do not understand them.

²⁸ See *id.* (“[W]e, as well as our third-party service providers, may obtain information about your online activity to provide you with advertising about products and services tailored to your individual interests. We also engage third party online advertising companies, advertisers and ad networks to help target our messaging to visitors through interest-based and contextual means. These third parties may collect information about your activities on our Dow Jones Services and on other websites or apps including offline purchases, for such purpose. The networks use this information to show you advertisements on the Dow Jones Services or other third-party websites and apps that may be tailored to your individual interests.”).

²⁹ Aleecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, 4 I/S: J.L. & POL’Y INFO. SOC’Y 543, 563 (2008).

³⁰ *Id.* at 564.

³¹ See Debra Cassens Weiss, *Chief Justice Roberts Admits He Doesn’t Read the Computer Fine Print*, A.B.A. J. (Oct. 20, 2010), https://www.abajournal.com/news/article/chief_justice_roberts_admits_he_doesnt_read_the_computer_fine_print/ [<https://perma.cc/QT32-NG4E>].

³² PRESIDENT’S COUNCIL OF ADVISORS ON SCI. AND TECH., *BIG DATA AND PRIVACY: A TECHNOLOGICAL PERSPECTIVE*, at xi (2014), [hereinafter PCAST REPORT], https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf [<https://perma.cc/qq4N-S3W9>].

³³ Joel R. Reidenberg et al., *Ambiguity in Privacy Policies and the Impact of Regulation*, 45 J. LEGAL STUD. S163, S166 (2016).

³⁴ *Id.* at S181.

³⁵ Joel R. Reidenberg et al., *Disagreeable Privacy Policies: Mismatches Between Meaning and Users’ Understanding*, 30 BERKELEY TECH. L.J. 39, 42 (2015).

³⁶ FTC, *MOBILE PRIVACY DISCLOSURES: BUILDING TRUST THROUGH TRANSPARENCY* 10 (2013).

The report also noted the danger of location collection:³⁷

Third, mobile devices can reveal precise information about a user's location that could be used to build detailed profiles of consumer movements over time and in ways not anticipated by consumers. Indeed, companies can use a mobile device to collect data over time and 'reveal [] the habits and patterns that mark the distinction between a day in the life and a way of life.' Even if a company does not intend to use data in this way, if the data falls in the wrong hands, the data can be misused and subject consumers to harms such as stalking or identity theft.

Mobile devices also create voluminous amounts of location data.³⁸ The Supreme Court itself has recognized that cell site location information (CSLI) is not shared voluntarily with phone companies:³⁹

Cell phone location information is not truly "shared" as one normally understands the term. In the first place, cell phones and the services they provide are "such a pervasive and insistent part of daily life" that carrying one is indispensable to participation in modern society. Second, a cell phone logs a cell-site record by dint of its operation, without any affirmative act on the part of the user beyond powering up. Virtually any activity on the phone generates CSLI, including incoming calls, texts, or e-mails and countless other data connections that a phone automatically makes when checking for news, weather, or social media updates. Apart from disconnecting the phone from the network, there is no way to avoid leaving behind a trail of location data. As a result, in no meaningful sense does the user voluntarily "assume [] the risk" of turning over a comprehensive dossier of his physical movements.

Some apps are far worse. It is undoubtedly convenient to use a phone for turn-by-turn directions; however, the service provider may obtain continuous location information during the drive.⁴⁰

Additionally, new privacy dangers are posed by the so-called "Internet of Things":⁴¹

Cars, door locks, contact lenses, clothes, toasters, refrigerators, industrial robots, fish tanks, sex toys, light bulbs, toothbrushes, motorcycle helmets—these and other everyday objects are all on the menu for getting "smart." Hundreds of small start-ups are taking part in this trend—known by the marketing catchphrase "the internet of things."

And these devices all pose risks: "There's just one catch, which often goes unstated: If their novelties take off without any intervention or supervision from the government, we could be inviting a nightmarish set of security and privacy vulnerabilities into the world. And guess what. No one is really doing much to stop it."⁴² Most of these devices lack screens, keyboards, and mice. If it is hard to know what a computer or phone is doing, how can one tell the behavior of an

³⁷ *Id.* at 3.

³⁸ See *Carpenter v. United States*, 585 U.S. 296, 301 (2018).

³⁹ *Id.* at 314 (citations omitted).

⁴⁰ It is in fact unclear, even to technically sophisticated users, what location information is shared and when; see Steven M. Bellovin et al., *It's Too Complicated: How The Internet Opens Katz, Smith, and Electronic Surveillance Law*, 30 HARV. J.L. & TECH. 1, 83–88 (2016) ("Whether a mapping application is sending its location to the application provider frequently, occasionally, or never need not manifest itself in the behavior of the software.").

⁴¹ Farhad Manjoo, *A Future Where Everything Becomes a Computer Is as Creepy as You Feared*, N.Y. TIMES (Oct. 10, 2018), <https://www.nytimes.com/2018/10/10/technology/future-internet-of-things.html> [<https://perma.cc/AUU3-4V5P>].

⁴² *Id.*

Internet-connected hairbrush⁴³ or thermometer⁴⁴? This particular use is innocuous, in that only zip code data is used for targeted advertising. But the collection mechanism exists, and future companies may not be as ethical.

The reason for all of this tracking is, of course, to monetize the information, and in particular monetize it by using it for targeted advertising.⁴⁵ It is no surprise that two of the most successful Internet companies, Google and Facebook, are also two of the biggest collectors of personal information; together, they control a majority of the online advertising market.⁴⁶ In fact, more than half of U.S. advertising is online,⁴⁷ and hence driven by personal information.

Some have called advertising “the original sin of the Internet”:⁴⁸

I have come to believe that advertising is the original sin of the web. The fallen state of our Internet is a direct, if unintentional, consequence of choosing advertising as the default model to support online content and services. Through successive rounds of innovation and investor storytime, we’ve trained Internet users to expect that everything they say and do online will be aggregated into profiles (which they cannot review, challenge, or change) that shape both what ads and what content they see. Outrage over experimental manipulation of these profiles by social networks and dating companies has led to heated debates amongst the technologically savvy, but hasn’t shrunk the user bases of these services, as users now accept that this sort of manipulation is an integral part of the online experience.

There are, of course, benefits to an ad-supported Internet: it makes it accessible to more people.⁴⁹ Facebook’s home page formerly bragged that “it’s free and always will be.”⁵⁰ But the

⁴³ John Kell, *L’Oreal’s “Smart” Hairbrush Wants to Help Solve a Huge Beauty Problem*, FORTUNE (Jan. 3, 2017), <https://fortune.com/2017/01/03/loreal-smart-hairbrush-ces/> [<https://perma.cc/JTR9-8QSH>].

⁴⁴ Sapna Maheshwari, *This Thermometer Tells Your Temperature, Then Tells Firms Where to Advertise*, N.Y. TIMES (Oct. 23, 2018), <https://www.nytimes.com/2018/10/23/business/media/fever-advertisements-medicine-clorox.html> [<https://perma.cc/F7V5-SRQ8>].

⁴⁵ See, e.g., Lucy Handley, *L’Oreal’s Smart Brush ‘Listens’ to Hair, Recommends Luxury Treatments*, CNBC (Jan. 4, 2017), <https://www.cnbc.com/2017/01/04/loreal-smart-brush-listens-to-hair-recommends-luxury-treatments.html> [<https://perma.cc/TFF6-FGAB>] (“It will then provide hair tips and recommend L’Oreal’s Kérastase products.”). Exactly what is collected will vary with the device. For example, smart TVs know what you’re watching; this is used for advertising. See, e.g., Mohamed Al Elew and Gabriel Hongdsutit, *Your Smart TV Knows What You’re Watching*, MARKUP (Dec. 12, 2023), <https://themarkup.org/privacy/2023/12/12/your-smart-tv-knows-what-youre-watching> [<https://perma.cc/SYQ2-2G22>] (“The data is then used for content recommendations and ad targeting.”).

⁴⁶ Rani Molla, *Google’s and Facebook’s Share of the U.S. Ad Market Could Decline for the First Time, Thanks to Amazon and Snapchat*, VOX (Mar. 19, 2018), <https://www.vox.com/2018/3/19/17139184/google-facebooks-share-digital-advertising-ad-market-could-decline-amazon-snapchat> [<https://perma.cc/ULV5-ZVRH>].

⁴⁷ Lucas Shaw, *Google, Facebook Lead Digital’s March to Half of U.S. Ad Market*, BLOOMBERG (Sept. 20, 2018), <https://www.bloomberg.com/news/articles/2018-09-20/google-facebook-lead-digital-s-march-to-half-of-u-s-ad-market> [<https://perma.cc/H52J-549P>].

⁴⁸ Ethan Zuckerman, *The Internet’s Original Sin*, ATLANTIC (Aug. 14, 2014), <https://www.theatlantic.com/technology/archives/2014/08/advertising-is-the-internets-original-sin/376041/> [<https://perma.cc/R3S3-6GTD>].

⁴⁹ *Id.* (“Charging users for the service would have blocked most of our potential customers—most of the world still doesn’t have a credit card today, and fewer did in 1995 . . . but because Tripod’s services were free and ad supported, users around the world found us and began posting webpages they could not host elsewhere”).

⁵⁰ Per the Wayback Machine at the Internet Archive, that brag apparently changed around August 7, 2019. *Contrast* <https://web.archive.org/web/20190806000724/https://www.facebook.com/> [<https://perma.cc/BK5V-NSNM>], from August 6, 2019, *with* <https://web.archive.org/web/20190808000039/https://www.facebook.com/> [<https://perma.cc/L48M-B2BP>], from August 8. News reports suggest that Meta plans to charge \$14–17 for ad-free access in Europe; see Javier Espinoza & Hannah Murphy, *Ad-Free Facebook, Instagram Access Planned for \$14 per Month in Europe*, ARS TECHNICA (Oct. 3, 2023), <https://arstechnica.com/tech-policy/2023/10/ad-free-facebook-instagram-access-planned-for-14-per-month-in-europe/> [<https://perma.cc/XZF4-KH5T>], though after questions were raised about the legality of this scheme under the GDPR they are now contemplating both that scheme at a lower price and a “a free model characterized as serving ‘less personalized ads’ with consent. The new option will focus on contextual advertising generated by content viewed during a singular browsing session. Meta will not employ broader user profiling to serve ads”; see Joe Duball, *Meta Reshapes EU Personalized Advertising Offerings Again*, IAPP NEWS (Nov. 12, 2024), <https://iapp.org/news/a/meta-reshapes-eu-personalized->

cost in privacy is considerable: Internet ads are perceived to work best when they are highly targeted; this means that advertising companies must collect as much information as they can about their users. That in turn has meant tracking them across sites, using tools such as “web beacons”⁵¹ and the ubiquitous Facebook “like” buttons. Together, this has made a mockery of user consent: if you cannot see a web beacon—by definition, they are often invisible—and if you don’t know that the existence of a Facebook “like” button on a page means that Facebook can track your visit there, it is impossible to control your information.

There are other facets of Internet advertising that render useless any attempts by Internet users to understand who has their data. Few Internet ads are actually displayed by the hosting sites; instead, they come from Internet advertising companies such as Google’s DoubleClick unit.⁵² This is indeed acknowledged by the privacy policies of many media companies. For example, Warner Bros. Discovery’s privacy policy says “[w]e may disclose Information with business partners and third parties (e.g., other companies, retailers, distributors, social media networks, research organizations, publishers, and non-profit organizations) for their own purposes, including marketing products or services to you.”⁵³ Similarly, Major League Baseball’s website says:⁵⁴

We allow and use select Service Providers and/or Other Companies that may act as third parties as defined under applicable data privacy laws, such as those who deliver advertisements, content, or social networking or which provide other services associated with the MLB Properties, to collect data and/or serve ads when you visit MLB Properties. We may also make available certain data you have given to us (e.g., IP address) or that we have collected through publicly available means to third parties engaged in targeting advertisements on the MLB Properties including for use in accordance with such third party’s privacy policy. Network advertisers may use information (not including your name, address, email address or telephone number) about your visits to MLB Properties in order to provide advertisements about goods and services of interest to you. This may happen through permission given to such third parties to set cookies, web beacons or similar technologies in certain locations within the MLB Properties, including but not limited to certain of our emails. *Third party companies that manage and deliver advertisements to websites and applications such as ours are commonly referred to as “network advertisers.”* A permitted network advertiser may use cookies, web beacons or similar technologies to collect information about your interaction with the MLB Properties in order to tailor certain advertisements and content delivered within the MLB Properties and on other websites within such network advertiser’s ad network.

In other words, to protect your privacy you must know the policies of not just the sites you visit, but the policies of their advertisers as well. Furthermore, it is not just a single layer; often, the primary advertising site will redirect you to a third, which can send you to a fourth, *ad*

advertising-offerings-again [https://perma.cc/K7ZD-3EEY].

⁵¹ See *Web Beacon*, IAPP, <https://iapp.org/resources/article/web-beacon/> [https://perma.cc/CWY3-WKRH] (“Also known as a web bug, pixel tag or clear GIF, a web beacon is a clear graphic image (typically one pixel in size) that is delivered through a web browser or HTML e-mail.”).

⁵² See, e.g., Dan Wallach, FTC workshop *The Big Picture: Comprehensive Online Data Collection* transcript, December 6, 2012, Session 1, at 22, available at <https://www.ftc.gov/sites/default/files/documents/videos/big-picture-comprehensive-online-data-collection-session-1/4/121206bigpicturept1.pdf> [https://perma.cc/F76T-HJ5N].

⁵³ *Warner Bros. Discovery Privacy Policy*, WARNER BROS. DISCOVERY (Mar. 27, 2024) (emphasis added) <https://www.warnermediaprivacy.com/policycenter/b2c/WMNS/en-us/> [https://perma.cc/C6T7-Y8NG].

⁵⁴ *MLB Privacy Policy*, MLB, <https://www.mlb.com/official-information/privacy-policy> [https://perma.cc/GL46-8UPF].

nauseum.⁵⁵ Even advertisers who do not participate in the actual ad display learn something about users, via a bidding process:⁵⁶

Rubicon is not just a sales platform for Web site operators. It's an analytics system that uses consumer data to help sites figure out how much their visitors are worth to advertisers. Most sites . . . compile data about their own visitors through member registration or by placing bits of computer code called cookies on people's browsers to collect information about their online activities. To those first-party profiles, Rubicon typically adds details from third-party data aggregators, like BlueKai or eXelate, such as users' sex and age, interests, estimated income range and past purchases. Finally, Rubicon applies its own analytics to estimate the fair market value of site visitors and the ad spaces they are available to see. The whole process typically takes less than 30 milliseconds.

The dynamic nature of the advertising ecosystem makes determining which sites are showing you ads—in other words, which sites' privacy policies you need to read—quite daunting, even for experts. This is best seen by noting how hard it is for even well-meaning websites to eliminate obnoxious (and probably fraudulent) ads:⁵⁷

Within about an hour we had successfully replicated the issue and pinpointed the source. Our AdOps team moved quickly to alert the vendor whose network was being used to serve the ad, and we blocked the source of the issue in Google's tools. By the end of the day we felt we had successfully blocked the ad and had stopped receiving reports of redirects for the day. Whoever was behind the ad, however, kept finding ways into the system throughout the week on Vox Media sites and many others around the web. Our tools for blocking this require us to identify the source of each malicious ad and block it, which is reactive and not preventative.

In other words, it took a team of professionals an hour to find the actual source of one ad, but the offender—a different web site—kept moving to different places. For our purposes, what matters is that each such site could have a different privacy policy, and that new sites are appearing constantly. An ordinary user would have no hope of finding even the immediate ad source, let alone the identities of any intermediaries.

Some sites explicitly list their possible partners. The transparency is good, but the numbers can be shocking. PayPal, for example, lists a huge number of sites with which it will sometimes share information.⁵⁸ From the site, it is clear that many of the entries are country-specific, and that many are unquestionably necessary, either to carry out their services or to comply with laws and regulations. But a fair number of companies are there for marketing or to “deliver personalised [sic] advertising.”⁵⁹

⁵⁵ Wallach, *supra* note 52.

⁵⁶ See Natasha Singer, *Your Online Attention, Bought in an Instant by Advertisers*, N.Y. TIMES (Nov. 17, 2012), <https://www.nytimes.com/2012/11/18/technology/your-online-attention-bought-in-an-instant-by-advertisers.html> [<https://perma.cc/UE7U-BJC8>].

⁵⁷ Winston Hearn, *Why Ads Keep Redirecting You to Scummy Sites and What We're Doing About It*, VOX MEDIA (Jan. 22, 2018), <https://product.voxmedia.com/2018/1/22/16902862/why-ads-redirect-to-giftcards-and-what-were-doing-to-secure-them> [<https://perma.cc/6CZV-T6C9>].

⁵⁸ See *List of Third Parties (Other Than PayPal Customers) with Whom Personal Information May be Shared*, PAYPAL (Oct. 1, 2024), <https://www.paypal.com/ie/webapps/mpp/ua/third-parties-list> [<https://perma.cc/6QP4-DLZS>]; see also Rebecca Ricks (@baricks), X (Jan. 18, 2018, 2:05 PM), <https://x.com/baricks/status/954067244321050624> [<https://perma.cc/2S9H-J4PF>] (providing a visualization of third parties PayPal shares customer information with, though the date of the source data is not stated).

⁵⁹ See *List of Third Parties (Other Than PayPal Customers) with Whom Personal Information May be Shared*, *supra* note 58. On a recent

The risks of tracking and dossier compilation go far beyond marketing. Recommendation engines are agnostic to what they're suggesting.⁶⁰ It may be something benign, such as what movie you might want to watch next. But it can also be used—and abused—to spread propaganda. Professor Zeynep Tufekci has argued that YouTube is a potent radicalizing engine: “[i]t seems as if you are never ‘hard core’ enough for YouTube’s recommendation algorithm. It promotes, recommends, and disseminates videos in a manner that appears to constantly up the stakes. Given its billion or so users, YouTube may be one of the most powerful radicalizing instruments of the 21st century.”⁶¹ There do not appear to be malicious political intentions at play here, but there certainly could be. Apart from the widely reported “fake news” phenomena during the 2016 election,⁶² where invented stories spread rapidly on social media, an experiment—done with the cooperation of Facebook—showed that Facebook messages could affect voter turnout.⁶³ More ominously, in Myanmar, the military apparently “turned the social network into a tool for ethnic cleansing, according to former military officials, researchers and civilian officials in the country.”⁶⁴ More than 700,000 Rohingya have fled the country in fear of their lives.⁶⁵

Finally, as Professor Paul Ohm has argued, the very existence of these digital dossiers is itself a serious risk.⁶⁶

In my work, I’ve argued that these databases will grow to connect every individual to at least one closely guarded secret. This might be a secret about a medical condition, family history, or personal preference. It is a secret that, if revealed, would cause more than embarrassment or

visit to London, a news website I visit regularly in the U.S. gave me more precise data in a privacy pop-up. A colleague in London confirmed the data in early 2025. Her experiments showed that 158 partners can store and/or access information on a device, 172 partners can conduct “personalised [sic] advertising and content, advertising and content measurement, audience research and services development,” 58 partners can “[u]se precise geolocation data,” without consent, 17 can “actively scan device characteristics for identification”, i.e., attempt to defeat attempts at privacy by, say, deleting cookies, 108 can “match and combine data from other data sources,” and 109 can “identify devices based on information transmitted automatically.” (on file with the author).

⁶⁰ The primary goal of a search engine is to deliver results perceived as useful by the querier. See, e.g., a white paper originally commissioned by Google and later published: Eugene Volokh & Donald M. Falk, *First Amendment Protection for Search Engine Search Results*, 8 J. OF L., ECON. & POL’Y. 883, 884 (2012) (“Third, and most valuably, search engines select and sort the results in a way that is aimed at giving users what the search engine companies see as the most helpful and useful information. (That is how each search engine company tries to keep users coming back to it rather than to its competitors”). In that, they do not necessarily care about anything else. However, see Timothy Graham & Mark Andrejevic, *A Computational Analysis of Potential Algorithmic Bias on Platform X During the 2024 US Election* (2024), <https://eprints.qut.edu.au/253211/> [<https://perma.cc/XN7L-WJLM>], for a discussion on how they can be biased as an exercise of their own First Amendment rights. see also *Anderson v. TikTok*, 116 F.4th 180 (3d Cir. 2024). Nonetheless, such biases or perceived ideological alignment can drive away customers; see, e.g., Callie Holtermann, *With Surge in New Users, Bluesky Emerges as X Alternative*, N.Y. TIMES (Nov. 12, 2024), <https://www.nytimes.com/2024/11/12/style/bluesky-users-election.html> [<https://perma.cc/9UVF-4CY5>] (“Others are frustrated by the type of content that seems to be rising to the top of their feeds. ‘You were getting this awful timeline of far-right, white supremacist, conspiracy theory posts—which the great majority of people don’t want to interact with on a daily basis,’ said Dr. McGregor.”).

⁶¹ Zeynep Tufekci, *YouTube, The Great Radicalizer*, N.Y. TIMES (Mar. 10, 2018), <https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html> [<https://perma.cc/Z5MT-KGRY>].

⁶² See, e.g., Andrew Higgins et al., *Inside a Fake News Sausage Factory: ‘This Is All About Income’*, N.Y. TIMES (Nov. 25, 2016), <https://www.nytimes.com/2016/11/25/world/europe/fake-news-donald-trump-hillary-clinton-georgia.html> [<https://perma.cc/5J7U-MCYT>].

⁶³ See *id.*

⁶⁴ Paul Mozur, *A Genocide Incited on Facebook, with Posts from Myanmar’s Military*, N.Y. TIMES (Oct. 15, 2018), <https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html> [<https://perma.cc/K4MN-WEZH>].

⁶⁵ *Id.*

⁶⁶ Paul Ohm, *Don’t Build a Database of Ruin*, HARV. BUS. REV. (Aug. 23, 2012), <https://hbr.org/2012/08/dont-build-a-database-of-ruin> [<https://perma.cc/8AHH-UCRQ>]; for a more detailed discussion of databases of ruin, see also Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701 (2010).

shame; it would lead to serious, concrete, devastating harm. And these companies are combining their data stores, which will give rise to a single, massive database. I call this the Database of Ruin.

As this author has often argued, something that does not exist cannot be stolen or otherwise misused.⁶⁷

Apart from the ubiquitous, invisible collection of data, there is another technological development that has had a great effect on our privacy: machine learning.⁶⁸ Machine learning (“ML”), sometimes called AI or artificial intelligence,⁶⁹ relies on “training” the system with a large amount of data.⁷⁰ The ML system finds patterns and, in effect, makes predictions.⁷¹ Thus, when Netflix suggests movies to a user, or Amazon says “people who bought this also bought this other thing”, those suggestions are not the result of human reasoning and curation. Rather, the algorithm—more precisely, the algorithm plus the training data—has found correlations that users hopefully find useful.⁷²

Some people find these recommendations annoying and intrusive. “Many consumers appreciate having computers delve into their hearts and heads. But some say it gives them the willies, because the machines either know them too well or make cocksure assumptions about them that are way off base.”⁷³ In one famous incident, Target used sales data *correlated with* pregnancy to send targeted ads; they identified one teenage girl as pregnant before her parents knew.⁷⁴ Generative AI⁷⁵ may make this issue worse, since these systems are prone to errors commonly known as “hallucinations.”⁷⁶

The fact that these large databases exist is troubling enough from a privacy perspective. What makes them problematic, though, is that they are dual use: they are used both intrusively and to deliver desired results. When you type the name of a restaurant into Google, even if it is a very

⁶⁷ See, e.g., Steven M. Bellovin, THINKING SECURITY: STOPPING NEXT YEAR’S HACKERS 106 (2015) (“[a] key that doesn’t exist on a machine can’t be stolen from it”).

⁶⁸ Soumia Zohra El Mestari, et al., *Preserving Data Privacy in Machine Learning Systems*, 137 COMPUT. & SEC., 103605, Feb. 2024.

⁶⁹ Artificial intelligence is a broad field, going back more than 60 years. Its goal is, roughly, to produce a computer that can think (whatever that means). Machine learning is one specific technology (more precisely, a set of technologies) for achieving AI. Because of how well it works, it is currently the most favored approach to achieving artificial intelligence; as a result, the two terms are often conflated. Clara Piloto, *Artificial Intelligence vs Machine Learning: What’s the Difference?*, MIT PRO. EDUC., <https://professionalprograms.mit.edu/blog/technology/machine-learning-vs-artificial-intelligence/> [<https://perma.cc/V7A9-ZWXL>].

⁷⁰ Timothy Lee, *How a Stubborn Computer Scientist Accidentally Launched the Deep Learning Boom*, ARS TECHNICA (Nov. 11, 2024), <https://arstechnica.com/ai/2024/11/how-a-stubborn-computer-scientist-accidentally-launched-the-deep-learning-boom/> [<https://perma.cc/ND7B-RNMY>] (“This speed allowed them to train a larger model—and to train it on many more training images”).

⁷¹ U.S. Patent No. 10,607,154 B2 fig.2 items 250 and 280.

⁷² See, e.g., Robert M. Bell, Yehuda Koren & Chris Volinsky, *All Together Now: A Perspective on the Netflix Prize*, 23 CHANCE 24, <https://www.tandfonline.com/doi/pdf/10.1080/09332480.2010.10739787> [<https://perma.cc/KMA9-QMCE>].

⁷³ Jeffrey Zaslow, *If TiVo Thinks You Are Gay, Here’s How to Set It Straight*, WALL ST. J. (Nov. 26, 2002), <https://www.wsj.com/articles/SB1038261936872356908> [<https://perma.cc/YN4E-TBMN>].

⁷⁴ Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES (Feb. 16, 2012), <https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html> [<https://perma.cc/AUN3-JUP4>].

⁷⁵ Haomiao Huang, *The Generative AI Revolution Has Begun—How Did We Get Here?*, ARS TECHNICA (Jan. 30, 2023), <https://arstechnica.com/gadgets/2023/01/the-generative-ai-revolution-has-begun-how-did-we-get-here/> [<https://perma.cc/EZ8U-86R2>].

⁷⁶ See, e.g., Ashley Belanger, *OpenAI Faces Defamation Suit After ChatGPT Completely Fabricated Another Lawsuit*, ARS TECHNICA (June 9, 2023), <https://arstechnica.com/tech-policy/2023/06/openai-sued-for-defamation-after-chatgpt-fabricated-yet-another-lawsuit/> [<https://perma.cc/4XVX-CYN8>].

generic restaurant name, you will likely be shown the web site of the one near you. How? Google's algorithms consider your IP address, which is correlated with your location,⁷⁷ and the fact that most people look for restaurants near where they are.

This, then, is the dilemma. The FIPPs state "There must be a way for an individual to find out what information about him is in a record and how it is used."⁷⁸ However, ML algorithms do not work that way. They do not provide any explanation for *why* a particular correlation exists and hence cannot say what in a particular record produced the given answer.⁷⁹ It is therefore impossible to assert or deny that a particular datum was used for some purpose.

The advent of machine learning raises two more fundamental issues with the FIPPs: the entire scheme is based on the notion of protecting personally identifiable information ("PII"). Four of the five principles, in fact, explicitly refer to individual records. However, ML algorithms do not need to know someone's identity to invade privacy. The "My Tivo Thinks I'm Gay" incident⁸⁰ is just one example, but in principle, most recommendation algorithms do not need PII.⁸¹ Furthermore, it is often possible to deanonymize records.⁸²

These algorithms may be able to infer PII, even if none of the data used contains explicit identifying information. Consider a turn-by-turn map application. Over time, the location that someone frequently leaves from in the morning or returns to in the evening is likely their home address. Does this count as "a record of identifiable information" about a particular person? What if the PII is obtainable only by combining one highly revealing dataset with another that contains the actual identifiers?

II. DATA SECURITY AND HARM

The fifth principle in the FIPPs requires data security, to "prevent misuse of the data."⁸³ This point and the harm that can result from security breaches are worth a separate discussion.

Empirically, our data is at great risk. Many large organizations have proven unable to protect themselves against attacks.⁸⁴ It is, of course, obvious that without data security, one cannot

⁷⁷ *Privacy & Terms*, GOOGLE, <https://policies.google.com/technologies/location-data?hl=en> [<https://perma.cc/99TK-5VEB>]. IP addresses—Internet Protocol addresses—are the Internet's analogue to phone numbers. For reasons beyond the scope of this note, IP addresses generally indicate one's rough locale—think of how, in the pre-cellular era, the area code and exchange of a phone number denoted the city. In both cases, the association is done for strong technical reasons.

⁷⁸ HEW REPORT, *supra* note 13, at xx.

⁷⁹ Bell et al., *supra* note 72 ("machine learning methods tend to center on algorithms (black boxes), where the focus is on the quality of predictions—rather than 'understanding' what drives particular predictions")

⁸⁰ Zaslow, *supra* note 73.

⁸¹ This may not be strictly true: PII may, in fact, be helpful. Some products are particularly appealing to certain demographics; knowledge of those could lead to more

accurate suggestions. Joseph A. Calandrino et al., "You Might Also Like:" *Privacy Risks of Collaborative Filtering*, in 2011 IEEE SYMPOSIUM ON SECURITY AND PRIVACY 231 (2011).

⁸² Michael Barbaro & Tom Zeller Jr., *A Face Is Exposed for AOL Searcher No. 4417749*, N.Y. TIMES (Aug. 9, 2006), <https://www.nytimes.com/2006/08/09/technology/09aol.html> [<https://perma.cc/M9SE-GNHM>]; see also Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1707–08 (2010).

⁸³ HEW REPORT, *supra* note 13, at xxi.

⁸⁴ Mike Isaac & Sheera Frenkel, *Facebook Security Breach Exposes Accounts of 50 Million Users*, N.Y. TIMES (Sept. 28, 2018), <https://www.nytimes.com/2018/09/28/technology/facebook-hack-data-breach.html> [<https://perma.cc/LB6J-77A4>].

guard against theft and hence misuse of data. Similarly, unauthorized modification of records leads to incorrect data about individuals. In other words, even if security were not called out specifically, it is implicit in the other principles. The FTC has used this theory to become a de facto privacy regulator.⁸⁵ Using its powers under Section 5 of the FTC Act,⁸⁶ the FTC has often moved against companies with inadequate data security, on the grounds that such behavior is *a priori* unfair competition; additionally, if consumers are promised that their data will be protected, failure to do so can constitute deception.

There are two caveats, however. First, the Act gives the FTC authority over “[u]nfair methods of competition . . . and unfair or deceptive acts or practices.”⁸⁷ Second, the offending behavior must be one that “causes or is likely to cause substantial injury to consumers.”⁸⁸ Both limitations are significant.

When can a security breach be considered as due to an unfair or deceptive act or practice? The FTC has never issued any regulations on this; however, in the past, it has held that failure to adopt “reasonable and appropriate” measures was in itself unfair.⁸⁹ This standard has been upheld by the Third Circuit,⁹⁰ though the Eleventh Circuit has disagreed.⁹¹ But the actual standards that companies must follow to avoid liability remain unclear: “[t]he painful reality is that we lack broadly applicable, specific standards, partially because of the ugly complexity of the problem. It involves not just technical standards but also corporate executive decision-making [sic] and reevaluation of practices over time.”⁹²

The second issue is how to define “substantial injury.” Traditionally, the FTC has held that “[m]onetary, health, and safety risks are common injuries considered ‘substantial,’ but trivial, speculative, emotional, and ‘other more subjective types of harm’ are usually not considered substantial for unfairness purposes.”⁹³ Courts have frequently held that mere disclosure of other information does not constitute harm.⁹⁴ Professor Danielle Citron has noted that “[f]or most courts, privacy and data security harms are too speculative and hypothetical, too based on subjective fears and anxieties, and not concrete and significant enough to warrant recognition.”⁹⁵

Harm can be cumulative, too. Professors Citron and Daniel Solove noted:

⁸⁵ See generally Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583 (2014).

⁸⁶ Federal Trade Commission Act, 15 U.S.C. § 45.

⁸⁷ 15 U.S.C. § 45(a)(1).

⁸⁸ 15 U.S.C. § 45(n).

⁸⁹ See Solove & Hartzog, *supra* note 85, at 643.

⁹⁰ *F.T.C. v. Wyndham Worldwide Corp.*, 799 F.3d 236, 255 (3d Cir. 2015) (“We thus conclude that Wyndham was not entitled to know with ascertainable certainty the FTC’s interpretation of what cybersecurity practices are required by § 45(a). Instead, the relevant question in this appeal is whether Wyndham had fair notice that its conduct could fall within the meaning of the statute.”)

⁹¹ *LabMD v. FTC*, 894 F.3d 1221, 1236 (11th Cir. 2018) (“In the case at hand, the cease and desist order contains no prohibitions. It does not instruct LabMD to stop committing a specific act or practice. Rather, it commands LabMD to overhaul and replace its data-security program to meet an indeterminable standard of reasonableness. This command is unenforceable.”)

⁹² Merritt Baer & Chinmayi Sharma, *What Cybersecurity Standard Will a Judge Use in Equifax Breach Suits?*, LAWFARE (Oct. 20, 2017), <https://www.lawfaremedia.org/article/what-cybersecurity-standard-will-judge-use-equifax-breach-suits> [https://perma.cc/RB8W-PSMM].

⁹³ Solove & Hartzog, *supra* note 85, at 639.

⁹⁴ *Id.* at n.48.

⁹⁵ Danielle Keats Citron, *The Privacy Policymaking of State Attorneys General*, 47 NOTRE DAME L. REV. 747, 798–799 (2016).

When privacy violations result in negative consequences, the effects are often small—frustration, aggravation, anxiety, inconvenience—and dispersed among a large number of people. When these minor harms are suffered at a vast scale, they produce significant harm to individuals, groups, and society. But these harms do not fit well with existing cramped judicial understandings of harm.⁹⁶

The problem of defining privacy harms is especially serious when it comes to federal court jurisdiction. The Supreme Court has repeatedly held that the judicial power of the courts is restricted to “cases” and “controversies,”⁹⁷ and that “[n]o principle is more fundamental to the judiciary’s proper role in our system of government than the constitutional limitation of federal-court jurisdiction to actual cases or controversies.”⁹⁸ But for there to be a “controversy” there must be harm, and concrete harm at that. As Justice Kavanaugh wrote in *TransUnion v. Ramirez*, “No concrete harm, no standing.”⁹⁹

Concrete harm being necessary for redress is inescapable, but the current societal consensus on what constitutes harm is inadequate. Apart from the risk of sensitive data theft resulting from the disclosure of private information, per the FIPPs, the inability to control disclosure of information is itself a problem. If privacy is, at root, “[t]he right of an entity (normally a person), acting in its own behalf, to determine the degree to which it will interact with its environment,”¹⁰⁰ undesired disclosure is by definition a privacy violation. However, the FTC and the courts may not have the authority to act.

III. ANALYSIS

We now revisit the five privacy principles originally spelled out¹⁰¹ and see why they no longer work.

There must be no personal data record-keeping systems whose very existence is secret. Few, if any, consumer-facing Internet companies are secret. That is, their existence is public, and at least some details about their data collection practices are known. Partly, this is out of necessity; California, a large consumer market, requires that privacy policies exist for online businesses.¹⁰² Nevertheless, though the information is available to those who know where to look, very few people are aware of these companies; as noted, the role of online advertising intermediaries is very hard to determine, even for skilled users. Furthermore, the California requirement which mandates privacy policies to be posted applies only to companies “that [collect] personally identifiable information through the Internet about individual consumers;”¹⁰³ other forms of data compilation and profiling are not covered. Thus, although this requirement is moderately satisfied in theory, in

⁹⁶ Danielle Keats Citron & Daniel J. Solove, *Privacy Harms*, 102 B.U. L. Rev. 793 (2022).

⁹⁷ U.S. CONST. art III, § 2, cl. 1.

⁹⁸ *Simon v. E. Kentucky Welfare Rts. Org.*, 426 U.S. 26, 37 (1976).

⁹⁹ *TransUnion LLC v. Ramirez* 141 S. Ct. 2190, 2200 (2021).

¹⁰⁰ Shirey, *supra* note 17, at 231.

¹⁰¹ See HEW REPORT, *supra* note 13.

¹⁰² Online Privacy Protection Act, CAL. BUS. & PROF. CODE §§ 22575–22579 (2004). California added considerably more requirements in the California Consumer Privacy Act (hereinafter CCPA); see generally CAL. CIV. § 1798.100–1798.199.100 (2020).

¹⁰³ Cal. Bus. & Prof. Code § 22575(a) (2013).

practice it is often not.

There must be a way for an individual to find out what information about him is in a record and how it is used. It is less clear that this requirement is satisfied, even in theory, except in California and a few other states.¹⁰⁴ The California Consumer Privacy Act (“CCPA”) does have stringent requirements, but these individual rights are only conferred to California residents.¹⁰⁵ Again, though, even if all companies make such information available, it is infeasible for consumers to request changes to a record they do not know exists.

There must be a way for an individual to prevent information about him that was obtained for one purpose from being used or made available for other purposes without his consent. There is no way for individuals to understand how their data is used. Much of it, of course, is used for advertising, but that covers a wide range of activities. Facebook has been accused of deliberate vagueness about how it establishes connections between people and how it uses information collected for two-factor authentication, a security feature, for targeting ads.¹⁰⁶ The risk of information leakage from targeted ads has even reached popular culture.¹⁰⁷

A 2020 study by Consumer Reports Digital Lab illustrated the difficulty that consumers have in actually exercising their rights under the CCPA.¹⁰⁸ For more than 40% of the sites they tested, at least some of their testers could not find the opt-out link;¹⁰⁹ indeed, a number of companies did not even provide one, despite the legal requirement for a “clear and conspicuous link” on their home pages.¹¹⁰ The opt-out process was often onerous, with more than half of the testers dissatisfied.¹¹¹ The result is that the companies’ default settings remained in effect. As Professor Gordon Hull noted, “[t]he only seemingly consistent rule is that the software will default to openness.”¹¹² More generally, default settings matter.¹¹³

The existence of partners muddies the issue even further. The most obvious case in point is the information that Cambridge Analytica obtained from Facebook.¹¹⁴ Facebook “routinely

¹⁰⁴ Several other states have adopted privacy statutes, and the number is constantly increasing; these are generally weaker than California’s, so we concentrate on the latter.

¹⁰⁵ See generally CAL. CIV. § 1798.115 (2020) (“Consumers’ Right to Know What Personal Information is Sold or Shared and to Whom”).

¹⁰⁶ See Kashmir Hill, *Facebook Is Giving Advertisers Access to Your Shadow Contact Information*, GIZMODO (Sept. 26, 2018), <https://gizmodo.com/facebook-is-giving-advertisers-access-to-your-shadow-co-1828476051> [<https://perma.cc/4GDD-AN9N>].

¹⁰⁷ See, e.g., “Family Planning,” KEVIN & KELL (Oct. 12, 2018), <https://www.kevinandkell.com/2018/kk1012.html> [<https://perma.cc/U9NQ-QNWY>].

¹⁰⁸ Maureen Mahoney, *California Consumer Privacy Act: Are Consumers’ Digital Rights Protected*, CR DIGIT. LAB (Oct. 1, 2020), https://innovation.consumerreports.org/CR_CCPA-Are-Consumers-Digital-Rights-Protected_092020_vf.pdf [<https://perma.cc/BCG4-2N3H>].

¹⁰⁹ *Id.* at 4.

¹¹⁰ Cal. Civ. Code § 1798.135(a)(1).

¹¹¹ Mahoney, *supra* note 108, at 5.

¹¹² Gordon Hull, *Successful Failure: What Foucault Can Teach Us About Privacy Self-Management in a World of Facebook and Big Data*, 17 ETHICS & INFO. TECH. 89, 93 (2015).

¹¹³ That defaults matter is commonly accepted in the software community. See, e.g., *id.* at n.7 (“Research indicates that most individuals do not change software or other defaults (as, for example, 401(k) participation, which can be raised dramatically by simply switching from an opt-into an opt-out default). The reasons for this are partly economic—changing defaults (especially on Facebook) takes time and effort, and partly normalizing: the default setting communicates what an ‘average’ or ‘reasonable’ user ought to prefer.”)

¹¹⁴ See, Kevin Granville, *Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens*, N.Y. TIMES (Mar. 19, 2018), <https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html> [<https://perma.cc/LC32-TLZN>].

allows researchers to have access to user data for academic purposes—and users consent to this access when they create a Facebook account.”¹¹⁵ Arguably, someone did violate agreements, since the researcher who originally obtained the data appears to have been barred from further redistribution.¹¹⁶ From a privacy perspective, though, users consented to the original transfer and then lost control of their data.

There must be a way for an individual to correct or amend a record of identifiable information about him. As with the second point, though it may be true in theory, in practice, people are unaware of where their data is stored and hence of how it may be corrected, if at all. Furthermore, there are so many data collectors, such that attempting to correct all erroneous records would be extremely time-consuming—and one never knows when the next data collector will spring up.

Any organization creating, maintaining, using, or disseminating records of identifiable personal data must assure the reliability of the data for their intended use and must take precautions to prevent misuse of the data. As we have seen, even the most well-known Internet companies, such as Facebook, are vulnerable to technical security failures. No security paradigm should be based on the assumption that large collections of data can be protected.

The FIPPs, then, and by extension the whole concept of notice and consent, no longer fit the modern world. Many collectors are effectively unknown; there are also technical means of data collection that are opaque even to technically sophisticated users. Furthermore, the same data and databases are used to both answer user requests and invade privacy. Our ability to consent has vanished.

Others, of course, have noted the failure of consent. Professor Fred Cate noted, “as theoretically appealing as this approach may be, it has proven unsuccessful in practice.”¹¹⁷ He and Professor Mayer-Schönberger noted that “it certainly is not the optimal mechanism to ensure that either information privacy or the free flow of information is being protected.”¹¹⁸ More recently, Professor Susan Landau and Dr. Patricia Vargas Leon have discussed how device telemetry, the import of which is all but unknown to people, have exacerbated the problem.¹¹⁹

IV. SOME WAYS FORWARD

We now look at two technical approaches that might, under certain circumstances, provide alternative paths. The first, use restrictions, is a generally applicable mechanism, but poses considerable challenges. The second, differential privacy and in particular local differential privacy, has fewer conceptual and legal barriers, but is only applicable in restricted circumstances.

¹¹⁵ *Id.*

¹¹⁶ *Id.*

¹¹⁷ FRED H. CATE, CONSUMER PROTECTION IN THE AGE OF THE ‘INFORMATION ECONOMY’ 341–43 (Routledge 2016).

¹¹⁸ Fred H. Cate & Viktor Mayer-Schönberger, *Notice and Consent in a World of Big Data*, 3 INT’L DATA PRIV. L. 67 (2013).

¹¹⁹ Susan Landau & Patricia Vargas Leon, *Reversing Privacy Risks: Strict Limitations on the Use of Communications Metadata and Telemetry Information*, 21.2 COLO. TECH. L.J. 225, 233 (2023).

A. Use Restrictions

In contrast to notice and consent, some, most notably President Obama’s Council of Advisors on Science and Technology (“PCAST”), have advocated for use restrictions.¹²⁰ That is, instead of asking individuals to consent to the collection of their data, they are asked to consent to how it is used.

Professor Landau gives several specific examples of where use restrictions are employed today.¹²¹ She notes that a National Academies study committee¹²² concluded that the only way to control bulk signal intelligence collection is to restrict how the data is actually used. Other examples she gives, e.g., the Shibboleth information-sharing system, are examples of “attribute credentials.”¹²³ Attribute credentials, as opposed to the more common identity credentials, say what you are allowed to do rather than who you are. They are thus (generally) privacy-preserving. The easiest analogy is cash versus a line of credit: the former lets you buy things; the latter does, too, but based on who you are; it is therefore not privacy-preserving. Later, she and Dr. Vargas Leon suggested use control as a way to protect information gathered from metadata and device telemetry.¹²⁴

In many ways, this is an attractive idea. However, the details are extremely important. Note the careful wording of the PCAST report’s first recommendation: “[p]olicy attention should focus more on the actual uses of big data and less on its collection and analysis.”¹²⁵ The group did not recommend any particular mechanism based on use restriction, and in particular said nothing about how consent should be given and by whom; rather, they suggested it as a policy direction. In other words, they recommend that policymakers consider it, but do not suggest any way in which it can be accomplished.

It turns out that there are a number of problems with trying to employ use restrictions in a broader commercial sense, and in particular in trying to craft regulations to instantiate them as an alternative to notice and consent. The first is defining what a “use” is. Is each advertising campaign a new use? Assuredly not. Online advertising versus door-to-door sales? That’s a closer call; most people would view the latter as far more intrusive. A natural answer is to define categories of uses, but that isn’t easy, either. Would-be data abusers would naturally prefer broad, inclusive categories; privacy-sensitive individuals would prefer narrow ones. But who should define the categories, and how? How would they be updated over time?

Professor Cate suggested restricting use based on whether data was “per se harmful,” “per se not harmful,” or “sensitive,”¹²⁶ and balancing those criteria and “benefit maximization.”¹²⁷

¹²⁰ PCAST REPORT, *supra* note 32, at 41.

¹²¹ Susan Landau, *Control Use of Data to Protect Privacy*, 347 SCIENCE 506 (2015).

¹²² *Id.*; see also NATIONAL RESEARCH COUNCIL, BULK COLLECTION OF SIGNALS INTELLIGENCE (2015).

¹²³ Landau, *supra* note 121, at 506; see also Carl Ellison et al., *SPKI Certificate Theory*, RFC 2693 (Sept. 1999) (describing different examples of attribute credentials).

¹²⁴ Landau & Vargas Leon, *supra* note 119 at 236–237.

¹²⁵ PCAST REPORT, *supra* note 32, at 49.

¹²⁶ CATE, *supra* note 117, at 370.

¹²⁷ *Id.* at 371.

However, it is not always clear what those categories mean. “Per se harmful” to some extent depends on the definition of harm, and advertisers would have us believe that seeing ads tailored to our interests is beneficial.

Professor Landau and Dr. Vargas Leon have proposed a restricted list of uses, but their focus is on metadata and telemetry. Significantly, and likely because their work is post-Covid pandemic, they suggest permitting use for “during publicly declared public health emergencies, providing information on the movement of people in the aggregate; this latter use for a maximum of one week.”¹²⁸ This observation was based on the work done in privacy-preserving exposure notification.¹²⁹

Another important question is who should grant consent for a specific use. Having a government data protection agency do so is more compatible with European approaches to privacy; it may not pass muster in the U.S. The obvious answer, individual consent, is problematic: how are the individuals to be located? Data collected long ago can still be valuable, see, e.g., a study on Alzheimer’s incidence today correlated with results from a 1960 aptitude test.¹³⁰ Locating the test subjects was problematic; the researchers only found 38% of the test-takers¹³¹ and were able to achieve that much only because they were able to tap into high school 50th reunion data.¹³² Nor did the researchers seek individual consent; they instead obtained a waiver from their institutional review board,¹³³ a solution more applicable to academic research than to industry efforts.

Furthermore, and as noted, many privacy violations occur independent of the existence of any PII. Consent to assorted categories could, presumably, be set at collection time, perhaps by drawing on browser or device defaults, but there would be no way to change such consent in the future, whether for new uses, new abuses, or changed personal preferences. Auditing compliance with user restrictions is also a problem.

The third problem is the potential for later misuse of collected data. It is a truism in the privacy community that the most serious abuses happen when data collected for one purpose is then used for another. Serious abuses have happened, e.g., the misuse of census data to intern Japanese Americans during World War II.¹³⁴ There is the additional risk of hacking.¹³⁵ On the

¹²⁸ Landau & Vargas Leon, *supra* note 119, at 327.

¹²⁹ *See id.* at 312.

¹³⁰ Alison R. Huang, Kiersten L. Strombotne, Elizabeth Mokyr Horner & Susan J. Lapham, *Adolescent Cognitive Aptitudes and Later-in-Life Alzheimer Disease and Related Disorders*, JAMA NETWORK OPEN 1 (Sept. 7, 2018).

¹³¹ *Id.* at 2.

¹³² Tara Bahrapour, *In 1960, about a Half-Million Teens Took a Test. Now It Could Predict the Risk of Alzheimer’s Disease.*, WASH. POST (Sept. 21, 2018), https://www.washingtonpost.com/local/social-issues/in-1960-about-half-a-million-teens-took-a-test-now-it-could-predict-whether-they-get-alzheimers/2018/09/20/fcbabebe-b864-11e8-a7b5-adaaa5b2a57f_story.html [<https://perma.cc/E8BD-DYSY>].

¹³³ Huang et. al, *supra* note 130, at 2.

¹³⁴ *See* Lori Aratani, *Secret Use of Census Info Helped Send Japanese Americans to Internment Camps in WWII*, WASH. POST (Apr. 6, 2018), <https://www.washingtonpost.com/news/retropolis/wp/2018/04/03/secret-use-of-census-info-helped-send-japanese-americans-to-internment-camps-in-wwii/> [<https://perma.cc/M6S7-ZJV6>].

¹³⁵ Getting illicit access to vast troves of personal data does not require the skills of a national intelligence agency. Even amateurs have been able to do it; *see, e.g.*, Thomas Brewster, *How An Amateur Rap Crew Stole Surveillance Tech That Tracks Almost Every American*, FORBES (2019).

contrary, data that does not exist cannot be abused.¹³⁶

Finally, and perhaps fatally for proposals to enact statutory restrictions on data use, there may be a First Amendment issue. The Supreme Court has held that “an individual’s right to speak is implicated when information he or she possesses is subjected to ‘restraints on the way in which the information might be used’ or disseminated.”¹³⁷ The full ruling is more nuanced and may leave room for regulation. The Court noted that the basis for its holding was that:¹³⁸

On its face, Vermont’s law enacts content- and speaker-based restrictions on the sale, disclosure, and use of prescriber-identifying information. The provision first forbids sale subject to exceptions based in large part on the content of a purchaser’s speech. For example, those who wish to engage in certain “educational communications,” § 4631(e)(4), may purchase the information. The measure then bars any disclosure when recipient speakers will use the information for marketing. Finally, the provision’s second sentence prohibits pharmaceutical manufacturers from using the information for marketing. The statute thus disfavors marketing, that is, speech with a particular content. More than that, the statute disfavors specific speakers, namely pharmaceutical manufacturers.

In other words, the issue in *Sorrell* was not just the restriction on speech, but a restriction that applied to particular content and particular speakers. A law that blocked all users, not just a few, and all uses, not just some, might withstand scrutiny. The opinion noted that the HIPAA¹³⁹ did pass muster: “[f]or instance, the State might have advanced its asserted privacy interest by allowing the information’s sale or disclosure in only a few narrow and well-justified circumstances.”¹⁴⁰ Use controls that applied to everyone, not just marketers, and applied to all uses, including, e.g., research, would be neither content- nor speaker-based.

A recent ruling by the Ninth Circuit shows more constitutional risks.¹⁴¹ In 2022, California enacted its Age-Appropriate Design Code Act,¹⁴² which among other things required that certain platforms “are designed in a manner that recognizes the distinct needs of children.”¹⁴³ The court found that this “clearly [compelled] speech by requiring covered businesses to opine on potential harm to children.”¹⁴⁴ Similarly, the Supreme Court noted that “we have repeatedly held that laws curtailing their editorial choices must meet the First Amendment’s requirements.”¹⁴⁵ It is possible that a requirement to classify data uses would run afoul of these same principles.

Alternatively, a solution that turned on conditional consent would likely suffice: a user might voluntarily turn over data only to entities that promised to abide by certain restrictions.

¹³⁶ See Bellovin, *supra* note 67.

¹³⁷ *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 568 (2011).

¹³⁸ *Id.* at 563–64.

¹³⁹ Health Insurance Portability and Accountability Act of 1996, 42 U.S.C. § 1320d-2.

¹⁴⁰ *Sorrell*, 564 U.S. at 573.

¹⁴¹ *NetChoice v. Bonta*, No. 23-2969, 2024 WL 3838423 (9th Cir. 2024).

¹⁴² Cal. Civ. Code §§ 1798.99.28–1798.99.40 (West 2022).

¹⁴³ *NetChoice*, at *1 (internal citation omitted).

¹⁴⁴ *Id.* at *9. The court, citing *Interstate Cir., Inc. v. City of Dallas*, 390 U.S. 676, also noted that the Supreme Court “has previously applied First Amendment scrutiny to laws that deputize private actors.”

¹⁴⁵ *Moody v. NetChoice, LLC*, Nos. 22-277 & 22-555, slip op. at 2 (U.S. July 1, 2024).

Violating that promise would be seen as unfair and deceptive by the FTC.¹⁴⁶

Use controls do seem to avoid the insoluble problems of collection limits and notice and consent. An approach based on categories with user consent to selected categories, either at time of collection or later, might be feasible if embedded in a suitable legal framework.

B. Differential Privacy

Differential privacy is a comparatively recent innovation.¹⁴⁷ Its importance lies in the fact that it can give mathematical guarantees of privacy: one can calculate just how private a database is.¹⁴⁸ There is a cost, of course: there is a tradeoff between privacy and accuracy.¹⁴⁹ The more privacy-preserving a database must be, the less accurate its contents will be.¹⁵⁰ Still, the importance of its mathematical guarantee of privacy was sufficient to induce the U.S. Census Bureau to adopt it for the 2020 census,¹⁵¹ since the Bureau is required by law to protect the personal data of individual respondents.¹⁵² That said, the loss of accuracy was controversial, and Alabama and others sued. No decision on the merits was reached; the request for an injunction was denied due to lack of standing, unripeness, and other legal issues, and the suit was eventually voluntarily dismissed.¹⁵³

The idea behind differential privacy is that “noise” is added to the data. That is, numerical values are perturbed according to a mathematical formula. For example, if an individual’s annual salary is \$100,000, that number could be changed to \$95,000, \$105,000, etc. The *average salary* of everyone in the database would still be approximately correct, but any individual record would probably not be. Just how the noise to be added is selected is a complex matter and is at the heart of the accuracy/privacy tradeoff.

Using differential privacy as described is problematic for three reasons. The first is that people may not want to trust the party collecting the data. Perhaps they are storing the raw data, which is unprotected and hence subject to compromise and the temptation to misuse it. Trying to bar the latter by law runs into the same *Sorrell* issues as use control would, though with the same more nuanced analysis. While one can attempt to avoid some types of data collection, accurate responses to the Census are required by law.¹⁵⁴

¹⁴⁶ See generally Solove & Hartzog, *supra* note 85; Letter from James C. Miller III to Hon. John D. Dingell, reprinted in In re Cliffdale Associates, Inc., 103 F.T.C. 110, 174 (1984) [hereinafter FTC Policy Statement on Deception], available at <https://www.ftc.gov/legal-library/browse/ftc-policy-statement-deception> [<https://perma.cc/XV3K-SGBT>].

¹⁴⁷ See generally Cynthia Dwork et al., *Calibrating Noise to Sensitivity in Private Data Analysis*, in THEORY OF CRYPTOGRAPHY: THIRD THEORY OF CRYPTOGRAPHY CONF. 265 (2006).

¹⁴⁸ Alexandra Wood et al., *Differential Privacy: A Primer for a Non-Technical Audience*, 21 VAND. J. ENT. & TECH. L. 209, 212 (2018).

¹⁴⁹ See *id.* at 255.

¹⁵⁰ See *id.*

¹⁵¹ Ron Jarmin, *Census Bureau Adopts Cutting Edge Privacy Protections for 2020 Census*, U.S. CENSUS BUREAU (Feb. 15, 2019), https://www.census.gov/newsroom/blogs/random-samplings/2019/02/census_bureau_adopts.html [<https://perma.cc/56JS-H3XR>].

¹⁵² 13 U.S.C. § 9.

¹⁵³ *Alabama v. U.S. Dep’t of Commerce*, 546 F. Supp. 3d 1057, 1057–58 (M.D. Ala. 2021). There were also claims that use of differential privacy would selectively disadvantage minorities; an analysis by myself and others showed that this was not a problem. Miranda Christ et al., *Differential Privacy and Swapping: Examining De-Identification’s Impact on Minority Representation and Privacy Preservation in the U.S. Census*, IEEE SYMP. ON SEC. AND PRIV. (2022).

¹⁵⁴ 13 U.S.C. § 221.

The second issue is that not all types of collected data are amenable to protection by differential privacy. One can't easily perturb search query histories, nor a list of URLs visited. Finally, for technical reasons, database records with many fields cannot be protected that way. The Census Bureau has concluded that, at least at the moment, the American Community Survey cannot be protected via differential privacy.¹⁵⁵

While the latter two issues cannot yet be dealt with technically, the first one can via a technique known as “local differential privacy.”¹⁵⁶ In local differential privacy, the noise is added at the source—by the person submitting the data; the central database never sees the accurate values.

V. RECOMMENDATIONS

We recommend that the following steps be adopted.

- First and foremost, we need a new paradigm for privacy, one that goes beyond notice and consent.¹⁵⁷ One can hearken back to Warren and Brandeis' definition, “the right to be let alone”;¹⁵⁸ however, that is not an operational definition in the same sense as the more modern ones: it does not tell us what rules or restrictions should be imposed. This is a challenging research question. As noted, PCAST,¹⁵⁹ Professor Landau,¹⁶⁰ and others have proposed use restrictions. Professor Ari Ezra Waldman suggests an approach to privacy by design based on liability law;¹⁶¹ he notes, though, that he is “not suggesting a new products liability tort for privacy-invasive design through which individuals could sue technology companies and data collectors.”¹⁶² Other scholars have suggested privacy torts,¹⁶³ and of course that hearkens back to the Warren and Brandeis article.¹⁶⁴ We suggest that the first step would be a study by the National Academies of Science, Engineering, and Medicine; the charge to such a study committee would be to identify existing proposals, to evaluate their advantages, disadvantages, and feasibility, and to define the parameters of future projects aimed at crafting a new paradigm. In addition, funding agencies such as the National Science Foundation should be encouraged to sponsor research on the topic.
- We may have to revisit the definition of “privacy.” Although early definitions noted that purpose was part of the meaning of privacy,¹⁶⁵ that aspect has been elided from modern

¹⁵⁵ Donna Daily, *Disclosure Avoidance Protections for the American Community Survey*, RANDOM SAMPLINGS (Dec. 14, 2022), <https://www.census.gov/newsroom/blogs/random-samplings/2022/12/disclosure-avoidance-protections-acs.html> [https://perma.cc/W2YR-22QS] (“Our ongoing research has made it clear that the science for a formally private solution for the ACS does not yet exist. It's also not clear that differential privacy would ultimately be the best option”).

¹⁵⁶ See generally Graham Cormode et al., *Privacy at Scale: Local Differential Privacy in Practice*, INT'L CONF. ON MGMT. OF DATA 1655 (2018).

¹⁵⁷ Westin, *supra* note 2; HEW REPORT, *supra* note 13.

¹⁵⁸ Samuel D. Warren & Louis Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 195 (1890).

¹⁵⁹ PCAST REPORT, *supra* note 32, at xii.

¹⁶⁰ Landau, *supra* note 121.

¹⁶¹ Ari Ezra Waldman, *Privacy's Law of Design*, 9 U.C. IRVINE L. REV. 1239, 1240 (2019).

¹⁶² *Id.* at 1244–45.

¹⁶³ *Id.* at 1263.

¹⁶⁴ Warren & Brandeis, *supra* note 158.

¹⁶⁵ HEW REPORT, *supra* note 13, at 61 (“Assure that no use of individually identifiable data is made that is not within the stated purposes

definitions.¹⁶⁶

- We desperately need a more modern definition of “injury” or “harm.” Limiting it to direct financial harm, e.g., through theft of credit card numbers, or to disclosure of health information, is inadequate. For example, simple disclosure of group membership can be very dangerous. In one incident, confusion over Facebook’s privacy policies led to some students being outed as gay to their parents, with devastating effects on their family ties.¹⁶⁷ Even for those who advocate less regulation, more clarity here would be useful; it would lend more predictability to the FTC’s enforcement actions.¹⁶⁸ The authority of the FTC to act against security and privacy breaches should be enhanced by statute. As some have pointed out:¹⁶⁹

[T]he FTC lacks the general authority to issue civil penalties and rarely fines companies for privacy-related violations under privacy-related statutes or rules that provide for civil penalties When the FTC does include fines, they are often quite small in relation to the gravity of the violations and the overall net profit of the violators. This is because any fines issued by the FTC must reflect the amount of consumer loss.

- Barring a new paradigm, use controls seem to be a promising approach, assuming that the constitutional issues can be resolved. Defining categories and consent mechanisms¹⁷⁰ is an open question, though arguably less daunting than a completely new paradigm. A study committee, possibly under the auspices of the FTC, should define a set of categories; data users would assign their effort to a particular category.¹⁷¹ Misassignment or misleading users would be a deceptive practice, per the FTC Act.¹⁷² It is crucial that user permissions are independent of each site. That is, what is protected is the data, not the source from which it came; to do otherwise is to fall back into the same traps as today’s notice and consent. Someone who wishes to opt out of, say, email marketing pitches should only have to do it once, or at most once per email address used. This might require a central registry of preferences, but we already use such for the “Do Not Call” list.¹⁷³ Local differential privacy is also helpful, when applicable.

- Once a new paradigm is selected, be it use restrictions or something else, privacy policies based on notice and consent should be phased out as swiftly as feasible. Although the GDPR may

of the system as reasonably understood by the individual”).

¹⁶⁶ Shirey, *supra* note 17.

¹⁶⁷ Geoffrey A. Fowler, *Watched: When the Most Personal Secrets Get Outed on Facebook*, WALL ST. J. (Oct. 3, 2012), <https://www.wsj.com/articles/SB10000872396390444165804578008740578200224> [<https://perma.cc/U2D6-Y3RW>] (“The two students were casualties of a privacy loophole on Facebook—the fact that anyone can be added to a group by a friend without their approval. As a result, the two lost control over their secrets, even though both were sophisticated users who had attempted to use Facebook’s privacy settings to shield some of their activities from their parents.”).

¹⁶⁸ The FTC has looked into the question, but there is no clear consensus so far. *See* FED. TRADE COMM’N, FTC INFORMATIONAL INJURY WORKSHOP (2018).

¹⁶⁹ Solove & Hartzog, *supra* note 85, at 605.

¹⁷⁰ *See* Section IV.

¹⁷¹ Ironically, there is the potential for privacy violations from this very practice intended to preserve privacy. Suppose there are 24 different use categories. That means there are 2²⁴ (about 16 million) different combinations of settings. Unusual-enough choices would themselves constitute a tracking mechanism. For related work on the privacy implications of too many possibilities, *see generally* Peter Eckersley, *How Unique Is Your Web Browser?*, 10 INT’L SYMP. ON PRIV. ENHANCING TECH. 1 (2010).

¹⁷² Federal Trade Commission Act, 15 U.S.C. § 45 (2012).

¹⁷³ 16 C.F.R. § 310.4(b)(iii)(B) (2024).

still require it, many websites have dual policies already. This is to comply with the GDPR when they must but to take advantage of looser American regulations when they can.

- Finally, if we have to stick with notice and consent, two major changes should be made. First, it should be mandatory to disclose privacy practices in a simple, standardized format, akin to nutrition labels on food. Research suggests that this approach is very promising.¹⁷⁴ Kelley et al. found that concise, standardized formats worked best.¹⁷⁵

We have shown here that it is not solely the table-based format, but holistic standardization that leads to success. Our standardized short-text policy left no room for erroneous, wavering, or unclear text, serving as a concise textual alternative to tabular formats.

While the standardized short text policy we developed was successful for most tasks, it is not as easy to scan as a table. Indeed, one participant suggested policies could be improved if they were set up “like a chart so you can scan it visually for answers instead of having to take the time to read it.”

They noted that “[t]he standardized formats performed the best overall, across the variety of the metrics [they] looked at. The accuracy, comparison, and speed results eclipse the results of the text formats in use today.”¹⁷⁶

Second, site operators’ privacy policies must disclose the policies used by any embedded sites. The site operator can, at least in principle, control this; the user cannot. Note that these two notions are linked: a simple-to-read privacy policy can easily be encoded in machine-readable form. This can be passed to advertisers, and they would have to comply with it.

The most important thing, though, is to act and to act now. Every day, more data is collected; every day, more abuses and leaks take place.

¹⁷⁴ See, e.g., Patrick Gage Kelley et al., *Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach*, SIGCHI CONF. ON HUM. FACTORS IN COMPUTING SYS. 1573 (2010); Patrick Gage Kelley et al., *A “Nutrition Label” for Privacy*, 5 SYMP. ON USABLE PRIV. AND SEC. 1 (2009).

¹⁷⁵ Patrick Gage Kelley et al., *Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach*, SIGCHI CONF. ON HUM. FACTORS IN COMPUTING SYS. 1573, 1580 (2010).

¹⁷⁶ *Id.* at 1581.