Santa Clara High Technology Law Journal

Volume 35 | Issue 1 Article 1

11-1-2018

Model(ing) Privacy: Empirical Approaches to Privacy Law and Governance

Barrett, Lindsey

Model(ing) Privacy: Empirical Approaches to Privacy Law & Governance

By Lindsey Barrett

Privacy can be difficult for people to conceptualize, including for the policymakers charged with designing, interpreting, and enforcing privacy law. In both consumer privacy law and Fourth Amendment jurisprudence, the privacy protections afforded to individuals are shaped by the ability of governmental decision-makers to assess privacy preferences, expectations, and behaviors, which they are rarely in a position to do accurately. While policymakers can have a hard time understanding the subtle factors influencing privacy decision-making or parsing seemingly contradictory privacy incentives, it is an area where new empirical approaches have begun to excel. Researchers have used empirical techniques like machine learning, natural language processing, and crowdsourcing to explain the complexities of privacy decision-making, and to illustrate the nuances of privacy preferences, expectations, and behaviors that many opinion surveys often fail to grasp. Recent work has focused on eliciting privacy norms through crowdsourcing, modeling individual privacy preferences and expectations using machine learning, extracting key terms from privacy policies through natural language processing, and modeling AI assistants based on context and user preferences to predict (or nudge) future decisions. Modeling privacy preferences, expectations and behavior can provide judges, regulators, and legislators with a more accurate and nuanced sense of privacy norms for future cases and policy discussions. Encouraging the implementation of proactive privacy tools, such as automated annotation of privacy policies and nudging assistants, can help bridge the gap separating user expectations, user behavior, and how both are understood under existing laws. While the use of this research in privacy law and policy cannot fundamentally transform the structural flaws that skew regulators' perceptions of societal norms, it can at least

Teaching Fellow & Staff Attorney, Communications & Technology Clinic, Institute for Public Representation, Georgetown University Law Center. I'm enormously grateful to everyone who took the time to provide me with invaluable feedback through various drafts, particularly Paul Ohm, Jonathan Frankle, Robert Walker, Joe Jerome, Andrew Selbst, participants at the 2018 Privacy Law Scholars Conference and participants at the Amsterdam Privacy Conference 2018.

correct the worst of those excesses, and facilitate policy that reflects how people actually think about privacy in the modern age.

CONTENTS

INTRODUCTION

"The makers of our Constitution . . . sought to protect Americans in their beliefs, their thoughts, their emotions and their sensations. They conferred, as against the Government, the right to be let alone – the most comprehensive of rights, and the right most valued by civilized men."

- *Justice Louis Brandeis, dissenting,* Olmstead v. United States[1]

"Your user agreement sucks."

[1] Olmstead v. United States, 277 U.S. 438, 478 (1928) (Brandeis, J., dissenting).

- *Senator John Kennedy to Mark Zuckerberg, Senate Judiciary Hearing on Facebook, Social Media, Privacy, and the Use and Abuse of Data*[2]

People care about privacy for different reasons, and to differing extents. As the volume of data sets about each of us continues to proliferate, and the uses of that information continue to evolve, gauging individual privacy expectations and broader societal norms has become increasingly challenging. Individuals make privacy decisions that seem to undermine their stated preferences, even as the risks to the fundamental interests linked to privacy, such as equality, autonomy, and intellectual freedom, only continue to grow. Largely to blame for these apparent contradictions are the ineffective standards that determine how privacy decision-making, expectations, and preferences are measured. The regulatory regime governing consumer privacy and the Fourth Amendment's protections for privacy from the government both rest on the idea that judges and policymakers can discern individual and collective privacy norms, when in reality, they are rarely able to do so accurately.

Consumer privacy law in the United States is molded around the idea of privacy as an economic good, where the degree of legal protection a person receives depends on her control over her information through notice and choice mechanisms, like app permissions or privacy policies. The notice and choice model relies on the idea that informing consumers in convoluted boilerplate of how their data is collected and used empowers them to make privacy decisions that reflect their preferences. Under this thinking, any failure to subsequently make privacy-protective choices indicates either apathy or a deliberate declaration of a contrary preference. In fact, it is exceedingly difficult for individuals to make choices that produce the privacy outcomes they prefer or expect due to cognitive and structural limitations. Phenomena like decision fatigue, learned helplessness, and lack of information about collection and tracking all impede individuals from making the privacy choices that correspond to what they hope (or believe) will happen to their information.

Many technologies have become so intertwined with daily life that even individuals with strong privacy preferences or expectations cannot make choices that suit those preferences. A person might want to avoid geolocational tracking but need to carry a cell phone to ensure an elderly parent can get in touch; someone else might wish to avoid

[2] *Transcript of Mark Zuckerberg's Senate Hearing*, WASH. POST (Apr. 10, 2018) https://perma.cc/Y7E3-PN5P.

web tracking but be required to use a school or employer-run Gmail account. In a memorable example, journalist Julia Angwin documented her attempts to avoid every form of surveillance she could by using burner phones whenever possible, abstaining from using any Google services, relying on a credit card under a fake name when she couldn't use cash, and carrying her smartphone in a makeshift Faraday bag to block the phone from sending or receiving signals.[3] She assessed her diligent efforts to avoid being tracked as "50% successful" – and this is a tech-savvy investigative journalist, who was solely dedicated to the task of protecting her privacy.[4] Others may have stronger privacy preferences or expectations repudiated by the actions of others that are beyond their control. In a recent and infamous example, Facebook provided the information of 87 million users to the political firm Cambridge Analytica, which then coordinated with the Trump campaign to target voters based on that data.[5] The firm acquired the information after just 30,000 users downloaded a quiz app, and Facebook's developer guidelines allowed the firm to access the data from every Facebook friend the quiz app users had. While users consent to a lot of things when they create a Facebook account, it is difficult to argue that the 86,970,000 users who did not download the app themselves expected their information to be used for voter targeting by a presidential campaign (or even that the 30,000 who did download the app would expect that result). Data breach after data breach further demonstrates that while companies may promise to protect their users' data in their privacy policy, they consistently fail to do so.[6] Chief Justice Roberts noted in *Carpenter v. United States,* this country of 326 million individuals is a country of 396 million cell phone accounts, and only the few without phones can escape such "tireless and absolute surveillance."[7]

Despite the enormous barriers to making choices that cohere with an individual's privacy expectations or preferences, the failure of individuals to make privacy-protective decisions is repeatedly declared

[3] Jacob Silverman, *'Dragnet Nation' Looks at the Hidden Systems that Are Always Looking at You*, L.A. TIMES (Mar. 6, 2014, 12:00 PM), https://perma.cc/3J4C-HQUS.

[4] Andrew Leonard, *Is Privacy Really Dead? Julia Angwin and the Quest to Escape Big Brother,* SALON (Mar. 2, 2014, 9:00 PM), https://perma.cc/UB7Y-GDDH.

[5] Kevin Granville, *Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens*, N.Y.T. (Mar. 19, 2018), https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html.

[6] 2017 saw a record high of 1,579 breaches, a 44.7% increase from the year before. More than half of these breaches involved Social Security numbers, with a total of nearly 158 million exposed. *2017 Annual Data Breach Year-End Review,* IDENTITY THEFT RESOURCE CTR., https://perma.cc/Y5C8-TKEQ (last visited July 30, 2018).

[7] Carpenter v. United States, 585 U.S. ___ , No. 16-402, slip op. at 1, 14 (U.S. June 22, 2018).

to be evidence that people do not care about their privacy, or to the extent that they do care, they do not care enough. In addition to leaving individual privacy at the mercy of broken mechanisms, the inefficacy of notice and choice warps privacy narratives in public policy by allowing non-privacy-protective behavior to be attributed to a conscious rejection by the marketplace. This lets a failed system masquerade as evidence of democratic consensus on permissive regulation. Motivated analysis of opinion surveys that measure privacy preferences, expectations, and behavior in an a contextual and leading manner further bolsters broad claims that most individuals care less about their privacy than they do about receiving goods and services for it, and that regulators should respond accordingly.[8] In a country where consumer privacy is considered a good to trade away and where consumer privacy is constantly juxtaposed against the primacy of American innovation, strong legal protections are easy to portray as stilted, counterproductive, and even undemocratic.[9] As regulators and legislators attempt to craft policy that reflects the will of the people they serve, this hijacked narrative continues to provide support for privacy governance that fails to protect privacy, including claims that permissive laws or self-regulation are the result that the public, rather than industry, truly wants.

In criminal law, the Fourth Amendment at least sets a textual threshold of a right to privacy. But the enforcement of that right also depends on a judge's understanding of what a reasonable expectation of privacy is,[10] and her application of common-law doctrines to new technological contexts. Here, too, misunderstanding of privacy preferences, expectations and behaviors leads to weaker protections for privacy, as a judge's perception of what a reasonable expectation of privacy is tends to differ from that of the average person.[11] Furthermore, as Justice Alito put it, "judges are apt to confuse their own expectations of privacy with those of the hypothetical reasonable person to which the *Katz* test looks."[12] Demographics may also play a part. Judges are older, whiter, more likely to be male, and better

[8] *See, e.g.*, Association of National Advertisers, Comment Letter on Protecting the Privacy of Customers of Broadband and Other Telecommunications Services, at 11 (May 31, 2016), https://perma.cc/YS3Y-2TVV (hereinafter, ANA Comments).

[9] *See* Paul M. Schwartz & Karl-Nikolaus Peifer, *Transatlantic Data Privacy Law*, 106 GEO L. J. 115, 132 (2017) (describing the "marketplace discourse" of privacy in the United States).

[10] Katz v. United States, 389 U.S. 347, 361 (1967) (Harlan, J., concurring).

[11] *See* Christopher Slobogin & Joseph E. Schumacher, *Reasonable Expectations of Privacy and Autonomy in Fourth Amendment Cases: An Empirical Look at "Understandings Recognized and Permitted by Society"*, 42 DUKE L. J. 727, 732 (1993).

[12] United States v. Jones, 132 S. Ct. 945, 962 (2012) (Alito, J., concurring).

educated than is the average defendant.[13] The judiciary is also more likely to be aware of technological risks,[14] less likely to have had personal interactions with the criminal justice system,[15] and likely to overestimate the average person's understanding of how privacy and surveillance works.[16] Whatever the combination of factors, judges tend to believe that the public holds a lower expectation of privacy than the public itself reports.[17]

One example of the space between the judge's reasonable expectation of privacy and that of the average person is the recent landmark privacy case, *Carpenter v. United States*.[18] The Supreme Court held that the government must obtain a warrant before obtaining seven or more days of cell site location information (CSLI) from a third-party service provider, in lieu of extending the third-party doctrine's principle that the defendant had assumed the risk (and extinguished any reasonable expectation of privacy) that law enforcement might obtain the information by entrusting it to the company.[19] But the logic of the holding depends on both the duration of the tracking, and the fact that the data collected was geolocational.[20] Other kinds of information transmitted through third parties – which, in the modern age, is nearly all the information we interact with – may still be accessed without a warrant. While the decision is an important

[13] *See* Bernard Chao, Catherine S. Durso, Ian P. Farrell & Christopher T. Robertson*, Why Courts Fail to Protect Privacy: Race, Age, Bias, and Technology*, 106 CAL. L. REV. 263, 290 (2018) (noting that the judiciary tends to be more male, white, affluent and educated than ordinary members of the public); Matthew Tokson, *Knowledge and Fourth Amendment Privacy*, 111 NW. U. L. REV. 139, 170 (2016).

[14] *See* Tokson, *supra* note 13, at 169-70 (observing that knowledge about technology tends to reach more highly educated people, and that judges are both more highly educated than average, and more likely to be better acquainted with both government surveillance and criminal procedure).

[15] *See* Chao et al., *supra* note 13, at 289-291 (discussing the prior literature on how demographics impact judicial perceptions, noting that the judiciary skews more male, white, educated, and older than the American population at large, and the relatively high rates of imprisonment and lower representation on the state and federal bench for African-Americans and Hispanics).

[16] Tokson, *supra* note 13, at 172 ("[J]udges are well-informed socioeconomic elites who are likely to systematically overestimate societal knowledge. Societal knowledge tends to be counterintuitively low, and tends to spread more quickly to elites than to the average citizen.") (citations omitted). *See, e.g., id.* at 174 (distinguishing the high level of understanding courts consider the average person to have of surveillance and web tracking from the average person's actual understanding of those technologies, which tends to be limited).

[17] Slobogin & Schumacher, *supra* note 11, at 731 (describing the evolution of the "societal understanding" of privacy norms as a function of the *Katz* test).

[18] Carpenter v. United States, 585 U.S. ___ , No. 16-402, slip op. at 12-13 (U.S. June 22, 2018).

[19] *Id.* at 11.

[20] *Id.* at 18 ("We do not disturb the application of *Smith* and *Miller* or call into question conventional surveillance techniques and tools, such as security cameras. Nor do we address other business records that might incidentally reveal location information.").

step towards updating the Fourth Amendment for the digital age, it's unlikely this distinction aligns with how most people think about the sensitivity of their information. In one study, 90% of participants believed that police should have to at least obtain a warrant before accessing the email addresses they had corresponded with, whereas only 55-60% believed that a warrant should be required for their geolocation information, and few varied in their answer according to the duration of the geolocation tracking.[21] *Carpenter's* protection for information relayed through third parties rests on factors that people care less about, as opposed to the kinds of information (like email addresses) left unprotected by the decision.

While *Carpenter* provided a rare glimmer of hope for those who wish to see a Fourth Amendment that fully grapples with the breadth and depth of technological change, most judicial attempts to apply analog doctrines to digital technology frequently ignore the extent to which the new context undermines the logic underlying older doctrines. Areas of Fourth Amendment jurisprudence like the third-party doctrine, the public view doctrine, and the content/non-content distinction insufficiently consider the role of context in privacy decision-making, and have been rendered inapt by technological realities.[22] For example, the majority's declination to extend the third-party doctrine to CSLI rested on the particular sensitivity of geolocation information, given that cellphones accompany their owners everywhere, in contrast with a car.[23] But modern vehicles come equipped with GPS and the ability to sync with the driver's phone; some even receive their connectivity from the same cell towers that collect CSLI, all while being subject to warrantless access by law enforcement under the Fourth Amendment's automobile exception.[24] The result is a privacy right tethered to the reasonable judge's expectation of privacy rather than that of the reasonable person, and doctrines that fail to provide the Constitutional guarantees of privacy they were designed to give.

This article does not intend to imply that consumer and criminal privacy law work in the same way, or are hobbled by identical deficiencies. The principles, objectives, and constraints of consumer

[21] Christine S. Scott-Hayward, Henry F. Fradella & Ryan G. Fischer, *Does Privacy Require Secrecy? Societal Expectations of Privacy in the Digital Age*, 43 AM. J. CRIM. L. 20, 52-54 (2015).

[22] *See infra* Part II(b)(ii).

[23] *Carpenter*, slip op. at 13.

[24] Lindsey Barrett, *Herbie, Fully Downloaded: Data-Driven Vehicles and the Automobile Exception*, 106 GEO L. J. 182, 185-187 (2017) (describing the tracking and collection capacities of connected cars and arguing that the privacy interest in the information they collect merits the use of a warrant).

privacy law and Fourth Amendment law are different, as are the institutional competencies of the governmental decision-makers creating and applying the law. Consumer privacy regulators must be able to accurately deduce privacy norms and apply them to legal standards, as a judge does. But they also use their understanding of norms to determine the fundamental level of protection consumers will receive from proactive policy measures, whether through enforcement actions, regulation, public education, or incentive programs. Correctly assessing privacy norms in a market-based privacy model is crucial, as there is no theoretical floor of what protections consumer privacy should receive. A narrative informed by a mistaken faith in economically rational privacy decision-making or misleading opinion surveys can impact how legislators and regulators view their obligations to the public.

In comparison, privacy in criminal law as guaranteed by the Fourth Amendment does have a fundamental floor, namely the Constitutional right the Framers provided for privacy.[25] But that too is ensured by a judge's understanding of societal expectations of privacy, when their experiences and perspectives rarely reflect those of the average person. Judges applying the law based on an accurate understanding of societal privacy norms is a fundamental component of ensuring that the Fourth Amendment's "reasonable expectation of privacy" reflects what the average reasonable person would expect.[26] Moreover, consumer privacy and Fourth Amendment privacy have become increasingly interdependent as the vast amounts of data that companies collect and store become the vast amounts of evidence the government can access. The myth of the rational privacy decision-maker in consumer privacy is also highly relevant to many Fourth Amendment doctrines.

In both areas of the law, the amount of privacy the law guarantees depends in large measure on how a person makes privacy decisions, given the information society expects them to have.[27] And while the failure of notice and choice as an effective data control mechanism facilitates companies' ability to collect data that can be accessed by law enforcement, permissive privacy regulation gives them little reason to

[25] U.S. CONST. amend. IV.

[26] Katz v. United States, 389 U.S. 347, 361 (1967) (Harlan, J., concurring).

[27] *See, e.g.*, Transcript of Oral Argument at 16, *Carpenter*, 585 U.S. ___ (No. 16-402) (2018) (Alito, J.) ("Well, I mean, that's a debatable empirical point whether people realize [that the cell phones companies are storing records of their movements], and there's reason to think maybe they do . . . . The contract, the standard MetroPCS contract seems to … advise the customer that we can disclose this information to the government if we get a court order. So I don't know whether that will hold up.").

restrain themselves. The technologies that are primarily regulated by consumer law are the instrumentalities of the surveillance the Fourth Amendment protects against. This article discusses both areas of the law to highlight those interdependencies and because the research discussed *infra* can be a valuable tool for judges and regulators alike in understanding the privacy norms that they often struggle to accurately interpret.

Computer science researchers have begun to attack the problem of understanding privacy norms with research that empirically measures privacy behavior and preferences, and which more heavily focuses on the role of context and other non-normative factors on privacy decision-making, such as by deducing preferences through modeling user behavior, or by training predictive models on crowdsourced answers to context-focused privacy vignettes. Empirical evidence on privacy norms already plays a role in shaping consumer privacy policy at the agency level, largely through opinion surveys. But the role of empirical evidence in privacy policy and judicial decision-making can be expanded and improved to better measure how people feel about their privacy, and why. In consumer privacy, where legal protections hinge on individuals' control of their information and regulators' perception of the norms their actions create, it is crucial to parse privacy expectations, preferences, and behavior, and to understand the barriers hindering consumers from making privacy decisions that reflect their preferences and expectations. Regulators would benefit from analyzing and incentivizing research that acknowledges the impact of those barriers and the role of context in privacy decision-making, particularly in lieu of broadly framed opinion surveys that ignore the existence of either. In Fourth Amendment law, the use of empirical research to guide governmental decision-making is much less common; empirical research on privacy norms could help correct judges' erroneous understanding of what a reasonable expectation of privacy is, and to consider how blunt precedents should be applied to new technology.

Though deeply flawed, the existing legal standards governing privacy in the United States are unlikely to change in a fundamental way anytime soon. The reasonable expectation of privacy test is a foundational pillar of Fourth Amendment doctrine, and notice and choice is equally inextricable from consumer privacy law. But as long as the degree of privacy protection a consumer is afforded depends primarily on engaging with notice and choice mechanisms, regulators need to understand how and why privacy decision-making fails to represent the outcomes consumers intend. As long as the Fourth Amendment rests on a judge's perception of societal norms and applying analog precedent to technologies that undermine the principles upon which they were based, judges need a more accurate understanding of the average person's reasonable expectation of privacy, and the extent to which technology shifts the applicability of older precedents to the modern day. A more empirically accurate approach to how those standards measures privacy norms could help straighten misguided narratives, restore weakened protections, and update basic policy assumptions that have grown obsolete.

This article will proceed in five parts. The first part will provide an overview of consumer and criminal privacy laws and the respective deficiencies of each that have resulted in misinterpretation of privacy preferences, expectations, behaviors, and norms by judges and regulators. The second will address the role of context in privacy decision-making and in Fourth Amendment doctrine. The third will provide an overview on relevant empirical privacy research, and discuss new research that uses machine learning, crowdsourcing, and natural language processing to provide an empirical basis for understanding privacy preferences, expectations, behaviors, and norms, as well as methods of operationalizing user preferences and expectations to make privacy decision-making more coherent. The fourth part will address the applications of that research in consumer and criminal law, recognizing the institutional competences and limitations of different branches of government in how empirical evidence may be used. The fifth part will address additional considerations, and the sixth will conclude.

I. PRIVACY NORMS & THE LAW

A. Consumer Privacy

As the United States lacks an omnibus privacy law, consumer privacy is protected through a fractal array of sector-specific statutes at the state and federal levels.[28] The Health Insurance Portability and Accountability Act (HIPAA) protects the privacy of health information;[29] the Genetic Information Nondiscrimination Act (GINA) protects the privacy of genetic information;[30] the Family Education Rights and Privacy Act (FERPA) covers educational information;[31]

[28] *See* Deirdre K. Mulligan & Jennifer King, *Bridging the Gap Between Privacy and Design*, 14 U. PA. J. CONST. L. 989, 990 nn.1-5 (2012) (detailing sector-specific privacy statutes); Lorrie Faith Cranor, *Necessary but Not Sufficient: Standardized Mechanisms for Notice and Choice*, 10 J. ON TELECOMM. & HIGH TECH. L. 273, 277 (2012).

[29] 45 C.F.R. § 160 (2017).

[30] 42 U.S.C. § 2000ff-5 (2012).

[31] 20 U.S.C. § 1232g (2012).

children's privacy is protected by the Children's Online Privacy Protection Act (COPPA);[32] financial privacy is protected under the Graham-Leach-Bliley Act (GLBA)[33] and the Fair Credit Reporting Act (FCRA);[34] the Privacy Act of 1974 provides protections for information held by the federal government.[35] Other laws, prompted by sporadic public uproar over individual events, protect varied forms of information like video rental records under the Video Privacy Protection Act (VPPA),[36] or state driver's licenses under the Driver's Privacy Protection Act (DPPA).[37]

Each of these is heavily influenced or directly predicated on the Fair Information Practice Principles (FIPPs), a framework of privacy governance principles that became the global touchstone of privacy regulation. First mentioned in a report from the U.S. Department of Health and Welfare, the FIPPs became more influential after being formalized in a report from the Organization for Economic Cooperation and Development (OECD) in 1980.[38] The FIPPs, particularly the collection limitation principle and the use limitation principle, emphasize the informational control of the user and spurred the development of notice and choice as a bedrock privacy safeguard in privacy law and policy.[39] Federal and state privacy statutes allow different treatment of information based on whether the data-collecting entity obtains consent.[40] The Federal Trade Commission (FTC), the United States' *de facto* data protection agency, has applied its deception authority under the FTC Act to privacy policies and promises by bringing enforcement actions against companies that obtain consent through a policy that is contrary or silent as to the actual practices of

[32] 15 U.S.C. § 6502 (2012).

[33] 15 U.S.C. § 6801 (2012).

[34] 15 U.S.C. § 1681 (2012).

[35] 5 U.S.C. § 552a (2012).

[36] 18 U.S.C. § 2710 (2012).

[37] 18 U.S.C. § 2721 (2012).

[38] ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT, OECD GUIDELINES ON THE PROTECTION OF PRIVACY AND TRANSBORDER FLOWS OF PERSONAL DATA (1980).

[39] *See* Rebecca Balebako, Cristian Bravo-Lillo & Lorrie Faith Cranor, *Is Notice Enough: Mitigating the Risks of Smartphone Data Sharing*, 11 I/S: J. L. & POL'Y FOR INFO. SOC'Y 279, 281-82 (2015) (describing the development of the Fair Information Practice Principles, and noting that only one, Integrity/Security, focuses solely on the data collector, rather than the information control of the subject); CHRISTOPHER HOOFNAGLE, FEDERAL TRADE COMMISSION PRIVACY LAW AND POLICY 152-53 (2016) (describing the FIPs as the "basis for virtually all information privacy regulation").

[40] Daniel J. Solove, *Privacy Self-Management and The Consent Dilemma*, 126 HARV. L. REV. 1880, 1880 (2013) ("Consent legitimizes nearly any form of collection, use, or disclosure of personal data.").

the company.[41] Notice and choice, particularly through privacy policies, is broadly understood to be the fundamental basis of the agency's privacy philosophy and an integral component of its policy toolkit.[42]

The primacy of notice and choice in consumer privacy law makes individual privacy decision-making the bellwether of privacy protection in the consumer sphere. How a user understands the privacy decisions she make, and a regulator's interpretation of those motivations is fundamental. The idea that providing an individual with notice of a company's practices, and choice over whether or not to use the service after being provided with information about those practices, is the foundational safeguard that a data-collecting entity should provide.[43] But while the notion of privacy as informational control is the foundation establishing consumer privacy law, it is as ineffective as it would be difficult to replace.[44] Critics have long challenged FIPPs on the basis that notice and choice merely confers an illusion of control, and that it reduces privacy compliance to check-the-box formalism with no real consumer guarantees.[45] The ever-expanding network of Internet of things (IoT) devices also makes notice and choice increasingly less workable from a practical standpoint.[46] How and

[41] Aleecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, 4 I/S: J. L. & POL'Y FOR INFO. SOC'Y 543, 545-48 (2008) (describing the development of the FTC's approach to privacy policies and enforcing unfair and deceptive practices); Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 600 (2014) (describing the FTC's entry into privacy enforcement).

[42] FED. TRADE COMM'N, PRIVACY ONLINE: A REPORT TO CONGRESS 7 (June 1998) (describing notice as the "fundamental" basis of the FTC's privacy approach). *See also* Solove & Hartzog, *supra* note 41, at 634 (describing notice and choice as the "central" aspect of the agency's work). *Cf.* Joel Reidenberg, N. Cameron Russel, Alexander J. Callen, Sophia Qasir & Thomas B. Norton, *Privacy Harms and the Effectiveness of the Notice and Choice Framework*, 11 I/S: J. L. & POL'Y FOR INFO SOC'Y 485, 491 (describing the failures of the notice and choice mechanisms).

[43] Woodrow Hartzog, *The Inadequate, Invaluable Fair Information Practices*, 76 MD. L. REV. 952, 952-53 (2017).

[44] *Id.* at 953-54; Solove, *supra* note 40, at 1880-82 (describing the cognitive and structural limitations hampering the efficacy of notice and choice as a method of effective privacy management, and noting that the solution that sufficiently grapples with the consent-based model remains "elusive").

[45] *See* Hartzog, *supra* note 43, at 953 n.1 (detailing critiques of the FIPs since their inception); Julie Cohen, *Turning Privacy Inside Out*, 20.1 THEORETICAL INQUIRIES IN LAW (forthcoming 2019) ("notice-and-consent protections, which function as the principal regulatory tool in the U.S. system and as an increasingly important backstop in the European system, simply do not work.").

[46] *See* Hosub Lee & Alfred Kobsa, *Privacy Preference Modeling and Prediction in a Simulated Campuswide IoT Environment*, 2017 INST. ELECTRICAL & ELECRONICS ENGINEERS INT'L CONF. ON PERVASIVE COMPUTING & COMM. 276, 276 (2017) (describing the difficulty of designing effective notice and choice mechanisms for IOT devices); Hartzog, *supra* note 43, at 953 (Hartzog describes the evolving risks that have made the FIPPs increasingly obsolete, including that "[t]he mass connectivity of the 'Internet of Things' and near ubiquity of mobile devices make the security and surveillance risks presented by the isolated computer terminals and random CCTV

when can a device provide substantive information about the company's data practices when the user is driving a connected car or if the interface of the device is too small to easily read? Other IoT devices with outward-facing sensors may also have to grapple with novel third-party consent problems.[47] Notice and choice has been the whipping boy of privacy scholars and advocates for a reason – it has not worked, and it will only continue to fail consumers as new forms of interactive, networked technology advance and spread.

1. Framing Privacy

Not only does notice and choice fail to provide consumers with a reliable way to manage their information, but the consumers' inability to meaningfully engage with privacy policies and permissions settings is one of the most frequently cited indications that individuals care very little about their privacy.[48] It is the symptom that is not only killing the patient, but also mistakenly convinces the doctor that the condition is benign rather than severe, compounding the damage the symptom itself inflicts. The prevalence of individual privacy behavior that is inconsistent with expressed preferences has given rise to the so-called "privacy paradox," which concludes that an individual's behavior that is less privacy-protective than their expressed preferences reveals a true preference against privacy in favor of other values, such as convenience, efficiency, or economic gain.[49] While the privacy

cameras of the '80s and '90s seem quaint."); Pardis Emami-Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor & Norman Sadeh, *Privacy Expectations and Preferences in an IoT World***,** 2017 SYMP. ON USABLE PRIVACY & SECURITY 399, 400 (noting the additional challenges that IOT devices pose for the notice and choice model, such as "obtaining consent for data collection, allowing users to control, customize, and choose the data they share, and ensuring the use of collected data is limited to the stated purpose").

[47] *See, e.g.*, Cara Bloom, Joshua Tran, Javed Ramjohn & Lujo Bauer, *Self-Driving Cars and Data Collection: Privacy Perceptions of Networked Autonomous Vehicles*, 2017 SYMP. ON USABLE PRIVACY & SECURITY 357, 361 (noting that few legal protections apply to the collection of data belonging to third parties, such as data collected by automotive sensors).

[48] *See, e.g.*, Kirsten Martin & Helen Nissenbaum, *Measuring Privacy: An Empirical Test Using Context to Expose Confounding Variables*, 18 COLUM. SCI. & TECH. L. REV. 176, 180 (2016) (describing the conflict between reported privacy preferences and privacy behavior, and the so-called "privacy paradox"); Caleb S. Fuller, *The Perils of Privacy Regulation,* 30 REV. AUSTRIAN ECON. 193, 197 (2017) ("Third, some appear puzzled by the lack of privacy protection that contracts provide. Yet, might the absence of such protection be evidence that consumers do not value it highly?").

[49] Idris Adjerid, Eyal Peer & Alessandro Acquisti, *Beyond the Privacy Paradox: Objective versus Relative Risk in Privacy Decision Making*, 42 MGMT. INFO. SYS. Q. 465, 469-470 (2018) (defining the privacy paradox and outlining the accompanying literature); HELEN NISSENBAUM, PRIVACY IN CONTEXT TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE 104-5 (2010) (describing the privacy paradox). *See, e.g.,* Alan McQuinn, *The Economics of "Opt-Out" Versus "Opt-In" Privacy Rules,* INFORMATION TECHNOLOGY & INNOVATION FOUNDATION (October 6, 2017), https://perma.cc/B2UH-GPD8 (citing several studies that asked survey respondents

paradox's diagnosis is incorrect, the existence of the symptoms it attempts to analyze is well-supported. People tend to assert that they value privacy in opinion surveys, but they nevertheless take advantage of free products and services in exchange for providing their information[50] or fail to take advantage of safeguards for their privacy, such as reading privacy policies,[51] rejecting privacy-invasive user settings, or failing to affirmatively choose privacy-protective user settings. In one study that ranked the privacy attitudes of participants as high, medium, or low concern, only 46% of the high concern group had informed themselves of the existence or content of a monitoring policy at their school or workplace, and 41% of that same group reported their frequency of reading privacy policies as "rare."[52] This is logically inconsistent, but it invites a far wider range of conclusions than simply deciding that actions must speak louder than words, and that the validity of expressed preferences is vitiated by contradictory behavior.

The Senate hearing in response to Facebook sharing 87 million users' data with Cambridge Analytica without their consent provides a recent example of how this logic is often employed.[53] As Senator Ron Johnson questioned Mark Zuckerberg about user responses to revelations that the social media had shared their information, Zuckerberg said that he had not seen a "dramatic falloff" of users leaving the platform. Johnson replied, "But it seems like Facebook users still want to use the platform because they enjoy sharing photos and they share the connectivity with family members, that type of thing. And that overrides their concerns about privacy."[54] This logic ignores the fact that Facebook, like many other technologies that create collateral privacy risks, offers a service that may be impractical or impossible for users to fully extricate themselves from, whether because it is the only way to communicate with family members, crucial for publicizing a small business, or any of the other needs that

whether they would pay for the privacy they already received for free, and concluding that the only reason why public opinion polls claim to support strong privacy legislation "is because these surveys rarely confront consumers with the price consequences of their choices," and claiming that as regulation would require consumers to pay more for privacy, they would not actually support strong privacy legislation despite saying that they do.).

[50] *See* Nissenbaum, *supra* note 49, at 105.

[51] *See* McDonald & Cranor, *supra* note 41, at 550; Solove, *supra* note 40, at 1884.

[52] Alessandro Acquisti & Jens Grossklags, *Privacy and Rationality in Individual Decision Making*, 3 INST. ELECTRIC & ELECTRONICS ENGINEERS SECURITY & PRIVACY 26, 28 (2005).

[53] *See Facebook, Social Media Privacy, and the Use and Abuse of Data: Hearing Before the S. Comm. on the Judiciary, S. Comm. on Commerce, Science, and Transportation*, 115th Cong. (2018).

[54] Transcript of Mark Zuckerberg's Senate Hearing, *supra* note 2.

technology companies have worked hard to fill.[55] For the many valid concerns the Cambridge Analytica debacle may have given users about their privacy, users may have equally convincing reasons not to leave the platform, all of which the Senator (and those who would employ similar logic) ignores.[56]

Similarly, in a report criticizing privacy opinion surveys and their role in policymaking, Jim Harper and Solveig Simpleton note that despite many surveys reporting a high level of respondent "concern" about giving out credit card information or social security numbers online, consumers still do so, which they conclude is a clear indication that the reported concerns are false.[57] This is an extreme example of privacy paradox myopia, as there is a wide range of goods, services, and opportunities that can only practically be acquired by submitting those numbers over the Internet, and the fact that consumers are able and willing to do so has enabled the same vibrant Internet economy that the authors otherwise champion. The vast majority of government benefits also require applicants and recipients to disclose their social security number – many of these have online portals, which a person might use despite a privacy concern out of convenience or necessity, like a disability benefit recipient who is unable to use a phone.[58] The rationale of the privacy paradox is simplistic, and it assumes a causal connection between preference and behavior without accounting for the factors that separate hypothesized cause from observable effect.

Some surveys attempt to infer apathy from inconsistency, while others simply frame their questions to elicit desired answers. One such study conducted in 2016 by the Digital Advertising Alliance (DAA), an umbrella organization of advertising trade groups, purported to

[55] *See* Aja Romano, *How Facebook Made It Impossible to Delete Facebook,* VOX (March 22, 2018, 2:00 PM EDT), https://perma.cc/5ZN9-FSDV.

[56] *See id. See also* April Glaser, *The Problem With #DeleteFacebook*, SLATE (March 21, 2018, 3:46 PM) https://perma.cc/EP62-UASD (arguing that telling users to delete their Facebook accounts ignores the extent to which the platform is just as inextricable from the lives of many of its users as the company fought hard to become).

[57] JIM HARPER & SOLVEIG SIMPLETON, WITH A GRAIN OF SALT: WHAT CONSUMER PRIVACY SURVEYS DON'T TELL US 4 (2001), https://perma.cc/8R82-FHW2.

[58] U.S. GOV'T ACCOUNTABILITY OFF., GAO-04-768T, SOCIAL SECURITY NUMBERS: USE IS WIDESPREAD AND PROTECTIONS VARY 11 (2004) (detailing the range of government benefits that require SSN disclosure); Ctr. on Budget & Policy Priorities, *Online Services for Key Low-Income Benefit Programs: What States Provide Online With Respect to SNAP, TANF, Child Care Assistance, Medicaid, CHIP, and General Assistance* (July 29, 2016), https://perma.cc/EM2F-U3Q3 (discussing and listing the various state benefit programs with online components). To receive unemployment benefits in the state of Maryland, for examples, recipients are able to file weekly forms through an online portal, WEBCERT, using their SSN as a username. *See* Md. Dept. of Labor, Licensing and Regulation, *Webcert Logon* (last visited Sept. 23, 2018), https://perma.cc/E74N-AZP5.

demonstrate that consumers actively want an Internet supported by advertising, which was then used to argue that privacy regulation that could hamper advertisers' pervasive tracking contravened consumers' wishes.[59] But the indirect logic here is by design, as the survey's questions provide questions designed to promote a framing of free services versus paid services, relying on the inference that if consumers prefer not paying for things then they do not value privacy or legal protections for it.[60] Similar arguments have been posed as to the precision of advertising, with industry representatives arguing that any survey demonstrating that consumers value tailored advertisements means that they do not support privacy regulation that could hinder that targeting in any way.[61] The very premise of the question ignores the existence of privacy risks engendered by companies collecting enormous amounts of data and using it in opaque ways, the ill regulation seeks to mitigate. In a world without data breaches, identity theft, and information misuse, I share most peoples's preference for free services. But use of a product or service cannot be extrapolated to indicate endorsement of its drawbacks, or support for permissive

[59] For example, in comments to the Federal Communications Commission arguing against implementation of the Commission's now-defunct broadband privacy rule, Dan Jaffe of the Association of National Advertisers argued that the rule would be contrary to the public interest as "consumers want, expect, and benefit from interest-based advertising." *See* ANA comments, *supra* note 8, at 11. Similarly, Luigi Mastria of the Digital Advertising Alliance cited statistics from a 2013 survey conducted by Zogby and DAA as evidence of why consumers do not support Internet browsers that block cookies, which are a fundamental component of web tracking. Luigi Mastria, *Hearing on A Status Update On The Development Of Voluntary Do-Not-Track Standards Before the S. Comm. on Commerce, Science & Transportation*, 113th Cong. 9 (April 24, 2013) (statement of Luigi Mastria); Alison M. Cheperdak, *Double Trouble: Why Two Internet Privacy Enforcement Agencies Are Not Better than One for Business or Consumers*, 70 FED. COMM. L. J. 261, 294 (2018) (arguing that "[t]he FCC's increased regulations will have a negative impact on consumers because most consumers are not opposed to sharing information with Internet business in exchange for free or discounted services," citing the 2016 DAA/Zogby survey, *see infra* note 60 and accompanying text, that reported "[m]ore than 85 percent of respondents said they preferred [that] ad-supported Internet model instead of paying for online content.").

[60] DIGITAL ADVERTISING ALLIANCE, ZOGBY ANALYTICS PUBLIC OPINION SURVEY ON VALUE OF THE AD-SUPPORTED INTERNET 5 (May 2016), https://perma.cc/9PM7-9RBT ("Question 11: Which of the following would you prefer: an Internet where there are no ads, but you have to pay for most content you read/see like blogs, entertainment sites, video content and social media, or today's Internet in which there are ads, but most content is free?"). Unsurprisingly, 85% answered "an ad-supported Internet where most content is free," and only 14% opted for "a *paid* Internet where *everything cost money* because there is no advertising." *Id.* at 5 (emphasis added).

[61] *See, e.g.*, Daniel Castro, *Stricter Privacy Regulations for Online Advertising Will Harm the Free Internet*, INFORMATION TECHNOLOGY & INNOVATION FOUNDATION (Sept. 2010), https://perma.cc/E5VH-8YLC (noting that found that stricter privacy laws in the EU diminished the effectiveness of advertising and other evidence to claim "[t]he evidence clearly suggests that the tradeoffs of stronger privacy laws result in less free and low-cost content and more spam (i.e. unwanted ads) which is not in the interests of most consumers.").

privacy regulation that allows those drawbacks to perpetuate. Other surveys have used similar framing devices to elicit responses that seem to undermine a public desire for privacy, like another DAA survey that asked "[w]ould you support a law that restricted how data is used for Internet advertising, but also potentially reduced the availability of free content like blogs and video sites online?"[62] This provides the evidence the study's funders intend to produce, but contributes no meaningful information about the respondents' actual thoughts on privacy legislation.

The use of deliberately framed survey responses as empirical evidence of widespread privacy apathy contributes to broad misunderstandings of privacy norms and enables a lax regulatory regime when used as ammunition against policy reforms. When an individual's control over their information through notice and choice mechanisms is used as a representative illustration of their privacy preferences and expectations, it is possible to conclude that the failure to exercise that control constitutes a deliberate rejection of privacy. In reality, a range of cognitive and structural forces impede the individuals' incentives and abilities to take privacy-protective steps, obstacles that tend to be ignored in the rhetorical haste to drive the last nail into privacy's coffin.

The following sub-parts will detail the structural and cognitive limitations that complicate the conclusions of the privacy paradox, and illustrate the difficulty of relying on user privacy decision-making (or on surveys that ignore the barriers to coherent privacy decision-making) as a referendum on public privacy preferences or expectations.

2. Structural Limitations

The information asymmetry between consumers and data-collecting entities is perhaps the greatest impediment to delivering functional notice or meaningful choice. Consumers do not understand how privacy policies are intended to work, and most privacy policies fail to effectively convey the information they are intended to provide. In many contexts, a company may not even be able to disclose the trajectory of where the consumer's information might travel because it does not know what third parties might receive the consumer's information or how they might use it.[63] Few people read privacy

[62] *Interactive Survey of US Adults*, DIGITAL ADVERTISING ALLIANCE (Apr. 2013), https://perma.cc/2WHM-G4TW.

[63] Researchers continually discover new ways in which consumer behavior is tracked online; activity that is usually unknown to consumers and sometimes unknown to both the consumer and the company. *See* Gunes Acar, Steven Englehardt & Arvind Narayan, *No Boundaries for User Identities: Web Trackers Exploit Browser Login Managers*, FREEDOM TO TINKER (Dec.

policies, and those who do are left with little basis to understand the uses of their data.[64]

When individuals do take the time to read privacy policies, they tend not to understand what they mean. Multiple studies have illustrated that most individuals have a tenuous grasp of what a privacy policy is actually intended to accomplish.[65] In one such study, more than half of respondents believed that "[w]hen a company posts a privacy policy, it ensures that the company keeps confidential all the information it collects on users."[66] The typical privacy policy does nothing of the sort, and is designed to provide regulatory cover for the data collecting entity as much as it is to inform the user; if anything, privacy policies are more accurately described as corporate disclaimers, rather than consumer guarantees. As Senator John Kennedy memorably described it to Mark Zuckerberg as he testified in the Senate, "[t]he purpose of that user agreement is to cover Facebook's rear end. It's not to inform your users about their rights … tell your $1,200 an hour lawyers … you want it written in English and non-Swahili, so the average American can understand it. That would be a start."[67] In one study that interviewed experts, knowledgeable users,

27, 2017), https://perma.cc/G6PR-F4QG (explaining "how a long-known vulnerability in browsers' built-in password managers [was] abused by third-party scripts for tracking on more than a thousand sites."); Steven Englehardt, Gunes Acar & Arvind Narayan, *No Boundaries: Exfiltration of Personal Data by Session-Replay Scripts*, FREEDOM TO TINKER (Nov. 15, 2017), https://perma.cc/R3BZ-PKKD (reporting that when third-party analytics scripts record user online behavior such as keystrokes and mouse behavior, "the extent of data collected by these services far exceeds user expectations," particularly when "[t]his data can't reasonably be expected to be kept anonymous."). *See also* Steven Englehardt, Gunes Acar & Arvind Narayan, *Website Operators Are in the Dark about Privacy Violations by Third-Party Scripts*, FREEDOM TO TINKER (Jan. 12, 2018), https://perma.cc/5NJZ-AKLG (discussing the trio's previous research and noting that many websites had no relationship with the third parties exploiting their user data, or any idea that the exploitation was occurring).

[64] Solove, *supra* note 40, at 1884 (citing research demonstrating how few people read privacy policies); Martin & Nissenbaum, *supra* note 48, at 180, n.4 (describing studies that document how little consumers understand how their data might be used); Joshua Gluck, Florian Schaub, Amy Friedman, Hana Habib, Norman Sadeh, Lorrie Faith Cranor & Yuvraj Agarwal, *How Short Is Too Short? Implications of Length and Framing on the Effectiveness of Privacy Notices*, 2016 SYMP. ON USABLE PRIVACY & SECURITY 321, 322 (noting that "It is fairly rare for individuals to read a privacy policy in its entirety" due to their complexity and length); JOSEPH TUROW, MICHAEL HENNESSY & NORA DRAPER, THE TRADEOFF FALLACY: HOW MARKETERS ARE MISREPRESENTING CONSUMERS AND OPENING THEM UP TO EXPLOITATION 8 (2015), https://perma.cc/WL5V-U5WV ("[S]ome lawyers who write the policies for large companies have acknowledged to Joseph Turow that the documents are legal tender not designed to be understood by ordinary people.").

[65] *See* Turow et al, *supra* note 64, at 8; Aaron Smith, *Half of Online Americans Don't Know What a Privacy Policy Is*, PEW RESEARCH CENTER (December 4, 2014) https://perma.cc/9GBK-H4HM .

[66] *See* Smith, *supra* note 65.

[67] Transcript of Mark Zuckerberg's Senate Hearing, *supra* note 2.

and untrained crowd-workers, frequent disagreements emerged between the experts regarding the meaning of key terms, as did disagreements between the expert consensus and the consensus of the crowdsourced workers.[68] This demonstrates both the opacity of privacy policies generally and their tendency to confuse the average consumer –both of which contradict the idea that informational self-determination through notice and choice reflects an informed decision-making process.

That privacy policies are long and difficult to understand prevents people from reading them, but shorter policies would likely fail to adequately describe the relevant practices a user should know.[69] Privacy policies are given the impossible task of providing information that is sufficiently complete, yet perfectly digestible for consumers, while fulfilling regulatory requirements.[70] Consumers are trapped between a rock and a hard place; privacy policies are too long to be worth reading, yet they fail to convey relevant information needed to make an informed decision. Moreover, they are highly unlikely to disappear from how consumer privacy is governed anytime in the near future.[71] At the same time, poor understanding of the Internet ecosystem also engenders misleading survey results.[72]

The same lack of understanding of the substance and applicability of privacy policies seems to persist in users' understandings of application permissions, corporate data practices, and surveillance

[68] *See* Joel Reidenberg, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T. Graves, Fei Liu, Aleecia McDonald, Thomas B. Norton, Rohan Ramanath, N. Cameron Russell, Norman Sadeh & Florian Schaub, *Disagreeable Privacy Polices: Mismatches Between Meaning and Users' Understanding*, 30 BERK. TECH. L. J. 39, 86-87 (2015). *See also* Cranor, *supra* note 28, at 274 (noting that privacy policies are "long, complicated, full of jargon, and change frequently.").

[69] *See generally* Gluck et al., *supra* note 64; Kanthashree Mysore Sathyendra, Shomir Wilson, Florian Schaub, Sebastian Zimmeck & Norman Sadeh, *Identifying the Provision of Choices in Privacy Policy Text,* 2017 CONF. ON EMPIRICAL METHODS NAT. LANGUAGE PROCESSING 2774, 2774, https://perma.cc/V33S-THRQ (noting that the level of education required to understand the average privacy policy is higher than that of the average individual).

[70] Jialiu Lin, Bin Liu, Norman Sadeh & Jason I. Hong, *Modeling Users' Mobile App Privacy Preferences: Restoring Usability in a Sea of Permission Settings* 2017 CONF. ON EMPIRICAL METHODS IN NAT. LANGUAGE PROCESSING 2774, 2774 (describing the competing requirements of privacy policies and the difficulty of simultaneously fulfilling all of them).

[71] *See generally* Solove, *supra* note 40, at 1885 (discussing the "fundamental dilemma" of notice).

[72] Turow et al., *supra* note 64, at 4-5 (reporting widespread misconceptions about advertising, for example, "49% of American adults who use the internet believe (incorrectly) that by law a supermarket must obtain a person's permission before selling information about that person's food purchases to other companies," and that 65% did not know that the statement "[w]hen a website has a privacy policy, it means the site will not share my information with other websites and companies without my permission" is false). *See also* Tokson, *supra* note 13, at 179-180 (distinguishing knowledge polls as less vulnerable to distortion than opinion polls, as they make fewer faulty assumptions based on information respondents do not possess).

practices. In one study, only 22% of participants understood that applications continue to run in the background when the user is not directly engaged with them and that an app can still access their information when not in use.[73] Even when individuals are fully informed and take proactive steps to protect their privacy, their actions may not have the desired impact. In one example, researchers at Princeton demonstrated that when a smartphone user turns off location services, her location can still be deduced from other sources of publicly-available information.[74] Another study demonstrated the limited efficacy of popular tracking blockers and browser plug-ins, as many users mistakenly believed they had successfully shielded their browsing activity.[75] And let's not forget the constant barrage of data breaches hitting companies large and small, obviating whatever promises those companies made to their users about the privacy or security of their information.[76] Other companies may employ technological workarounds to infer information that their privacy policies claim they refrain from collecting, making the disclaimers in their privacy policies, should consumers take the time to read them, hollow.[77] These structural forces – from the lack of consumer knowledge about how privacy policies work, to the frequent inability of companies to effectively disclose practices outside their control – make it all but impossible for consumers to make meaningful decisions that exert effective control over their information.[78]

[73] Lynn Tsai, Primal Wijesekera, Joel Reardon, Irwin Reyes, Serge Egelman, David Wagner, Nathan Good & Jung-Wei Chen, *Turtle Guard: Helping Android Users Apply Contextual Privacy Preferences*, 148, 2017 SYMP. ON USABLE PRIVACY & SECURITY 145, 148, *citing* Christopher Thompson, Maritza Johnson, Serge Egelman, David Wagner & Jennifer King, *When It's Better to Ask Forgiveness than Get Permission: Attribution Mechanisms for Smartphone Resources*, 2013 SYMP. ON USABLE PRIVACY & SECURITY 1, 3.

[74] Arsalan Mosenia, Xiaoliang Dai, Prateek Mittal & Niraj K. Jha, *PinMe: Tracking a Smartphone User around the World*, 4 INST. ELECTRICAL & ELECTRONIC ENGINEERS TRANSACTIONS ON MULTI-SCALE COMPUTING SYSTEMS 420, 420 (2017).

[75] William Melicher, Mahmood Sharif, Joshua Tan, Lujo Bauer, Mihai Christodorescu, & Pedro Giovanni Leon, *(Do Not) Track Me Sometimes: Users' Contextual Preferences for Web Tracking*, 2016 PROC. ON PRIVACY ENHANCING TECH. 135, 137, *citing* Pedro Leon, Blase Ur, Richard Shay, Yang Wang, Rebecca Balebako & Lorrie Cranor, *Why Johnny Can't Opt Out: A Usability Evaluation of Tools to Limit Online Behavioral Advertising*, 2012 CONF. ON HUMAN FACTORS COMPUTING SYSTEMS.

[76] *See, e.g.*, Christopher Mele, *Data Breaches Keep Happening. So Why Don't You Do Something?*, N.Y.T. (Aug. 1, 2018), https://www.nytimes.com/2018/08/01/technology/data-breaches.html (describing the millions of Americans affected by the barrage of hacks and the difficulty of both breach response and prevention).

[77] Hoofnagle, *supra* note 39, at 170, n. 92 (describing "reverse enhancement," where a data broker uses one form of information and other databases to deduce a consumer's home address).

[78] *See* Schwartz & Peifer, *supra* note 9 149 ("disclosure is a ritual to be endured"), *quoting* OMRI BEN-SHAHAR & CARL E. SCHNEIDER, MORE THAN YOU WANTED TO KNOW: THE FAILURE OF MANDATED DISCLOSURE 10 (2014).

3. Cognitive Limitations

While the informational asymmetry discussed presents a baseline obstacle to rational privacy decision-making, cognitive limitations also skew the process, further preventing privacy behavior from accurately reflecting users' privacy preferences or expectations. The logic of notice and choice assumes a baseline consumer rationality in privacy decision-making that is rarely, if ever, present[79] – and if privacy decision-making is not rational, the logic that individuals do not want privacy protections because they fail to engage with notice and choice is fundamentally unsound.

Social scientists and privacy scholars have described the cognitive phenomena that impede rational decision-making, such as hyperbolic discounting, which posits that people assign a lesser value to less ascertainable, far-off rewards, and higher value to rewards that are easily acquired.[80] In the case of privacy decision-making, when a privacy policy is densely opaque, and the possible benefit of preventing a hypothetical privacy harm is juxtaposed with the immediate benefit of a sales discount or free WiFi, hyperbolic discounting is one of the reasons why the overwhelming majority of the population would choose the latter.[81] The possibility that a snoop might be intercepting an individual's web traffic, and that the traffic could be directly or inferentially valuable enough to the snoop to somehow cause her harm, seems tenuous and remote compared to the concrete and immediate reward of Internet access.

A similar phenomenon is the idea of bounded rationality, which posits that humans are fundamentally limited in the amount of information and skills we can harness and apply to a given decision, and therefore use simplified metrics and approximations that are ultimately unrepresentative, resulting in an irrational outcome.[82] Otherwise put, it contradicts the idea that consumers will generally act in a rational way, maximizing utility and their own self-interest at any

[79] *See* Acquisti & Grossklags, *supra* note 52, at 26.

[80] *See generally* Solove, *supra* note 40, 1883; Acquisti & Grossklags, *supra* note 52, at 31.

[81] Acquisti & Grossklags, *supra* note 52, at 31.

[82] Alessandro Acquisti, *Privacy in Electronic Commerce and the Economics of Immediate Gratification*, 2004 PROC. ASS'N COMPUTING MACHINERY CONF. ON ELEC. COM. 21, 23; Acquisti & Grossklags, *supra* note 52, at 30; Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, Yang Wang & Shomir Wilson, *Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online,* 50 ASS'N COMPUTING MACHINERY COMPUTING SURVEYS Art. 44, at 2 (2017) (defining bounded rationality).

given turn.[83] Privacy decisions are particularly shaped by bounded rationality, as privacy choices are usually abstract and based on incomplete information, making a correct mental model of the decision all the more difficult to summon. In short, individuals often fail to engage in exactly the kind of clear-eyed, conditional thinking upon which the conclusion of the privacy paradox depends.[84]

Another cognitive phenomenon that affects individual decision-making is the common difficulty of accurately assessing the impact of cumulative risk. A phenomenon that may feel relatable to many readers (or at least to the author) occurs when people often discount the ultimate danger created by the series of incremental steps that will ultimately permit that danger to take place. In conjunction with hyperbolic discounting, this means a user is not only favoring the immediate reward over the possible risk when assessing a possibly privacy-invasive scenario, but failing to correctly ascertain that initial risk, even when the decision may not produce a reward. Examples include assessing the danger of small decisions like repeatedly using unsecured public WiFi networks, or failing to use strong and different passwords for a range of different online accounts against the risk of identity theft.[85] These risks may be even more challenging for connected devices where the context of data collection may seem more limited, but the possibilities for abuse are not.[86] Conversely, while an amorphous risk and the value of dignity may be difficult for consumers to quantify when assessing a transaction, the data collector faces no such difficulty in quantifying the value of consumers' information.[87] The way that individuals process risk is highly relevant for understanding both how notice and choice actually functions, and for correctly interpreting surveys that purport to measure which privacy risks individuals care about.

[83] Shaun B. Spencer, *Reasonable Expectations and the Erosion of Privacy*, 39 SAN DIEGO L. REV. 843, 899 (2002) (describing bounded rationality and its impact on privacy decision-making).

[84] Laura Brandimarte, Alessandro Acquisti & George Lowenstein, *Misplaced Confidences: Privacy and the Control Paradox*, 4 SOC. PSYCHOL. & PERSONALITY SCI. 340, 341 (2014).

[85] For another example, see Spencer, *supra* note 83, at 897. "For a consumer buying an appliance, the cost of [providing] her address will probably seem trivial compared to the cost of not buying the appliance. Consumers generally make their data-sharing decisions within the framework of each incremental transaction in which they participate, while merchants base their practices on the realities and economies of scale of the data profiling business." The risks compound, but each transaction creating it does would not alone justify the kind of calculation that the privacy paradox would require a "pro-privacy" consumer to make.

[86] Florian Schaub, Bastian Konings & Michael Weber, *Context-Adaptive Privacy: Leveraging Context Awareness to Support Privacy Decision Making*, 14 INST. ELECTRICAL & ELECTRONIC ENGINEERS PERVASIVE COMPUTING 34, 34 (2015).

[87] Spencer, *supra* note 83, at 897-98.

Decision fatigue further limits the efficacy of notice and choice. Consumers are faced with a constant barrage of privacy decisions, which compounds the difficulty of making privacy decisions that suit their expectations or preferences. These decisions include reading privacy policies, terms of service, checking application permissions, and others. As these decisions are nearly constant, a user engaging with them in a meaningful way, reading every privacy policy or going to the website of every app or IoT device, is not only unlikely but an impractical use of one's time.[88] One study famously estimated that it would take the average American internet user forty-minutes per day on average to read every privacy policy she came across, at a cost of $2,533-$5,038 a year.[89] Declining to read a privacy policy could be attributed to rational ignorance, when a consumer makes an economically logical decision not to seek out a sufficient explanation about a given scenario, such as deciding to forego the requisite time to read a privacy policy that will ultimately fail to effectively convey the necessary information anyway.[90] To make matters more confusing, providing too much control, by providing consumers with exhaustive information as to every way in which their data is used and every time it is used, can have the paradoxical effect of paralyzing individuals, rather than empowering them.[91]

The cognitive phenomena and informational asymmetry that prevent consumers from making privacy decisions that suit their preferences are considerable. But even when consumers are knowledgeable about how their information is used, the risks a given decision can create, and the basic steps they can take to mitigate those risks, consumers still fail to take those mitigating steps due to resignation, rather than apathy towards their privacy or an affirmative decision to subordinate their privacy preference to another value.[92] The difference between a person's privacy preference and their privacy expectations can further help diagnose the incoherence between commonly expressed privacy preferences and ultimate privacy behavior. A person with a preference for privacy acts in the hope that her action will protect her information; a person with an expectation of

[88] McDonald & Cranor, *supra* note 41, at 544.

[89] *Id.* at 563-64.

[90] Tokson, *supra* note 13, at 167 (defining rational ignorance generally and as a flaw in the reasonable expectation of privacy test's reliance on individual knowledge of privacy and technology).

[91] Hartzog, *supra* note 43, at 975.

[92] Turow et al., *supra* note 64, at 13, 17 (defining resignation, and reporting that "[T]he more a person knows about information collection in the marketing world, the more likely the resignation.").

privacy is convinced that it will. The convergence of a strong privacy preference, with the conviction that any pro-privacy steps will fail to accomplish their stated goal, illustrates a key part of the "paradox"—the resignation of the informed consumer.[93] "Learned helplessness" describes when an individual in a negative situation with no recourse to change it accepts the situation as a coping mechanism – lack of power to change the cause of the negative circumstance shapes the individual's response to it.[94] A study on Americans' attitudes towards online privacy and tradeoffs reported that the majority of those who were willing to trade their information for discounts under certain scenarios did so resigned to the possible risks, rather than as a rational decision weighing privacy, utility, and possible adverse outcomes.[95] The study further reported that those who are the most informed about online privacy are the most likely to be resigned to its abuse.[96]

When privacy skeptics and motivated industry coalitions use flawed methodologies to arrive at desired policy conclusions, their evidence appears convincing; it seems that consumers do not actually *want* the privacy protections that regulators would give them, and as privacy is a good to barter away rather than a right that must be shielded at a basic threshold from certain harms, any privacy enforcement (such as a new rule, law, or interpretation of policy) should be heavily influenced by those wishes.[97] Surely, if consumers will offer their email addresses in exchange for the use of public WiFi, they have rationally weighed the costs and the benefits of the transaction and concluded that they do not want to pay for the services they have grown accustomed to receiving for free. To ignore their choice in a free market would seem paternalistic and undemocratic.[98] And unlike other areas of the law,

[93] *Id.* at 17.

[94] Bloom et al., *supra* note 47, at 367 (discussing learned helplessness, and remarking that learned helplessness could have impacted the response of participants who rated a potentially privacy-invasive scenario as 'likely' who were more likely to be comfortable with that scenario).

[95] Turow et al., *supra* note 64, at 4, 14 (58% were "resigned" to the misuse of their information, and 72% rejected the statement "what companies know about me from my behavior online cannot hurt me.").

[96] *Id.* at 9. As limiting as anecdotal proof is, this phenomenon seems logical to me, as I write this paper on privacy behavior and expectations while connected to the open WiFi network of my bus.

[97] *See, e.g.*, Meredith Kapushion, *Hungy, Hungry HIPAA: When Privacy Regulations Go Too Far*, 31 FORDHAM URB. L.J. 1483, 1491 (2004) (criticizing HIPAA for not accommodating divergent consumer desires for their health data to be monetized).

[98] *See, e.g.*, Adam Thierer, *A Framework for Benefit-Cost Analysis In Digital Privacy Debates*, 20 GEO. MASON L. REV. 1055, 1068 (2013) ("Paternalistic claims clash mightily with the foundational principles of a free society – namely, that individuals are autonomous agents that should be left free to make choices for themselves, even when some of those choices strike others as unwise . . . . [A]ssertions that people cannot be trusted to look out for themselves … would imply that the benefits of regulation are virtually boundless and that the costs should generally be ignored in order to essentially save consumers from their own choices.").

where stronger guardrails exist to prevent consumers from being fleeced by an unconscionable transaction they purportedly want to engage in, privacy would seem to present no such countervailing concern, with the exception of the most concrete and egregious physical and monetary harms.

But as this article has hopefully begun to demonstrate, these arguments depend on a rational privacy decision-making process that simply do not exist. Opinion surveys fail to account for the heavy role that context plays in privacy decision-making, the lack of information most individuals have about the information ecosystem, and the cognitive phenomena that impact privacy decision-making that all contribute to the fallacy of the informed and rational privacy decision-maker. Advertising-dependent industries have strong incentives to keep privacy as unregulated as possible, and to make motivated arguments that a system dependent on a fallacious behavioral model is the only possible way to preserve freedom and innovation. Those arguments are politically useful, and they've often worked. As long as consumer privacy is considered to be a good that individuals should always have the right to trade away, motivated analysis of flawed methodologies will continue to provide support for arguments that dress commercial or ideological objectives in democratic clothing. The simple fact that someone might answer that they care about their privacy in a survey, and yet use public WiFi, will continue to be offered as proof that whatever privacy law aims to protect, consumers neither need nor want it.

Introducing new types of privacy research into policymaking will not affect the incentives of the stakeholders making those arguments. But empirical research that is able to measure privacy preferences, expectations, and behavior accurately – by accounting for the context of different privacy decisions, the loaded premises behind broad questions about complex technology, and the fallacy of the rational privacy decision-maker – can help illustrate for regulators and the public which policy prescriptions are medicinal, and which are snake oil designed to benefit the companies that sell them, in addition to illuminating privacy preferences, expectations and behaviors that are often difficult to discern.

Of course, the possibility remains that when individuals make decisions that do not prioritize privacy, they are simply prioritizing other values such as increased convenience, functionality, or a good or service. Not every privacy decision is an indictment of the notice and choice system. Like most goods, the value of privacy in a given circumstance is finite; like most rights, it is rarely absolute. But for that assessment to be accurate, it must be based on informed and meaningful privacy decision-making which, due to the cognitive and structural limitations discussed above, rarely exists in practice.[99] Asserting that decision-making under these circumstances reflects a free and informed choice is a facile distortion of how real people think and act, and should be repudiated when used as evidence of popular enthusiasm for permissive privacy regulation. For the use of notice and choice to provide a remotely faithful snapshot of individual privacy preferences and expectations – and for empirical research to do the same – the barriers to coherent decision-making must be fully accounted for.[100]

B. The Fourth Amendment

Whereas social science is often poorly used in consumer privacy legal and policy discussions, it is rarely used to inform assessments of a reasonable expectation of privacy under the Fourth Amendment. Regulators and legislators ask the public for input on how they should approach privacy issues, which creates a space for research to answer their questions: judges usually do not. In both areas, government decision-makers struggle to discern privacy norms that they need to understand in order to interpret or create the law – evaluating what constitutes a reasonable expectation of privacy under the Fourth Amendment has proven particularly challenging in the case of emerging technology.[101]

As a baseline, the Constitution protects privacy from government intrusion in the Fourth Amendment, which states that the government must obtain a warrant based on probable cause before searching or seizing persons, papers, places, or effects.[102] This has been interpreted through subsequent cases to mean that the government must obtain a warrant in order to search or seize something in which a person has a subjective expectation of privacy that society would deem objectively

[99] *See, e.g.*, Turow et al., *supra* note 64, at 16 (finding that most Americans don't have the basic knowledge to make informed cost-benefit choices).

[100] *Id.* at 3 ("Marketers justify their data-collection practices with the notion of tradeoffs, depicting an informed public that understands the opportunities and costs of giving up its data and makes the positive decision to do so . . . . This image of a powerful consumer has become a way to claim to policymakers and the media that Americans accept widespread tracking of their backgrounds, behaviors, and lifestyles across devices, even though surveys repeatedly show they object to these activities."); *id.* at 20 ("The findings indicate substantial tensions in a central area of society's public sphere that cannot be swept away by executives' assertions of consumer autonomy and rational choice.").

[101] *See generally* Orin S. Kerr, *An Equilibrium-Adjustment Theory of the Fourth Amendment*, 125 HARV. L. REV. 476, 481 (2011),

[102] U.S. CONST. amend. IV.

reasonable, unless a warrant exception applies.[103] While there have been a few landmark cases that have recognized the need for new rules,[104] most Fourth Amendment jurisprudence is hamstrung by doctrinal principles that far predate the technology to which the principle is applied, resulting in outcomes that undermine the underlying values of the Amendment.[105] While a common law system necessarily requires extending old principles to new facts, judges have to be able to accurately assess societal norms, which they are rarely positioned to be able to do well.[106] Fourth Amendment protections end up being tied to the reasonable judge's expectation of privacy, not the reasonable person's. The result is a similar mismatch of what legal doctrine presupposes and how human beings actually think and act.

1. Analog Doctrines

One aspect of why evaluating a modern reasonable expectation of privacy is so difficult is that technology has outpaced the applicability of the logic supporting many Fourth Amendment doctrines, creating the need for new rules that will uphold, rather than contravene, the privacy protections the Constitution is intended to confer.

Perhaps the most glaring instance of a blunt Fourth Amendment standard that fails to account for the categorically disruptive impact of new technology is the third-party doctrine. This is the idea that by entrusting information to a third party, the sender extinguishes her expectation of privacy in the information. The doctrine originally developed in two cases, *United States v. Miller*, and *Smith v. Maryland*,

[103] Katz v. United States, 389 U.S. 347, 361 (1967) (Harlan, J., concurring); Riley v. California, 134 S. Ct. 2473, 2483-84 (2014) (describing Fourth Amendment exceptions). *See also* Slobogin & Schumacher, *supra* note 11, at 731 (describing the evolution of the "societal understanding" of privacy norms as a function of the *Katz* test).

[104] *See e.g.*, *Riley*, 134 S. Ct. at 2491 (holding that a cell phone is not a container for search incident to arrest purposes, and that the Fourth Amendment requires a warrant to search one); Kyllo v. United States, 533 U.S. 27, 40-41 (2001) (holding that a search of the home using a device that is not in public use is presumptively unreasonable); United States v. Warshak, 631 F.3d 266, 286 (6th Cir. 2010) (finding that defendant had a reasonable expectation of privacy in his emails despite the fact that they were stored on a third-party server).

[105] *See* United States v. Jones, 132 S. Ct. 945, 958 (Alito, J., concurring in the judgment) ("The Court argues – and I agree – that 'we must "assur[e] preservation of that degree of privacy against government that existed when the Fourth Amendment was adopted."' But it is almost impossible to think of late 18th-century situations that are analogous to what took place in this case. (Is it possible to imagine a case in which a constable secreted himself somewhere in a coach and remained there for a period of time in order to monitor the movements of the coach's owner?)" (citations omitted)).

[106] Andrew D. Selbst, *Contextual Expectations of Privacy*, 35 CARDOZO L. REV. 643, 649 (2013) (describing the flaws of binary Fourth Amendment doctrines and arguing that a context-based approach can provide a more accurate descriptive measure of privacy norms, as well as a more coherent structure to judicial analysis of Fourth Amendment cases).

involving the privacy interest in the routing numbers sent to the defendant's bank, and the privacy interest in the phone numbers sent to the defendant's phone company respectively.[107] But in the digital age, it is essentially impossible to communicate without entrusting a range of information to digital service providers. Sending a text message requires the cell phone's signal to be collected by the nearest cell tower, and for a smart phone, often to a provider of cloud storage services as well. The use of email or the Internet requires routing traffic through an Internet service provider.[108] The mulishly bifurcated reasoning of the third-party doctrine creates a scenario where a text message cannot contain a reasonable expectation of privacy in the situation where a folded paper note would because in application the doctrine fails to consider the context separating various new applications.

Other tenets of Fourth Amendment doctrine are subject to similar critique. The automobile exception provides that under most circumstances, a police officer does not require a warrant to search a vehicle. This logic is predicated on the exigency concern of searching a mobile entity, as well as the idea that the pervasive regulation of automobiles undermines a reasonable expectation of privacy in one.[109] It is also about as binary a standard as they come – the degree of mobility of an entity does not control whether it is considered a "vehicle" for the purposes of the exception, including cases like an immobile trailer.[110] But the vast majority of vehicles are now equipped with GPS, such that the vehicle keeps a careful record of the drivers' location history. Infotainment systems also allow drivers to sync the contact information in their cellphones, or the information from intelligent assistant devices located in their homes, an area that is afforded the strongest possible Fourth Amendment protections.[111] This categorically changes the privacy interests involved in searching a

[107] United States v. Miller 425 U.S. 435, 442-43 (1976); Smith v. Maryland, 442 U.S. 735, 743-44 (1979).

[108] *See generally* Aaron Rieke, David Robinson & Harlan Yu, *What ISPs Can See*, UPTURN (March 2016), https://perma.cc/T5TZ-KDBM.

[109] *See* California v. Carney, 471 U.S. 386, 391 (1985) ("Even in cases where an automobile was not immediately mobile, the lesser expectation of privacy resulting from its use as a readily mobile vehicle justified application of the vehicular exception.").

[110] *See* United States v. Navas, 597 F.3d 492, 499 (2d Cir. 2010).

[111] Erin Biba, *How Connected Car Tech Is Eroding Personal Privacy*, BBC (Aug. 9, 2016), https://perma.cc/R6DS-AA2V.

vehicle,[112] and could not have been taken into account in the case that established the exception in 1925.[113]

Another blunt standard rendered inapt by technology is the content/non-content distinction, a line of jurisprudence that holds that unlike the content of communications, related "non-content" information associated with the communications have no Fourth Amendment privacy interest.[114] This distinction arose through a series of cases, starting with *Ex parte Jackson*, which distinguished the contents of a letter from the information on its envelope,[115] and was established more fully in *Smith*, which distinguished the contents of a telephone conversation from the telephone numbers dialed by the caller.[116] The content/non-content distinction is also present in the Stored Communications Act (SCA), a law designed to limit the government's access to the electronic information held by network intermediaries.[117] The statute requires the government to seek a warrant for the contents of electronic communications, but permits it to obtain "records or other information pertaining to a subscriber" on the basis of a subpoena showing that there are "reasonable grounds" that the information is "relevant and material" to the investigation.[118] The binary distinction is determinative here; content requires a warrant while non-content does not. However, subscriber information, customer records, and other non-content can be far more revealing that the distinction implies, particularly in aggregate.[119] Metadata – "data

[112] Lindsey Barrett, *supra* note 24, at 194 (arguing that the privacy implications of the data collected by connected cars and automated vehicle undermine the logic of the Fourth Amendment's automobile exception, such that a warrant should be required to search vehicle data).

[113] Carroll v. United States, 267 U.S. 132, 149 (1925).

[114] *See, e.g.*, Paula Kift & Helen Nissenbaum, *Metadata in Context – An Ontological and Normative Analysis of the NSA's Bulk Telephony Metadata Collection Program*, 13 I/S: J. L. & POL'Y FOR INFO SOC'Y 333, 372 (2017) ("[T]he main assumption underlying the NSA's justification for the bulk telephony metadata collection program – namely, that metadata is equivalent to non-sensitive data – no longer makes sense. Indeed, in light of the theory of contextual integrity, it never made any sense to begin with.").

[115] *Ex parte* Jackson, 96 U.S. 727, 733 (1877).

[116] Smith v. Maryland, 442 U.S. 735, 743-44 (1979).

[117] 18 U.S.C. § 2703 (2012).

[118] *Id.* § 2703(d).

[119] Ian Samuel, *The New Writs of Assistance,* 86 FORDHAM L. REV. 2873, 2892 (2018) (noting that "[t]he government can and does use its power under the SCA to force the disclosure of enormous amounts of sensitive information about the users of network services," such as personal addresses, IP addresses, and account information for tweets relating to a leak of classified information); William Jeremy Robison, *Free at What Cost?: Cloud Computing Privacy Under the Stored Communications Act,* 98 GEO. L. J. 1195, 1208 (2010) (noting that for both electronic communications service providers and remote communications service providers, the two types of service providers subject to the SCA, "personal identifying information about the user such as her name, physical or e-mail addresses, and IP address, is entitled to little protection"). *See also*

about data," like the size of a Microsoft word document, the location a text was sent from, or the time stamp on an email[120] – constitutes non-content, and can thus be acquired by the police on a standard akin to reasonable suspicion, not probable cause, by virtue of that designation. Yet metadata itself can be tremendously revealing, in addition to the fact that it creates structured datasets that are relatively easy for computers to analyze.[121]

The Fourth Amendment's forced binary of the public/private distinction is another example of a blunt test that ignores the transformative impact of technology on the facts driving older doctrines and differences of modern expectations.[122] The public view doctrine provides that there is no expectation of privacy in something public, such as marijuana plants on private property visible from a low-flying airplane.[123] However, there is a reasonable expectation of privacy in information so difficult to obtain that law enforcement must use technology that is not in "general public use" to perceive it.[124] Here, too, the distinctions are more nuanced than that rule might suggest. In *United States v. Knotts*, when the police attached a tracking device to the defendants' vehicle, there was no reasonable expectation of privacy that merited a warrant, as the defendants' movements were in public.[125] Though a later GPS tracking case was decided on the basis of the trespass, two concurrences questioned the warrantless invasion of privacy that was permitted by twenty-eight days of tracking a vehicle, regardless of the fact the movements were public.[126] The prodigious spread of public surveillance technologies, such as drones, CCTVs,

Gabriel R. Schlabach, *Privacy In The Cloud: The Mosaic Theory and the Stored Communications Act*, 67 STANFORD. L. REV 677, 697 (2015) (arguing that an amendment to the SCA to incorporate mosaic theory is needed to adequately protect user privacy). *But see* Orin S. Kerr, *The Mosaic Theory of the Fourth Amendment*, 111 MICH. L. REV. 311, 320, 347 (2012) (proposing and rejecting the viability of an approach to Fourth Amendment searches that evaluated government conduct in aggregate, rather than sequentially, acknowledging the outsized privacy implications that low-cost surveillance methods enabled by new technology without being defined as a "search" under current doctrine).

[120] *See* Kift & Nissenbaum, *supra* note 114, at 336 (defining metadata).

[121] *See, e.g.*, Lee & Kobsa, *supra* note 46, at 288 (describing the creation of datasets to "build machine learning models to predict future privacy decisions).

[122] *See* Helen Nissenbaum, *Toward an Approach to Privacy in Public: Challenges of Information Technology*, 7 ETHICS & BEHAV. 207, 208 (1997); Joel R. Reidenberg, *Privacy In Public*, 69 U. MIAMI L. REV. 141, 142 (2014) (arguing that the spread of ubiquitous public surveillance undermines Fourth Amendment privacy protections); Selbst, *supra* note 106, at 650.

[123] California v. Ciraolo, 476 U.S. 205, 214-15 (1986).

[124] Kyllo v. United States, 533 U.S. 27, 40 (2001).

[125] United States v. Knotts, 460 U.S. 276, 281-82 (1983).

[126] United States v. Jones, 132 S.Ct. 945, 953 (2012); *id.* at 954 (Sotomayor, J., concurring); *id.* at 958 (Alito, J., concurring).

automated license plate readers, and facial recognition databases have further eroded the public-private distinction.[127]

2. The Reasonable Judge's Expectation of Privacy

Perhaps the most challenging aspect of coalescing Fourth Amendment doctrine with the realities of technology is that the reasonable expectation of privacy test requires a judge to determine what privacy norms are, a difficult task that they are poorly equipped to do.[128] Studies attempting to provide empirical answers for what constitutes a reasonable expectation of privacy have found that the norms that judges rely on in their cases often contradict the norms reported by average individuals.[129] For example, in *United States v. Forrester*, the Ninth Circuit found that there was no expectation of privacy in the websites a person visits or the email addresses with which they correspond, analogizing to the phone numbers dialed in *Smith*.[130] The court reasoned that on the basis of assumption of the risk, and the distinction between content (the email) and non-content (the address whence it was sent), there was no privacy interest protected by the Fourth Amendment, and to whatever extent there was, the defendant repudiated it by trusting a third party, the service provider. This would likely surprise the majority of individuals who think that a warrant should be required to access the email addresses they have

[127] *See generally* Reidenberg, *supra* note 122, at 147. *See also* GEO. L. CTR. ON PRIVACY & TECH., THE PERPETUAL LINEUP 25-26 (2016), https://perma.cc/8V9G-XHWT (detailing the pervasive use of facial recognition technology in city, state, and federal law enforcement).

[128] *See* Spencer, *supra* note 83, at 846-47; Slobogin & Schumacher, *supra* note 11, at 732 ("[T]he results strongly suggest that some of the Court's decisions regarding the threshold of the Fourth Amendment and the warrant and probable cause requirements do not reflect societal understandings. Indeed, some of the Court's conclusions in this regard may be well off the mark.").

[129] *See* Slobogin & Schumacher, *supra* note 11, at 732 (describing the results of an empirical study on privacy expectations, and concluding that the Supreme Court's understanding of those expectations often contradicts the norm); Susan F. Mandiberg, *Reasonable Officers vs. Reasonable Lay Persons In The Supreme Court's Miranda And Fourth Amendment Cases*, 14 LEWIS & CLARK L. REV. 1481, 1499 (2010) (arguing that judges rely on contradictory standards for the reasonable officer and the reasonable lay person); Marc McAllister, *The Fourth Amendment and New Technologies: The Misapplication of Analogical Reasoning*, 36 S. ILL. U. L. J. 475, 477-78 (2012) (arguing that misplaced analogical reasoning often skews judicial perception of societal expectation of privacy, including with respect to attitudes concerning assumption of the risk and the third party doctrine); Steven L. Chanenson, *Get The Facts, Jack! Empirical Research and the Changing Constitutional Landscape of Consent Searches*, 71 TENN. L. REV. 399, 437 (2004) (arguing for increased use of empirical evidence to provide more accurate information about societal perceptions about consent searches, namely under what circumstances a reasonable person would feel free to terminate an encounter with the police).

[130] United States v. Forrester, 512 F. 3d 500, 504 (9th Cir. 2007).

corresponded with, contradicting the basis for the third-party doctrine (and the holding of *Forrester*).[131]

The survey also reported other ways in which judges tend to hold different perspectives from the average person that are likely to impact their understanding of reasonable expectations of privacy. In contrast with the holding of *Knotts* and other cases that there is no reasonable expectation of privacy for movements in the public view, only 24% of respondents would find warrantless GPS surveillance of their movements for more than ten days to be acceptable.[132] Most people also believe they have a reasonable expectation of privacy in both GPS and CSLI, despite frequent judicial insistence that the emission of such information is both knowing and voluntary.[133] Justice Gorsuch underlined the divergence of doctrine and popular perception in his *Carpenter* dissent, highlighting research that most people distinguish between invasions of privacy for more serious crimes over more minor ones, whereas Fourth Amendment doctrine does not.[134]

In a comprehensive treatment on the subject of knowledge and the Fourth Amendment, Matthew Tokson documents a number of ways in which judges tend to overestimate the average person's understanding of online privacy and government surveillance.[135] While courts have held that users understand how ISPs operate and how email is sent, erroneous understandings of the internet ecosystem abound (such as the 61% of respondents in a Consumer Reports survey who believed that what they do online is never shared without their permission).[136] Courts have used the existence of privacy policies to determine that users were aware of the privacy interests they gave up, when for a range of reasons, most users neither read nor understand the privacy policies.[137] In all fairness to judges, rapidly shifting norms regarding privacy and technology make a reasonable expectation of privacy difficult for anyone to deduce, as this article has indicated. Nevertheless, the space between what the reasonable judge expects and

[131] *See* McAllister, *supra* note 129, at 480, 498 (reporting that 86.1% of respondents did not agree with the idea that exposing otherwise private information to a third-party provider constituted knowingly providing that information to law enforcement).

[132] *Id.* at 491.

[133] Brief for Empirical Fourth Amendment Scholars in Support of Petitioner, as Amici Curiae Supporting Petitioner at 5, Carpenter v. United States, 585 U.S. __ (2018) (No.16-402) (citing a range of studies reflecting that "more people expect privacy in their cell-site location information than do not.") (hereinafter *Carpenter Empirical Brief*).

[134] Carpenter v. United States, 585 U.S. ___ , No. 16-402, slip op. at 7-8 (U.S. June 22, 2018) (Gorsuch J., dissenting).

[135] Tokson, *supra* note 13, at 174-75.

[136] Tokson, *supra* note 13, at 174. *See also supra* Parts (I) (A) (2) & (3).

[137] Tokson, *supra* note 13, at 174. *See also supra* Parts (I) (A) (2) & (3).

what the reasonable defendant expects undermines the very premise of the reasonable expectation of privacy test – that it is possible for a judge to correctly discern societal privacy norms and adjudicate defendants' Constitutional claims accordingly.

II. THE ROLE OF CONTEXT

An important thread unifying the flaws in consumer and criminal privacy governance is widespread failure to sufficiently account for the role of context in how individuals consider a given privacy scenario. Context greatly impacts how individuals consider privacy decisions, from deciding to skip over a privacy policy to sending an incriminating letter or email. It also changes the value and accuracy of opinion survey results, depending on whether the survey questions take into account how differently people think about their privacy under different circumstances. Privacy decisions tend to be the product of a wide range of factors, such as the relationship between the discloser and the recipient, perceived risk of disclosure, the location of the disclosure, and the information disclosed, as well as the cognitive factors that might affect a given decision.

The bedrock theory asserting the relationship between context and privacy is Helen Nissenbaum's theory of contextual integrity, which states that the key to effectively evaluating privacy norms and creating rules that adhere to individual expectations must consider the context of a given flow of information.[138] She defines privacy as the appropriate flow of personal information, rather than a right to secrecy, obscurity, or control.[139] What is appropriate in a given scenario is defined as adherence to informational norms, which are measured by the contextual integrity framework. A privacy event is analyzed on two levels, normative and descriptive.[140] For the descriptive layer, the privacy scenario is broken down into four basic elements – actors, contexts, attribute (type of information), and transmission principles.[141] If the norm described by the scenario does not conform with established norms of appropriate information flows, the normative layer analyzes the change to determine the social and moral impact, and whether the new norm would better serve the underlying values and objectives of the social context.[142]

[138] *See* Nissenbaum, *supra* note 49, at 127.

[139] *Id.* at 127.

[140] *Id.* at 127-28.

[141] *Id.* at 127. Transmission principles are the "constraint[s] on the flow … of information," including knowledge, permission, or a breach of a contract. *See id.* at 145.

[142]*See id.* at 150.

Contextual integrity builds on a historical and philosophical acknowledgment that in their day-to-day lives, a person often occupies multiple roles and exists in different spheres and contexts; no man is a monolith, and our expectations, decisions, and behavior are shaped by the particular sphere in which our actions occur.[143] By breaking a privacy decision down to its constitutive elements, it is possible to trace the individualized factors that impact a given privacy decision, such as cognitive limitations, the source of the information, or the environment in which the decision is made. Contextual analysis provides the nuance that an accurate evaluation of norms demands, rather than the clunky conjecture presumed by the privacy paradox.[144]

In consumer privacy law, this could mean evaluating the contextual factors – such as type of data collected, the identity of the collector, the reason for collection, the location of the collection – to consider why the consumer made the choices she made. The context of privacy decision-making could also be considered more broadly to encompass not just the circumstances of the decision, but additional factors, such as the cognitive limitations that shape decision-making and informational asymmetries. In criminal law, this means both evaluating the impact of various contextual factors in a given privacy scenario (the kind of information, the recipient, and so forth), but also evaluating the impact of the new contextual factors that the use of technology introduces to old doctrines (for instance, how the real time dossier of your movements created by a cell phone is distinguishable from the markings on a physical envelope). What changes about the circumstances of a privacy action when a transmission of information is in a written letter, as opposed to an email? Courts would say that while a person has a reasonable expectation of privacy in a list of the people she corresponded with through the mail, the same is not true for the email addresses, as the emails passed through an Internet service provider.[145] The majority of the population would likely feel differently.[146]

Understanding privacy norms within their underlying context is an enormously important part of interpreting them accurately. Analysis of a person's privacy decision-making cannot be mindlessly

[143]*See id.* at 130. *See also* Schaub et al., *supra* note 86, at 3-4 (describing the social and historical predicate for a contextual understanding of privacy).

[144] *See* Martin & Nissenbaum, *supra* note 48, at 190-92 (noting that contextual integrity requires that multiple parameters are equally impactful on a given privacy decision).

[145] *See* United States v. Forrester, 512 F. 3d 500, 510 (9th Cir. 2008).

[146] *See* Scott-Hayward et al., *supra* note 21, at 54 (claiming nearly 90% of survey respondents believed that law enforcement should be prevented from accessing email addresses absent probable cause).

extrapolated beyond the circumstances that created it, particularly when that analysis would shape the development of future law or trigger the liability of existing ones. Trading your ability to shield your location from an application may be worth an hour of free WiFi; but this is a fact-contingent calculation, which must be considered in light of the factual circumstances that shape it. More accurate research on privacy norms could enable policymakers to counter the sweeping assertions based on conjectural tautology with a more systematized approach, based on empirical evidence of the contextual expectations and preferences that privacy critics so often purport to be able to divine.

III. EMPIRICAL PRIVACY RESEARCH

To better understand privacy norms, judges and policymakers need to understand the inherent limitations impacting privacy decision-making, and to examine the context surrounding those choices at a more granular level than they currently do. Empirical research already informs consumer privacy policymaking, but different forms of research that address the role of context, considers the cognitive and practical limitations of privacy decision-making, and differentiates among privacy preferences, expectations, and behavior will provide more accurate insights for policymakers than opinion surveys that ignore the impact of context on how individuals consider privacy decisions, or assume the existence of a perfectly informed, economically rational decision-maker. Empirical research that attempts to account for the flaws in how we currently measure privacy norms in law and policy, such as cognitive limitations and information asymmetries, can guide judges and policymakers in their application of current law and shape the development of better ones in the future. In some cases, empirical research may assist consumers in making privacy choices that adhere to their expectations, such as through managing the volume of decisions or the way privacy policies are worded. A more granular analysis of privacy decision-making can help explain the divergence between stated privacy preferences and privacy behavior that blunter standards fail to grasp. Careful, detailed analysis can serve as a critical counterpoint to the facile argument that the contradictions between privacy preference and observed behavior compel the conclusion that the preference is meaningless, or a deliberate repudiation of the preference.

Much of the new research exploring these ideas relies on various artificial intelligence (AI) techniques, such as the use of machine learning to create a predictive model of privacy norms or guiding a user’s choices according to their previous ones. General AI, the hazy finish line of omnipotent, human-level cognition that has driven the research field for decades, is an unknown distance from current capabilities – an omniscient Skynet that can seamlessly manage our privacy concerns for us remains the stuff of science fiction.[147] But narrow AI, autonomous or semi-autonomous processes that can complete tasks within a certain scope, such as image or voice recognition, has been steadily and rapidly improving, and has the potential to help coalesce blunt privacy standards with the realities of multi-dimensional privacy decision-making.[148] Various types of AI are differentiated by the type of logic they employ to accomplish a given task.[149] Machine learning algorithms improve through iteration on a set of training data, and then based on the connections between features in the data, designate their own rules for classifying a given input.[150] Natural language processing refers to a host of machine-learning techniques that interpret text using statistical inference or the use of neural networks, which are inspired by the knowledge trajectory framework of the human neural system.[151] While not generally considered a form of artificial intelligence, another technique that has been often used in tandem with AI-focused privacy research is crowdsourcing, which uses human expertise, such as crowd-workers answering survey questions or labeling training data, to expand what automated processes can accomplish.[152] Current examples include Wikipedia, which is driven by the information provided by human

[147] STAN. U., ARTIFICIAL INTELLIGENCE AND LIFE IN 2030 at 4-5 (2016), https://perma.cc/8ZS2-SP47 (hereinafter, *Stanford Report*).

[148] ALEX CAMPOLO, MADELYN SANFILIPPO, MEREDITH WHITTAKER & KATE CRAWFORD, AI NOW 2017 REPORT 3 (2017), https://perma.cc/96MM-2K9R ("[R]eal-world applications [of AI] have only accelerated in the last decade due to … better algorithms, increases in networked computing power and the tech industry's ability to capture and store massive amounts of data.").

[149] *See generally id.* (including neural networks, Naïve Bayes classifiers, decision trees, and logistic regression).

[150] *See* Stan. U., *Machine Learning*, COURSERA, https://perma.cc/XNG7-G2VS (last visited Sep. 12, 2018) ("Machine learning is the science of getting computers to act without being explicitly programmed."); Nikki Castle, *An Introduction to Machine Learning Algorithms*, DATA SCIENCE (June 29, 2017), https://perma.cc/BQ2K-2EPG (explaining machine learning and providing an overview of random forests, neural networks, logistic regression, and kernel methods.); Jenna Burrell, *How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms*, BIG DATA & SOC'Y, Jan.–June 2016, at 5.

[151] *See generally* Peng Lai Li, *Natural Language Processing*, 1 GEO. TECH. L. REV. 98 (2016); YOAV GOLDBERG, NEURAL NETWORK METHODS IN NATURAL LANGUAGE PROCESSING at xvii (2017) ("Natural language processing (NLP) is a collective term referring to automatic computational processing of human languages. This includes both algorithms that take human-produced text as input, and algorithms that produce natural looking text as outputs."); *Stanford Report*, *supra* note 147, at 14.

[152] *Cf. Stanford Report*, *supra* note 147, at 9 (defining crowdsourcing as a form of artificial intelligence that "investigates methods to augment computer systems by making automated calls to human expertise to solve problems that computers alone cannot solve well").

editors, or the Amazon Mechanical Turk system, which provides access to human participants for a given task (such as answering survey questions) in exchange for a small fee.[153] Crowdsourcing provides a relatively rich and inexpensive source of data to populate the training sets computer scientists can then use to train predictive models.

The value of this new empirical privacy research is that it can provide a more accurate basis for privacy law and policy by explaining why people make the privacy decisions they make, after taking into account decision-making flaws, the impact of context, and the distinctions between preferences, expectations, and behavior. While opinion surveys have been a common tool of privacy law and policy, they often fail to capture a complete picture, fueling anti-privacy narratives like the privacy paradox's conclusion that anti-privacy actions must speak louder than pro-privacy words. For example, in a recent paper, Helen Nissenbaum and Kirsten Martin dissect the methodology of privacy opinion surveys, and demonstrate how the fact that individuals modulate their privacy behavior to contextual factors contradicts previous results that fueled the narrative of the privacy paradox.[154] But research that designs around the flaws of past methodologies may be able to eclipse those failures, such as more granular opinion surveys, or methodologies that measure behavior directly. In addition, crowdsourcing can be used to facilitate broader opinion surveys. When trying to ascertain a descriptive understanding of existing norms – for example, whether most people on average would assume that they have an expectation of privacy in their cell site location information – opinion surveys are still enormously valuable, and crowdsourcing may assist in expanding their reach, particularly as response rates for traditional landline phone surveys continue to plummet, impacting the quality of the available data.[155] Surveys are not inherently flawed as a methodology, provided they do not rely on the

[153] *See Stanford Report*, *supra* note 147, at 16. *See also* James Vincent, *Mozilla Is Crowdsourcing Voice Recognition to Make AI Work for the People*, VERGE (July 7, 2017, 10:17 AM EDT), https://perma.cc/78JC-YC9E.

[154] Martin & Nissenbaum, *supra* note 48, at 180-81 ("[A] nuanced view of privacy is able to explain away a great deal of what skeptics claim is a divergence of behavior from stated preference and opinion . . . . When respondents are given a chance to offer more fine-grained judgments about specific information-sharing situations, these judgments are quite nuanced.").

[155] *See* Ben Casselman, *Are Wage Gains Picking Up? Stalling? Questionable Data Makes It Hard to Say*, N.Y.T. (March 12, 2018), https://www.nytimes.com/2018/03/12/business/economy/wage-data.html (noting that "Americans are increasingly refusing to respond to government surveys" on employment and that "similar problems have affected other government and private-sector surveys" are resulting in diminished quality of data available to study); Kevin D. Haggerty & Amber Gazso, *The Public Politics of Opinion Research on Surveillance and Privacy*, SURVEILLANCE & SOC'Y, 174-5 (2005) (discussing falling response rates, study fatigue, and the resulting impact on survey data quality).

same misplaced assumptions about privacy that have characterized previous ones, such as conflating preferences and expectations, assuming the respondent is deeply informed about the internet ecosystem, or divorcing broad privacy value statements from situational context. Instead, these new methods acknowledge the limits of asking someone broad questions about how they value their privacy writ large. They also design around the prior flaws of previous survey methods by focusing on the context of privacy decisions and parsing the *kinds* of decisions being made in a given scenario (whether a true preference, a resigned expectation, or an action impacted by lack of information or decision fatigue).

I have divided the current research into descriptive research, the primary objective of which is to provide accurate information on privacy expectations and norms; and operative tools, which are designed to spur user privacy behavior that is more consistent with user expectations. Many of these studies fall into both categories, such as algorithms that model user preferences in order to provide an operative tool,[156] or a project focused on using semi-automation to streamline reading privacy policies, which intends to observe trends in common mismatches between policy commitments and user expectations.[157] The taxonomy is for simplicity.

A. Descriptive Research

Descriptive empirical privacy research – research that uses a range of methods to isolate the factors that impact individual privacy preferences and expectations, as well as illustrating broader norms – could be enormously helpful for policymakers and judges as they attempt to better understand the inconsistencies and ambiguities of modern privacy behavior and beliefs.

In a recent paper, a group of researchers from New York University and Princeton University used crowdsourcing and other methods to automate the discovery of privacy norms.[158] The researchers used Helen Nissenbaum's five-element contextual integrity framework – sender, subject, attribute, recipient, transmission principle – to capture norms that applied to variations of scenarios in an educational setting.[159] They automatically generated 1,411 questions

[156] Tsai et al., *supra* note 73, at 145-46.

[157] Lin et al., *supra* note 70, at 204.

[158] YAN SHVARTZSHNAIDER, SCHRASING TONG, THOMAS WIES, PAULA KIFT, HELEN NISSENBAUM, LAKSHMINARAYANAN SUBRAMANIAN & PRATEEK MITTAL, LEARNING PRIVACY EXPECTATIONS BY CROWDSOURCING CONTEXTUAL INFORMATIONAL NORMS 3 (2016), https://perma.cc/CXA5-6VYU.

[159] Example roles included students, professors, teaching assistants, and the registrar; attributes

using the framework, where a given question might be the following: "Is it acceptable for the Professor (sender) to share the student (subject)'s grades (attribute) with the class (recipient) with the student's permission (transmission principle)?"[160] The participants could respond with "yes," "no," or "does not make sense" (to remove the automatically generated scenarios that would be unlikely to occur in the real world, such as a student sending a teacher a TA's grades).[161] The researchers developed three metrics to gauge how the survey group perceived a given question; a norm-approval score, a use-approval score, and a divergence score.[162] The first score measures whether a norm was approved by the respondents, such as a 50% or 66% overall approval score. The second measures how many norms were approved by a given respondent, and the third measures the variance of one respondent's answers from the answers of all the others – i.e., their dissatisfaction with the established norm.[163] The researchers then encoded the privacy norms established by the respondents' answers into formal logic and used automated theorem provers to check the norms for semantic and transitive consistency.[164] One widely approved norm was a professor sending graduate schools a student's attendance, with her permission; a widely disapproved norm was a TA sending the class a student's grades if the student was performing poorly.[165]

The researchers were thus able discover population consensus about a range of privacy interactions, and test the strength, stability, and consistency of the norms established. They also note that, having derived a method for encoding norms that reflect broad consensus about privacy into formal logic, the framework they built could be amplified through the use of machine learning.[166] Namely, having created the formal framework to capture and express granular privacy norms in a way that can be used to train a predictive model, that model could predict additional norms involving different scenarios – such as

included grades, transcript, name, email, level of participation, and photo; and transmission principles included knowledge, permission, or a triggering event on breach of contract. *Id.* at 3.

[160] *Id.* at 3.

[161] *Id.* at 3-4.

[162] *Id.* at 4.

[163] *Id.* at 4.

[164] A norm had semantic consistency if the information flow that causes a norm to be disapproved is excluded from the other norms that have been approved. Transitive flow consistency entails that the property defined by one norm should be present in similar scenarios – a violation of transitive flow consistency demonstrates a clash between the participant's expectations and the logical consequence of that expectation. For example, a participant stating that she never shares electronic information with strangers, and subsequently stating that she uses email would violate transitive flow consistency. *Id.* at 7.

[165] *Id.* at 6 tbl.2.

[166] *Id.* at 2.

what consumers think about a particular type of tracking an agency is considering regulating, or how people distinguish among different kinds of sensitive information that are sent through a third-party service provider.

Another study at the University of California, Irvine built on similar ideas. The researchers built a machine-learning model to predict participants' future privacy preferences based on participants' reactions to various privacy scenarios in real time involving a connected device.[167] The researchers developed an app for Google Glass where the glasses would display a given IoT scenario based on the participant's location, and he or she would then respond to corresponding survey questions about their preferred privacy outcomes in the app. Similar to the contextual integrity framework, the questions broke down a privacy interaction into contextual parameters – where the monitoring occurs, what is monitored, who is monitoring, why the monitoring is taking place, and the frequency of the monitoring. They also asked about participants' desire to be notified, their willingness to accept the monitoring, and their level of comfort, associated risk, and appropriateness of the monitoring involved.[168] Having the participants answer questions at the locations they were being asked to evaluate through projecting the scenarios onto Google Glass has the added value of giving participants a more realistic and less abstract reminder of how they might truly react to each scenario in the real world.

Results included a range of preferences, such as the fact that participants would not allow videotaping of their movements without a clear purpose, and that they found still photography to be relatively more acceptable, but were still concerned about still photography with a purpose if the purpose is to determine a specific characteristic.[169] Monitoring for safety or social reasons (such as a service recommending a friend) was considered invasive, whereas monitoring for health purposes was more likely to be deemed acceptable.[170] The researchers then used a clustering algorithm to determine the impact of the different contextual factors (such as location, basis for monitoring, and so forth), and trained a machine learning classifier on the survey responses to predict user privacy preferences based on context (where, what, who, the reason for monitoring, and the persistence thereof).[171] The model was 77% accurate in predicting a binary privacy decision

[167] *See* Lee & Kobsa, *supra* note 46, at 1-2.

[168] *Id.* at 3, 4 tbl.1.

[169] *Id.* at 7.

[170] *Id.* at 7.

[171] *Id.* at 3.

(simply accepting the monitoring or rejecting it), and the researchers' use of an interpretable method of machine learning meant that they could determine why the model made the decisions it did.[172]

Another study also examined privacy preference modeling for IoT. The researchers used crowdsourcing to show participants context-based privacy vignettes design to measure the impact of eight factors on privacy decisions – type of data collected, the location where the data was collected, whether the user benefits from the data collection, the device that collects the data, the purpose of collection, retention time, whether the data is shared, and whether additional information could be inferred.[173] After being shown the vignette, participants were asked them how often they would want their phones to notify them about that type of collection, how comfortable they were with that data collection, and whether they would allow or deny the collection.[174] The researchers then used statistical regression to determine the significance of various factors, and used the data to train two machine-learning models, one to predict the individual's comfort level with a particular data collection scenario, and the other to predict what the individual's ultimate decision to allow or deny data collection would be.[175] The researchers were thus able to analyze which factors were the most influential for both comfort level and ultimate data collection decision.[176] As comfort level and collection decision were separate prediction models, the researchers could parse when comfort level and desire to permit or deny collection were distinct, and found that a high comfort level for a given collection scenario did not always mean that the user would choose to have their data collected in that scenario, if given the choice.[177] The study also found that users were more likely to accept collection when their data was being put to "beneficial" use, either for their own personal benefit, or for the "greater good."[178] The collection decision and comfort level models were 76-80%[179] and 81%

[172] *Id.* at 8.

[173] Emami-Naeini et al., *supra* note 46, at 401.

[174] *Id.* at 401.

[175] *Id.* at 402.

[176] *Id.* at 402 tbl.1 (describing the factors influencing comfort level); *id.* at 406 tbl.7 (describing the factors influencing grant or denial of permission).

[177] *Id.* at 406-7 ("For example, we found that the interaction between data type and location was the most helpful factor in the allow/deny model, but this factor was shown to be non-significant in explaining the comfort level. This suggests that being comfortable with a specific data collection instance does not automatically mean that someone would allow it to occur, given the choice.").

[178] *Id.* at 410.

[179] *Id.* at 406 tbl.6.

accurate, respectively.[180] These results are helpful for separating the factors and result of privacy decision-making that are often assumed, as well as illuminating which ones are more important than others.

B. Operative Research

Other researchers have focused on using different artificial intelligence techniques to guide user behavior in a more privacy-protective way, such as through providing the information necessary for coherent privacy decision-making in a more understandable form, or by mitigating the effects of decision fatigue and other cognitive phenomena. This includes different approaches to streamlining the process of reading privacy policies, and modeling user preferences to build predictive algorithms that nudge the user's privacy decisions accordingly.

One group investing considerable efforts towards researching and building privacy-protective user tools is the Useable Privacy Project, an ongoing research collaboration at Carnegie Mellon University.[181] One aspect of their work focuses on developing techniques using natural language processing, crowdsourcing, and machine learning to automate (or semi-automate) the process of reading a privacy policy, in order to provide users with better information, mitigate the cognitive biases that hinder privacy decision-making, and build a database of privacy policy trends to illustrate where expectation mismatches are most prevalent, what common policy clauses are the most misleading for readers, and other impediments to coherent privacy decision-making.[182] In one study, the researchers examined the mismatch between users' privacy expectations online, and what the privacy policy of a given website actually permitted the company to do.[183] The

[180] *Id.* at 405.

[181] NORMAN SADEH, ALESSANDRO ACQUISTI, TRAVIS D. BREAUX, LORRIE FAITH CRANOR, ALEECIA M. MCDONALD, JOEL R. REIDENBERG, NOAH A. SMITH, FEI LIU, N. CAMERON RUSSELL, FLORIAN SCHAUB & SHOMIR WILSON, THE USABLE PRIVACY POLICY PROJECT: COMBINING CROWDSOURCING, MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING TO SEMI-AUTOMATICALLY ANSWER THOSE PRIVACY QUESTIONS USERS CARE ABOUT 3-5 (Dec. 2013), https://perma.cc/U75U-CMZN (describing the goal of the project as "to develop, evaluate and deploy new technologies in the context of a novel, practical framework that empowers users to more meaningfully control their privacy without any additional cooperation from website operators other than the natural language privacy policies that they already have in place," including semi-automated understanding of privacy policies, privacy preference modeling for usable privacy disclosures, mitigating deleterious and cognitive behavioral biases in privacy disclosures, and privacy policy analysis).

[182] *See Towards Effective Web Privacy Notice and Choice*, USABLE PRIVACY, https://perma.cc/SD4P-QJ99 (last visited October 20, 2018).

[183] Ashwini Rao, Florian Schaub, Norman Sadeh, Alessandro Acquisti & Ruogu Kang, *Expecting the Unexpected: Understanding Mismatched Privacy Expectations Online*, 2016 SYMP. ON

researchers similarly found that a range of highly contextual factors impacted what participants' expectations were and why, including website type (health, financial, or dictionary) and user characteristics – things like privacy knowledge, privacy concept familiarity, age, whether or not the participant had an account with the website, or had used it recently.[184] Isolating the instances where privacy policies are the most likely to contravene privacy expectations could be enormously helpful in reducing decision fatigue in users, by highlighting only the portions of a given policy that they will need to read.[185]

In another study, another group of researchers built a classifier to automatically locate provisions of choice in privacy policies, such as the ability of the user to opt out of collection by clicking on a URL in the policy.[186] As opting out is a concrete, effective action the user can take, it is one of the more significant aspects to tease from a dense privacy policy and bring to the user's attention.[187] Highlighting the most important aspects of privacy decision-making for users, such as when they can opt out, prevents them from sapping their mental energy on cognitive tasks that are less important for their privacy choices, such as reading an entire privacy policy when actually reading it is unlikely to sway the user's decision to use the service. The ultimate goal would be to create a browser plug-in that would identify opt-out hyperlinks to the user.[188] The researchers also relied on manual annotation of privacy policies to build logic representations for future automation, and crowdsourced annotations to train a machine-learning classifier to identify when a policy would be clear to human users.[189] Another project, Polisis, uses natural language processing and neural-network classifiers to power a browser plug-in that breaks down the privacy policy of a given website for the user.[190]

In yet another study, researchers at the University of California, Berkeley used machine learning to predict how a user would respond to a privacy decision – namely, an application permission request – to predict their response to another privacy decision.[191] The researchers

USABLE PRIVACY & SECURITY 77, 77.

[184] *Id.* at 85-86.

[185] *Id.* at 87; Sadeh et al., *supra* note 181, at 3.

[186] *See* Sathyendra et al, *supra* note 69, at 2775.

[187] *See id.* at 2775.

[188] *See id.* at 2778.

[189] *See* Sadeh et al., *supra* note 181, at 8.

[190] Hamza Harkous, Kassem Fawaz, Rémi Lebret, Florian Schaub, Kang G. Shin & Karl Aberer, *Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning*, 2018 SYMP. ON USABLE PRIVACY & SECURITY 531, 531-32.

[191] Primal Wijesekera, Arjun Baokar, Lynn Tsai, Joel Reardon, Serge Egelman, David Wagner &

first measured the difference in accuracy of the user's decision, (whether the permission granted or denied corresponded with the user's expressed preference, when compared with what the privacy policy accomplishes) when users were given the permissions decision on first using the app, rather than on first installing it.[192] The idea here was that an ask-on-first-use permission would be made when the user had greater context for the decision, rather than at the moment the user downloaded the app.[193] The hypothesis appeared correct, as the ask-on-first-use decisions were 84% accurate, while the ask-on-first-install decisions were 25% accurate.[194] They then built a classifier that incorporates passively observed information relating to the individual's privacy preferences (behavioral information, responses to runtime information, and responses to permissions requests[195]), and reports the prediction and a confidence score, which would determine whether or not the user was prompted to confirm the preference selected by the algorithm.[196] The researchers' goal was to eliminate the volume of decisions a user has to make in addition to aligning the decisions more closely with the user's stated preferences.[197] By using past decision history and other information, the algorithm could infer when the privacy decision was aberrant from decisions the person made in the past, and automatically make the decisions where the algorithm reported a high confidence in its prediction.[198] This prevents the user from being overwhelmed with low-impact choices she is unlikely to care about, while preventing her lack of information from skewing the choice she made, and still giving her the opportunity to make the harder choices that she would want to make herself.

In a similar experiment, researchers tested an Android permissions manager, Turtleguard, which balanced the dual objectives of grounding each privacy decision in context and while preventing the user from being overwhelmed by the volume of decisions she needs to

Konstantin Beznosov, *The Feasibility of Dynamically Granted Permissions: Aligning Mobile Privacy with User Preferences*, 2017 INST. ELECTRICAL & ELECTRONICS ENGINEERS SYMP. ON PRIVACY & SECURITY 1077, 1078.

[192] *See id*. at 1080.

[193] *Id.* at 1077.

[194] *Id.* at 1078.

[195] *Id.* at 1080 tbl.2. Runtime information included when a platform switched to a new activity; permissions requests included when an app ask for a sensitive permission; and behavioral information included a wide range of factors, such as the user changing security settings, their use of two-factor authentication, enabling or disabling the speakerphone, or changing the developer options.

[196] *Id.* at 1078.

[197] *Id.* at 1077-78.

[198] *Id.* at 1077-78.

make.[199] Turtleguard uses a classifier to respond to the permissions request based on whether the requesting application was actually being used at the time of the request, which the researchers found was a reliable indicator of whether participants thought a permission request was invasive or not.[200] These decisions are logged in the Turtleguard portal, and can be audited by the user to calibrate the classifier for future requests. Another study trained a classifier to discern when a user would be comfortable being tracked by an advertiser, based on the properties of the web page, the user's demographics, and attitude towards tracking.[201] While this study was preliminary and on a smaller scale, it demonstrates another area in which users display finely-grained, contextual reasons for the privacy decisions they make, and that those reasons can be inferred and used by an algorithm to finesse privacy decision-making.

Together, these studies point to the possibility of using empirical methods to gain new insights into privacy preferences, expectations, behaviors, and norms, and the factors that influence each. Modeling crowdsourced norms, like the studies using contextual integrity vignettes[202] or the research on IoT preference modeling,[203] can be enormously helpful in eliciting why people make the privacy decisions they do. Contextual modeling could be used as a baseline to map the average reasonable person's expectation of privacy in a range of scenarios, particularly in cases of emerging technology where judges or regulators have a limited basis of comparison. With machine learning, studies that prioritize explainable methods – such as the study that used clustering to identify important factors on participants' privacy perceptions,[204] or the study that isolated comfort level and desire to permit tracking[205] – can help illustrate for regulators what kinds of practices users generally consider deceptive or unfair.[206] Proactive tools, like semi-automated privacy policy annotations and nudging privacy assistants, can help counteract problems like information asymmetry and decision fatigue, and empower users to make coherent privacy decisions that suit their expectations. In turn, tools that better enable individuals to make privacy choices that reflect their preferences will ideally help researchers and regulators to more

[199] *See* Tsai et al., *supra* note 73, at 145-46.

[200] *Id.* at 148.

[201] *See* Melicher et al., *supra* note 75, at 136.

[202] *See, e.g.*, Shvartzshnaider et al., *supra* note 158, at 2-3.

[203] *See, e.g.*, Emami-Naeini et al., *supra* note 46, at 401.

[204] *See* Lee & Kobsa, *supra* note 46.

[205] *See* Melicher et al., *supra* note 75, at 136.

[206] *See* Emami-Naeini et al., *supra* note 46, at 401.

accurately infer privacy preference preferences and expectations from their privacy behavior. An ideal world isn't the one in which we happen to live, but absent an overhaul of the entire notice and choice system, creative tools that help counteract the flaws of the user control model can at least help users in their seemingly impossible quest to obtain effective notice and meaningful choice. The following section will detail how both operative and descriptive research can be implemented and incentivized in the development and application of privacy law.

IV. APPLICATIONS

The value of empirical research in privacy law and policymaking would, appropriately enough, depend on the context in which it is used. Courts, regulatory agencies, and legislatures have different legal, practical, and philosophical limitations on the kinds of evidence that they can rely on as the basis for decision-making. Due to the different roles and histories of these institutions, policymakers' reactions to the value of empirical research will depend on their role, and consequently, so will the viability of these proposals.[207] The following section will discuss the value of different types of empirical privacy research for different areas of privacy law and policy.

A. Descriptive Research & Fourth Amendment

Modeling crowdsourced norms could provide a more accurate and empirical counterpoint to the privacy expectations that judges attempt to approximate, as well as demonstrate the instances when the implications of new technology diverge sharply enough from the logical underpinnings of precedent to merit a new approach. Many scholars have called for judges to anchor their legally determinative assessments of societal facts in empiricism, particularly in Fourth Amendment cases.[208] In one of the earliest such appeals, Christopher Slobogin and Joseph Schumacher conducted a survey of 217 people to determine how the average person actually perceives encounters that

[207] *See* J. Alexander Tanford, *Law Reform by Courts, Legislatures, and Commissions Following Empirical Research on Jury Instructions*, 25 L. & SOC'Y REV. 155, 156 (1991) (detailing the literature arguing that the impact of empirical research on public policy decision-making depends on the particular institutional role of the policymaker).

[208] *See, e.g.*, *Carpenter Empirical Brief*, *supra* note 133, at 1-2 (arguing that empirical evidence of people's privacy expectations should inform judicial analysis of the Fourth Amendment, and detailing past calls for an empirical approach and data illustrating the discrepancy between privacy expectations as reported by judges and precedent, and the privacy expectations reported in the survey results). *See generally* Tracey L. Meares, *Three Objections to the Use of Empiricism in Criminal Law and Procedure – And Three Answers*, 2002 U. ILL. L. REV. 851, 852-53 ("That empiricism is relevant to criminal law and criminal procedure is a point so obvious that it seems almost banal.").

implicate a reasonable expectation of privacy, and ultimately discovered a wide range of situations where the doctrine reflected assumptions contrary to the norms participants reported, particularly a lower expectation of privacy in a range of scenarios.[209] Lior Strahilevitz has advocated for the reasonable expectation of privacy in tort cases to consider an empirical evaluation of how information travels through social networks, ideally providing a sense of how likely certain information was to be made public based on the structure of the network and cultural variables.[210] Paul Ohm proposed that equilibrium adjustment theory be anchored in statistical measurements of how technology has shifted the balance between observer and observed, such as through the length of investigations, or number of indictments.[211] Andrew Ferguson has observed that location-based predictive policing algorithms could produce empirical data on what makes an area "high-crime" for Fourth Amendment purposes, possibly reducing the ability of the capacious term to be used as a stand-in for disadvantaged or minority neighborhoods.[212] Perhaps the most relevant to the present analysis is Matthew Tokson's empirical examination of the role of knowledge in the reasonable expectation of privacy test, in which he argues that anchoring the test to individual or societal knowledge fundamentally prevents the Fourth Amendment from evolving in step with technology.[213] Empirical research on the kinds of throwaway assumptions that often guide a judge's decision in a Fourth Amendment case – such as how common a certain kind of search is, or under what circumstances most people would feel able to terminate an encounter with the police – could play an enormous role in anchoring amorphous, norm-based questions to facts, a crucial development when radical technological developments make those norms difficult to gauge, and quick to shift.

Empirical research can also be used to render the fuzzy judicial approximations of privacy more accurate, as most judges are limited in their understanding of what the expectations of the "average" person are.[214] Generalized assumptions that judges may hold about societal

[209] Slobogin & Schumacher, *supra* note 11, at 732, 733-35 (describing hypothesis and sample population).

[210] *See* Lior Jacob Strahilevitz, *A Social Networks Theory of Privacy*, 72 U. CH. L. REV. 919, 970-71 (2005).

[211] *See* Paul Ohm, *The Fourth Amendment in a World without Privacy*, 81 MISS. L. J. 1309, 1313, 1352-53 (2012).

[212] *See* Andrew Ferguson, *The "High-Crime Area" Question: Requiring Verifiable and Quantifiable Evidence for Fourth Amendment Reasonable Suspicion Analysis*, 57 AM. UNIV. L. REV 1587, 1593-94 (2008)

[213] Tokson, *supra* note 13, at 139.

[214] *See* Strahilevitz, *supra* note 210, at 920-21 (noting that the judiciary "cannot agree on a

norms are narrowed by the range of experiences they themselves have held, which typically do not reflect the experience of the defendants before them.[215] To give a recent example, in the oral arguments for *Carpenter v. United States,* Justice Kennedy inquired as to how well the average individual knows what the retention and access policies of a cell service provider are, joking that "if I know [that a cell phone company has my data], everyone does."[216] The Justice was attempting to poke fun at himself by indicating that if someone as technologically unsavvy as himself is aware of cell phone company practices, the rest of the population must be. But in reality, the converse is more often true; a judge's experience is, unsurprisingly, more often a higher bar than that of the average person. The judiciary is older, more urban, better-educated and wealthier than most of the population,[217] and this demographic skew likely has a correspondingly transformative effect on how judges conceptualize what "reasonable" is, as they are more likely to be aware of surveillance practices than defendants, and unlikely to recognize that their perspective is unrepresentative.[218] Indeed, Tokson's study found that the majority of cell phone users do not know that their location is being tracked and collected, and 15% believed that no collection occurs at all.[219]

framework" for evaluating privacy norms in tort cases); Scott-Hayward et al., *supra* note 21, at 49 (describing the results of previous, similar studies and opinion polls reflecting that "the public has higher expectations of privacy than those recognized by the courts in most Fourth Amendment jurisprudence"); Jonathan Simon, Katz *at Forty: A Sociological Jurisprudence Whose Time Has Come*, 41 U. C. DAVIS L. REV. 935, 948 (2008) (arguing that the lack of substantive guidance shaping how a reasonable expectation of privacy creates "an invitation to judicial 'self-indulgence' in declaring what society is in fact prepared to recognize as reasonable.") (citing Richard Posner, *The Uncertain Protection of Privacy by the Supreme Court*, 1979 SUP. CT. REV. 173, 187); McAllister, *supra* note 129, at 31 (arguing that empirical evidence is a better indicator of societal expectations of privacy, and rather than reasoning by often-inapt analogy, courts should simply "ask society").

[215] *See generally* David L. Faigman, *"Normative Constitutional Fact-Finding": Exploring the Empirical Component of Constitutional Interpretation*, 139 U. PENN. L. REV., 541, 545 (noting that "[h]istorically, most constitutional fact-finding depended on the Justices' best guess about the matter," such as Chief Justice Marshall's observation in *Gibbons v. Ogden* that "[a]ll America understands, and has uniformly understood, the word 'commerce,' to comprehend navigation") (citations omitted); Erik Luna, *The* Katz *Jury*, 41 U. C. Davis L. Rev. 839, 846-7, 849 (2008) (arguing that "[the Court's] decisions on Fourth Amendment applicability do not seem to comport with the expectations of privacy held by the common citizen").

[216] Transcript of Oral Argument at 27, *Carpenter*, 585 U.S. ___ (No. 16-402) (2018)

[217] *See* Chao et al., *supra* note 13, at 290.

[218] *See Carpenter Empirical Brief*, *supra* note 133, at 11-12 (noting the statistical relationship between age and receptiveness to the third-party doctrine); Tokson, *supra* note 13, at 166 (noting that judges are likely to attribute their level of knowledge to defendants, given the tendency of individuals generally to impute their level of knowledge on a subject to others); *id*. at 169-171 (discussing the uneven distribution of knowledge).

[219] Tokson, *supra* note 13, at 177.

Analyzing his 810-person survey on individuals' understanding of cell phone tracking, Tokson notes that the average person's knowledge of cell phone tracking is substantially less than what judges have ruled it is, and the average person's expectation of privacy in that information is higher.[220] Analyzing the results against past judicial holdings on societal expectations of privacy, he further concluded that while judges are frequently correct in ruling on the *doctrinal* perception of privacy norms, they tend to underestimate what most people actually expect, and overestimate what they actually know.[221] Judges have no problem applying the standards the way the law instructs; but the law relies on the idea that the judge's perception of a reasonable expectation of privacy is fungible with that of the average defendant, which is rarely the case. Another study also found that that community consensus on privacy expectations frequently diverged starkly from Supreme Court precedent.[222]

A judge's bias about a given defendant's role in society may also skew outcomes. Avani Mehta Sood has noted that motivated cognition, the unconscious human tendency to reason towards preferred outcomes, may affect judicial applications of the exclusionary rule in Fourth Amendment cases.[223] For example, he found that while in an experiment using lay participants the severity of the crime decreased the likelihood that the participants would allow the information to be suppressed.[224] In testing that hypothesis on judges themselves, he also discovered that some may be conscious of motivated cognition and attempt to mitigate to its effects, while others are simply susceptible to the phenomenon.[225] Using a contextual framework to model informational norms can provide judges with a source of accurate data beyond their own experience for what norms are widely accepted by society, rather than what norms are widely accepted by judges.

In addition to providing judges with a more accurate perception of what the average reasonable expectation of privacy actually is, context-based modeling can also balance Fourth Amendment values in a way that is faithful to the principles of the precedent, while accounting for the decisive impact of new technology. Andrew Selbst

[220] *Id.* at 168-69.

[221] *Id.* at 179.

[222] *See* Scott-Hayward et al., *supra* note 21, at 58 (arguing that "rather than relying on their own perspectives on what 'society' is prepared to accept as reasonable, judges should turn to studies like these, which present the perspectives of members of the general public").

[223] Avani Mehta Sood, *Applying Empirical Psychology to Inform Courtroom Adjudication – Potential Contributions and Challenges*, 130 HARV. L. REV. F. 301, 303 (2017).

[224] *Id.* at 303-4.

[225] *Id.* at 312.

has argued for a contextual-integrity-based approach to Fourth Amendment doctrine, arguing that Nissenbaum's theory of appropriate information flows is better-equipped to respond to the shifting, granular privacy expectations of the modern reasonable person in the digital age.[226] While this article does not go quite so far as to propose replacing Fourth Amendment analysis with contextual integrity altogether, Selbst's prescient proposal demonstrates the value of contextual integrity to a Fourth Amendment inquiry, as well as the role that context-based modeling could play in making blunt legal tests coalesce with the realities of technological norms. Nissenbaum herself has also applied the contextual integrity framework to the Fourth Amendment context, arguing that the Fourth Amendment doctrines concerning metadata – namely, the content/non-content distinction, the public view doctrine, and the third-party doctrine – violate contextual integrity because technology has disrupted the original information flow in a way that is normatively undesirable.[227]

Crucially, context-based models like contextual integrity are deliberately flexible enough to capture the full range of interests that the Fourth Amendment is intended to measure, beyond merely the weight of the privacy expectation being assessed and whether that expectation would be deemed valid by society.[228] The range of scenarios can capture the full spectrum of interests that Fourth Amendment tests are intended to balance, including when warrantless search or seizure implicates a lesser privacy interest (such as consent searches,[229] or evidence in plain view[230]), or where the governmental prerogative is greater (such as search incident to arrest,[231] national security,[232] or heightened danger[233]) while still providing a more accurate report of what privacy expectations are reasonable.

Carpenter demonstrates how valuable contextual modeling could be to help judges assess nuanced privacy norms, particularly as the chasm between the limits of what Fourth Amendment principles purport to protect and what the doctrine actually covers is continually

[226] Selbst, *supra* note 106, at 649-50.

[227] *See generally* Kift & Nissenbaum, *supra* note 114.

[228] For instance, the survey by Scott-Hayward et al., *supra* note 21, at 52-52 tbl.1, respondents described how their feelings about a given surveillance activity would change on five escalating degrees of evidence, from anytime (without any proof), to never (even if probable cause to search existed).

[229] *See* Schneckloth v. Bustamonte, 412 U.S. 218, 243 (1973)

[230] *See* Coolidge v. New Hampshire, 403 U.S. 443, 465 (1971).

[231] *See* Chimel v. California, 395 U.S. 752, 762-63 (1969).

[232] *See* United States v. Ramsey, 431 U.S. 606, 616 (1977) (noting that as the government's national security interest is at its zenith as the border, searches are "per se reasonable").

[233] *See* McDonald v. United States, 335 U.S. 451, 455-56 (1948).

widened by technological change.[234] The Court held that the collection of more than seven days' worth of CSLI by the government without a probable-cause warrant violated the Fourth Amendment, given the revealing nature of the information.[235] The majority "declined to extend" *Smith* and *Miller* to CSLI, while leaving the third party doctrine untouched, if marked for death, as it pertains to any other information beyond the seven-days' worth of CSLI at issue in *Carpenter*.[236] But determining exactly how different types of information that seem to implicate the same kinds of privacy concerns as CSLI should be considered will be a difficult and fact-intensive task for judges. As Justice Kennedy raised in his dissent, the majority's basis for requiring a warrant does not provide much clarity as to what distinguishes the sensitivity of CSLI from financial information held by digital intermediaries, for one example.[237] Chief Justice Roberts claimed that however it is that the Court must assess what constitutes a reasonable expectation of privacy, its lodestars of seeking to secure "the privacies of life" against "arbitrary power" and "plac[ing] obstacles in the way of a too permeating police surveillance"[238] must mean that the Constitution requires the government to obtain a warrant before it can access "detailed, encyclopedic, and effortlessly compiled" geolocation information.[239]

But how will judges distinguish which other information passed through a third party (that is, any information transmitted over the internet or through any kind of digital service provider) provides the "intimate window into a person's life" that the majority held CSLI specifically provides?[240] Fitbit data that shows a person's sleep cycle or sexual habits, transactions on a financial app like Venmo, the times of day that a person adjusts the smart thermostat in their bedroom – each of these could implicate the "familial, political, professional,

[234] Or as Paul Ohm observed, "What might have seemed like a slow and partial degradation of the Fourth Amendment appears instead to be a full evisceration." *See* Ohm, *supra* note 211, at 1311.

[235] Carpenter v. United States, 585 U.S. ___ , No. 16-402, slip op. at 11 (U.S. June 22, 2018).

[236] As Paul Ohm put it, "a moment of silence, please, for the *nearly* departed." *See* Paul Ohm, *The Broad Reach Of Carpenter v. United States*, JUST SECURITY, (June 27, 2018), https://perma.cc/9SDJ-LRW3.

[237] *Carpenter*, slip op. at 2 (Kennedy, J., dissenting) ("Cell-site records, however, are no different from the many other kinds of business records the Government has a lawful right to obtain by compulsory process."); *id.* at 18 (Kennedy, J., dissenting) ("The troves of intimate information the Government can and does obtain using financial records and telephone records dwarfs what can be gathered from cell-site records.").

[238] *Id.* at 6 (majority opinion) (citations omitted).

[239] *Id.* at 10.

[240] *Id.* at 12.

religious, and sexual"[241] associations that the majority felt warranted protection for CSLI but is not directly tied to geolocation, and thus excluded from *Carpenter*'s new warrant requirement. The ability of the government to obtain information about a suspect in hindsight as opposed to the need to surveil prospectively is the same for this kind of non-geolocational data, and it is similarly cheap for companies or law enforcement to gather and store.[242] It is exactly here that contextual modeling could be useful to help judges identify which kinds of information presents particularly sensitive privacy concerns for the average person. Variations of different scenarios could be modeled just as the researchers modeled education privacy vignettes in the NYU and Princeton study.[243] While Chief Justice Roberts describes the case as deciding "only that a warrant is required in the rare case where the suspect has a legitimate privacy interest in records held by a third party,"[244] given how much information passes through third parties by default and how much of that information can implicate the legitimate interests that the Chief Justice himself cites, distinguishing when individuals do have a reasonable expectation of privacy will be a difficult, finnicky task. Context-based modeling could be invaluable in capturing the privacy norms that have left the assumptions of Fourth Amendment doctrine far behind, and which judges can struggle to discern.[245]

Descriptive research on privacy norms can also impact how searches and seizures are conducted before a case is even brought. The Office of Legal Policy in the Department of Justice issues guidance to federal law enforcement and component prosecutors about its policy views on certain matters, such as whether there are sensible reasons not to pursue enforcement in certain areas, or only under certain conditions.[246] For example, the Department issued guidance in 2015 on the use of cell site simulators, directing federal agents to acquire a warrant before a simulator is used, despite the fact that existing

[241] *Id*. at 12 (citations omitted).

[242] *Id.* at 12.

[243] *See* Shvartzshnaider et al., *supra* note 158.

[244] *Id*. at 21.

[245] As Justice Kennedy notes, "Whether credit card records are distinct from bank records; whether payment records from digital wallet applications are distinct from either; whether the electronic bank records available today are distinct from the paper and microfilm records at issue in *Miller*; or whether cell-phone call records are distinct from the home-phone call records at issue in *Smith*, are just a few of the difficult questions that require answers under the Court's novel conception of *Miller* and *Smith*." *Id.* at 21 (Kennedy, J., dissenting).

[246] *See Office of Legal Policy: Functions*, DEPT. OF JUSTICE (Jan. 24, 2018), https://perma.cc/R6P8-846L..

precedent did not clearly require them to do so.[247] The Attorney General releases memoranda to United States attorneys recommending approaches and prosecutorial tools, and the office also testifies to Congress on its recommendations.[248] Empirical research on privacy norms could inform such recommendations, and have a far-reaching impact on how searches and seizures are conducted—and defended—on the ground, impacting cases before they reach the courtroom, or preventing them from being filed at all.

Of course, the use of empirical modeling would never determine what constitutes a reasonable expectation of privacy in lieu of a judge. Even if legal or practicable, such a solution would require a blithe faith in the perfectibility of technology that past mistakes have repeatedly proven to be misguided. But using contextual modeling to examine areas where the juxtaposition of new technology with analog precedent has made the average reasonable expectation of privacy difficult to discern can lend some basis of empirical accuracy to broad judicial assessments of protean technological norms. While the Supreme Court is typically loath to overturn precedent, they have been known to do so when technology distorts the applicability of existing precedent to a sufficient extent.[249] It is also important to acknowledge the uphill battle that proposals to incorporate empirical approaches to judicial decision-making often face. Judges are typically leery to rely on "legislative facts," or factual determinations that are the product of the real world and may inform the legal determinations in a case.[250] This is often for

[247] *See, e.g.,* DEPT. OF JUST., POLICY GUIDANCE: USE OF CELL-SITE SIMULATOR TECHNOLOGY (2015).

[248] *See., e.g.*, OFFICE OF THE ATT'Y GEN., MEM. TO U. S. ATTORNEYS, GUIDANCE REGARDING USE OF CAPITAL PUNISHMENT IN DRUG-RELATED PROSECUTIONS (2018); *Reforming the Electronic Communications Privacy Act Before the S. Comm. on the Judiciary*, 114th Cong. 4, 6 (2015) (statement of Elana Tyrangel, Principal Deputy Assistant Attorney General) (recommending, among other things, that Congress update the statute's 180-day rule and the definition of remote computing service, and clarify other parts of the law).

[249] Brandeis and Warren's seminal law review article making the case for a right to privacy was written in response to the introduction of the mass market camera. *See* Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 195 (1890). Subsequent Supreme Court cases have overruled Fourth Amendment cases when the introduction of a new technology appeared to make the applicability of an old doctrine obsolete. *See* Riley v. California, 134 S. Ct. 2473, 2492 (2014) (holding that the search incident to arrest exception could not be applied to cell phones); Kyllo v. United States, 533 U.S. 27, 40 (2001) (holding that novel surveillance technology "not in the public use" could transform a government investigative action into a search).

[250] *See* FED. R. EVID. 201(a) (advisory committee note) (1972 Proposed Rules). *See also* Allison Orr Larsen, *Judicial Factfinding In an Age Of Rapid Change: Creative Reforms From Abroad*, 130 HARV. L. REV. F. 316, 316 n.4 (2017) ("A 'legislative fact' – a phrase coined by Kenneth Culp Davis in 1942 – is a generalized fact about the world as opposed to a 'whodunit' fact about what happened in any one case.") (citations omitted).

good reason, as without sufficient support from fact-finding entities or unbiased experts, they may rely on those facts to their detriment, and to the detriment of the development of precedent based on accurate conclusions.[251] Some judges seem even more leery to rely on legislative facts resulting from social science in particular.[252] However, the broad distrust of fact formed outside of the record is necessarily limiting. Judges cannot be expected to be a one-stop shop for all given information, particularly in cases involving complex technologies.[253] Others have further noted that reliance on empirical evidence has previously played a key role in Fourth Amendment and other cases, and that both members of the Supreme Court and lower courts have acknowledged its potential.[254]

It should also be noted that the value of using contextual modeling to discern what constitutes a reasonable expectation of privacy depends on the rule governing how that question should be evaluated—a rule that the Supreme Court has yet to categorically provide.[255] While an in-depth treatment of the issue is outside the scope of this article, it does assume that a judge's understanding of what the average person

[251] *See* Ryan Gabrielson, *It's a Fact: Supreme Court Errors Aren't Hard to Find,* PROPUBLICA, (Oct. 17, 2017) https://perma.cc/7BSH-KQU4.

[252] *See* Luna, *supra* note 215 at 848-49 (noting that "[t]he Court has shown no inclination to take a more empirically grounded approach to the preliminary question of search and seizure analysis") (citing Slobogin & Schumacher, *supra* note 11, at 761); Eduardo Bonilla-Silva, *ASA President Eduardo Bonilla-Silva Responds to Chief Justice John Roberts*, AM. SOCIOLOGICAL ASS'N (October 10, 2017), https://perma.cc/Y9UV-SXVD (responding to Justice Roberts' characterization of social science data as "sociological gobbledygook," noting examples of how sociological research has contributed to key legal and policy revelations, such as "clear evidence that separate is not equal . . . modern public opinion polling . . . [and] evidence of gender discrimination in the workplace"). *But see* Simon, *supra* note 214, at 937-38 (discussing the backlash to the sociological basis for *Brown v. Board of Education*, and arguing for the value, and possibility of, a sociological approach to the Fourth Amendment).

[253] *See generally* Larsen, *supra* note 250 (discussing the breadth of highly technical areas that general subject-matter courts are compelled to decide and proposing solutions to bridge that gap of expertise).

[254] *See Carpenter Empirical Brief*, *supra* note 133, at 12.

[255] *See, e.g.,* Carpenter v. United States, 585 U.S. ___ , No. 16-402, slip op. at 5-6 (U.S. June 22, 2018) (noting that "no single rubric definitively resolves which expectations of privacy are entitled to protection."); *id.* slip op at 7 (Gorsuch, J., dissenting), ("We don't even know what [*Katz's*] 'reasonable expectation of privacy' test *is*. Is it supposed to pose an empirical question (what privacy expectations do people *actually* have) or a normative one (what expectations *should* they have)? Either way brings problems.") (emphasis in original). *See also*, Orin S. Kerr, *Four Models of Fourth Amendment Protection,* 60 STAN. L. REV. 503, 503-4 (2007) (stating that "no one seems to know what makes an expectation of privacy constitutionally 'reasonable,'" that the Supreme Court has "repeatedly refused to offer a single test," and arguing that the Court alternatively relies on one or more of four possible models, a flexibility from which the jurisprudence benefits); Chao et al., *supra* note 13, at 273-75 (arguing for a reasonableness test predicated on "ordinary beliefs," and discussing past scholarship and cases that have supported the same).

believes a reasonable expectation of privacy to be is relevant to how the test should be implemented. Under that assumption, the more accurate a judge's understanding of the average person's expectations is, the more likely future Fourth Amendment jurisprudence will avoid decisions that make sense to esteemed legal scholars, but contravene the expectations of the majority of the population, upon which the doctrine purports to depend.

B. Descriptive Research & Consumer Privacy

The idea that empirically-driven determinations can spur better policymaking and regulation is also not a new concept in consumer privacy law, or in regulation more generally. For example, the Commission on Evidence-Based Policymaking was established by a bipartisan law in 2016, and seeks to "develop a strategy for increasing the availability and use of data in order to build evidence about government programs."[256] The Commission released a report of its findings in September of 2017, outlining strategies to improve the way data is collected at all levels of the government, and used to support evidence-based decision-making at all levels of authority.[257] The report is remarkably comprehensive in surveying the history of evidence-based policymaking in the federal government, the development of applicable laws governing government collection and use of data and statistics, and how evidence-based policymaking is currently implemented.[258] The Commission also emphasized that its recommendations concerning data collection must necessarily be accompanied by strong privacy protections and transparency requirements for the data collected.[259] Other recommendations included recommending that Congress and the President direct federal departments to develop long-term learning agendas to support the generation and use of evidence, having the Office of Management and Budget coordinate evidence-building efforts across departments, and securing sufficient research for the same.[260] Adoption of the Commission's recommendations to improve data collection practices and implement broader use of evidence in policymaking would be an

[256] *About CEP,* COMM'N ON EVIDENCE-BASED POLICYMAKING (Jul. 10, 2017), https://perma.cc/G4HF-TGTT.

[257] COMM'N ON EVIDENCE-BASED POLICYMAKING: THE PROMISE OF EVIDENCE-BASED POLICYMAKING 1-3 (2017) (hereinafter *Evidence-Based Policymaking Report*).

[258] *See id.* at 13-15.

[259] *Id.* at 18 ("A theme that runs throughout this report is that access to confidential data for evidence-building purposes should be increased, but only in the context of a modern legal framework providing for strengthened privacy protections and increased transparency.").

[260] *See id.* at 5.

integral part of isolating the lapses in current privacy standards, and discovering new ways to correct them.

Some laws directly depend on a policymaker's assessment of particular privacy norms, for which modeling can provide a more accurate portrait. For example, the Safety Act requires that motor vehicle safety standards be 'practicable,' which includes that there is sufficient public acceptance of the technology proposed by the agency such that wide adoption is possible.[261] As vehicles increasingly incorporate connected features and the development and testing of automatic vehicles continues to accelerate, vehicle privacy is likely a relevant consideration to future safety requirements that the National Highway Traffic Safety Administration adopts. The agency noted in its analysis that quelling public concerns about vehicle privacy and cybersecurity would be necessary for public acceptance.[262] Crowdsourcing and machine learning techniques could provide a more accurate and nuanced basis for determining what the public will in fact accept.[263]

Empirical research on privacy expectations and preferences already shapes FTC policy guidance and enforcement priorities, and both a heavier emphasis on its value, and incorporating more effective forms of research, can better inform those initiatives. The contours of the FTC's analysis under its deception authority, the aspect the agency most broadly relies on in privacy enforcement,[264] makes understanding consumer motivations and norms an important form of inquiry for the agency. When investigating whether a business practice is deceptive, the FTC requires that the practice be likely to mislead the consumer, then analyzes how that practice would be perceived from a consumer acting reasonably under the circumstances, and whether the conduct in question was material.[265] The first two prongs in particular require FTC staff to understand how a consumer thinks when she engages in a

[261] NAT'L HIGHWAY TRAFFIC SAFETY ADMINISTRATION, DOT HS 812 014, VEHICLE-TO-VEHICLE COMMUNICATIONS: READINESS OF V2V TECHNOLOGY FOR APPLICATION 53 (2014) (noting that in Pac. Legal Found. v. Dept. of Transp., 593 F.2d 1338, 1345-46 (D.C. Cir.), cert. denied, 444 U.S. 830 (1979), the D.C. Circuit held that public acceptance is one of the criteria for practicability, one of the factors the agency must demonstrate have met in requiring new safety standards under the Safety Act); Stephen P. Wood, Jesse Chang, Thomas Healy & John Wood, *The Potential Regulatory Challenges of Increasingly Autonomous Motor Vehicles*, 52 SANTA CLARA L. REV. 1423, 1457 (citing same opinion, and concluding "[t]hus, as a part of the agency's considerations, a standard issued by the agency will not be considered practicable if the technologies installed pursuant to the standard are so unpopular that there is no assurance of sufficient public cooperation to meet the safety need that the standard seeks to address").

[262] *See* Nat't Highway Traffic Safety Administration, *supra* note 261, at 133-35.

[263] *See id.* at 133-34

[264] *See* Solove & Hartzog, *supra* note 41, at 599.

[265] FED. TRADE COMM'N, POLICY STATEMENT ON DECEPTION 1 (Oct. 14, 1983).

privacy transaction with a company, and whether or not their decision-making process is reasonable in light of a range of factors, including societal norms.[266] Using surveys to understand consumer norms is nothing new for the agency, which already conducts surveys on consumer attitudes on various topics, using the results to "better understand how consumers perceive statements or representations made to them by businesses," and to help the FTC develop appropriate policy guidance and enforcement strategies.[267]

The agency should sharpen its focus on empirical privacy research and incorporate these new kinds of research that may be more effective at capturing nuanced privacy norms than opinion surveys alone. Modeling privacy behavior and norms may be particularly helpful for informing agency policy guidance, such as which novel practices consumers might be the most likely to consider invasive.[268] Empirical research on consumer preferences, expectations, and behaviors could also be invaluable in agency rulemakings, both in terms of providing substantive information to regulators as they consider how to shape a new policy, and to populate a robust record to avoid challenges that a new privacy rule or policy was created on the basis of insufficient information or demonstration of privacy concern from the public.[269] This research can also play an important role in "soft law" policy guidance, which helps shape industry practice in areas where few concrete laws may appear to apply with clear certainty, such as with emerging technology.[270] In such guidance, the FTC could report the results of the privacy norms of the studies it and others conduct on

[266] *See* Hoofnagle, *supra* note 39, at 123-25; JEF I. RICHARDS, DECEPTIVE ADVERTISING: BEHAVIORAL STUDY OF A LEGAL CONCEPT 14-16 (2010) (describing how the FTC has interpreted how to apply the deception standard to a range of consumer standards, from "unthinking" to "reasonable," since the inception of the Act).

[267] FED. TRADE COMM'N, PRIVACY IMPACT ASSESSMENT FOR FTC CONSUMER SURVEYS 2 (Aug. 2018) (hereinafter *Privacy Impact Assessment*). *See also* FED. TRADE COMM'N, STAFF REPORT, AN EXPLORATION OF CONSUMERS' ADVERTISING RECOGNITION IN THE CONTEXTS OF SEARCH ENGINES AND NATIVE ADVERTISING 1 (Dec. 2017) (describing the FTC's survey on consumer attitudes regarding native advertising); Hoofnagle, *supra* note 39, at 124 (describing how the agency determines whether a claim was misleading in implied representation cases, and noting that the Commission can, but is not required to, use surveys).

[268] *See* Schwartz & Peifer, *supra* note 9, at 149-150.

[269] *See, e.g.*, ANA Comments, *supra* note 8, at 9 (arguing that the basis the FCC offered for the broadband privacy rule, that it is justified to materially advance an interest in protecting the privacy of customer information as the vast majority of adults deem it important to control who can get information about them, and the number of data-collecting entities continues to grow and puts them at risk of information misuse, was not sufficiently proven in the record).

[270] *See* Solove & Hartzog, *supra* note 41, at 625-26 (noting that while FTC policy guidance materials "may not be exactly akin to advisory opinions, but they can come quite close. Companies take the guidance in these materials seriously," and comparing their effect on shaping conduct to dicta in judicial opinions).

privacy norms and behavior, and address how the insights gleaned from them can be incorporated into industry practice.[271] In addition to the cues that FTC guidance provides to industry, it can also shape the enforcement authorities of state and local consumer agencies, further facilitating the potential impact of any steps it takes to promote this kind of research and the use of its findings.[272]

The FTC is also well-equipped to conduct empirical privacy research. As the *de facto* US data protection agency, the Commission has both the authority and the imperative to share its technical expertise and insight on specific policy areas with relevant stakeholders, including industry, consumer advocates, consumers themselves, and other agencies. Its role in shaping privacy practices will only continue to grow as a range of industries increasingly involve technology, particularly as the agencies that have never needed institutional expertise in privacy find themselves in need of exactly that.[273] Empirical research on privacy norms is germane to the FTC's institutional role, and can have a decisive impact in shaping privacy policy as the FTC guides the policymaking of other enforcement bodies.

The relevant questions of the role empirical research may play in a policymaker's decision-making is perhaps the easiest to answer for a legislative body. In terms of the traditional sense of institutional parameters, as the branch with the most popular accountability, the legislature is least constrained in its prerogative to rely on certain kinds of evidence over others. Congress may overrule court decisions with which it disagrees, and it has done so, such as with the passage of the

[271] *See Privacy Impact Assessment*, *supra* note 267, at 2 (describing how the Commission uses surveys to shape enforcement and mold policy guidance for industry).

[272] Danielle Citron, *The Privacy Policymaking of State Attorneys General*, 92 NOTRE DAME L. REV. 747, 753 (2016) (discussing the impact of the FTC on privacy enforcement by state attorney generals, and quoting one AG as describing the FTC as "the 'mother ship' on data security issues because it has unique technical know-how that would be hard to reproduce at the state level").

[273] For example, the FTC has recently collaborated with the Department of Education, on student privacy. *See Student Privacy and Ed Tech*, FTC (Dec. 1, 2017), https://perma.cc/3D3Q-YCZ2. The FTC also joined forces with the National Highway Traffic & Safety Administration on connected cars. *See FTC, NHTSA Workshop to Focus on Privacy, Security Issues Related to Connected Cars*, FTC (June 27, 2017), https://perma.cc/N449-VBBD. The FTC also provided guidance on mobile health apps in conjunction with the Food and Drug Administration and the Department of Health and Human Services. *See* Press Release, Fed. Trade Comm'n, *FTC Releases New Guidance for Developers of Health Apps*, FTC (Apr. 5, 2016), https://perma.cc/KQK7-2AL3. The Commission has been repeatedly invoked as privacy guarantor in ongoing debates about broadband privacy. *See* Press Release, Fed. Comm. Comm'n, *Statement of FCC Chairman Ajit Pai on Congressional Resolution of Disapproval of FCC Broadband Privacy Regulations* (Mar. 28, 2017), https://perma.cc/A36V-UHFU. The agency's expertise will likely only continue to be needed as technology becomes relevant to the sectoral domains of other agencies.

1934 Communications Act in response to the Supreme Court's decision in *Olmstead*,[274] or the passage of the Wiretap Act[275] in response to *Katz*.[276] That same accountability makes them subject to public pressure, particularly when public pressure comes in the form of special interest groups with a heavy incentive to sway a legislator's perception of what the public wants.[277] Congress can and should incentivize further research and remove existing barriers, but the likelihood that research on privacy norms will impact the position of any legislator not already predisposed to believe its findings is slim at best. That said, members of both the House and Senate have shown interest in the both the benefits and risks of artificial intelligence,[278] and privacy writ large is even more of an established issue on the Hill.[279] Reviving Congress's Office of Technological Assessment could be a helpful step towards enabling the Hill to tackle nuanced issues of privacy and technology,[280] though it would not in and of itself affect the underlying incentives and pressures from lobbying that have prevented the majoritarian branch from legislating stronger privacy protections,[281] giving more resources to the FTC, or any of the other measures that Congress could accomplish to spur this kind of research.

[274] Communications Act of 1934 § 605, Pub. L. No. 73-416, 48 Stat. 1064, 1103 (1934) (codified at 47 U.S.C. § 605).

[275] Omnibus Crime Control and Safe Streets Act of 1968, sec. 802, § 2511, Pub. L. No. 90-351, 82 Stat. 197, 213 (1968), *amended by* Electronic Communications Privacy Act of 1986, Pub. L. No. 99-508, 100 Stat. 1848 (1986) (current version at 18 U.S.C. § 2511 (2012)).

[276] *See* Richard C. Turkington, *Protection for Invasions of Conversational and Communication Privacy by Electronic Surveillance in Family, Marriage, and Domestic Disputes Under Federal and State Wiretap and Store Communications Acts and the Common Law Privacy Intrusion Tort*, 82 NEB. L. REV. 693, 700-2 (2004) (discussing Congressional responses to *Olmstead* and *Katz*).

[277] *See, e.g.*, Alvaro Bedoya, *Why Silicon Valley Lobbyists Love Big, Broad Privacy Bills*, N.Y.T. (April 11, 2018), https://www.nytimes.com/2018/04/11/opinion/silicon-valley-lobbyists-privacy.html.

[278] In May of 2017, Representatives John Delaney and Pete Olsen launched a bipartisan Artificial Intelligence Caucus. *See* Press Release, Rep. John K. Delaney, *Delaney Launches Bipartisan Artificial Intelligence (AI) Caucus for 115th Congress* (May 24, 2017), https://perma.cc/4Y9C-3X8X. The caucus produced the FUTURE of Artificial Intelligence Act of 2017, which addressed the risks to privacy posed by AI, as well as concerns of bias. *See* FUTURE of Artificial Intelligence Act of 2017, H.R. 4625, 115th Cong. (1st Sess. 2017); FUTURE of Artificial Intelligence Act of 2017, S. 2217, 115th Cong. (1st Sess. 2017).

[279] In the Senate, there's the Judiciary Subcommittee on Privacy, Technology, and the Law. Other committees that frequently address privacy issues, the Senate Select Committee on Intelligence Responsibilities and Activities and the Senate Committee on Commerce, Science, and Transportation. The House contains similar coalitions, like the Digital Commerce and Consumer Protection Subcommittee and the House Intelligence Committee.

[280] The Office of Technology Assessment provided Congress with expert advice on complex technological issues, until it was unfunded in the 1990's. *See* ZACH GRAVES & KEVIN KOSAR, R STREET POLICY STUDY NO. 128, BRING IN THE NERDS: REVIVING THE OFFICE OF TECHNOLOGY ASSESSMENT 1-2 (2018), https://perma.cc/6Z32-NQYB .

[281] *See, e.g.*, Bedoya, *supra* note 277.

The technical agencies are ultimately the best positioned to conduct and incentivize empirical privacy research, and the most likely to be in a position where their efforts will have any observable effects.

C. Consumer Privacy & Operative Tools

The ideal objective of descriptive research is to provide more accurate information on privacy expectations, preferences, and norms to guide the enforcement of current laws, and the development of future ones. In contrast, operative tools aim to transform user behavior under the existing rules shaping user behavior, rather than necessarily transforming the application of current rules. However, tools that guide how users interact with privacy-invasive technology may have the most immediate effect on privacy behavior. Tools like browser extensions that flag aberrant privacy policies, smart privacy assistants, or permissions managers can accomplish the concrete work of accounting for notice and choice's deficiencies, and ideally restore some degree of meaningful control and choice to the user.

The federal government should support and incentivize this kind of user-centric privacy research. In some areas, it may be appropriate for policy to be responsive to the development of user-friendly tools, such as requiring some degree of uniformity in privacy policies , such that they can be machine-readable.[282] Agencies can lend visibility to researchers through workshops and conferences, providing grants to allow these studies to be expanded upon and replicated, or even simply encouraging the pursuit of such research through outreach.[283] During the Obama administration, the White House Office of Science and Technology Policy established the National Privacy Research Strategy, which created objectives for federally-funded privacy research, and provided guidance to agencies on how to further promote privacy research.[284] Even further, the adoption of the proposals made in the Evidence-Based Policymaking Commission report would spur research in the private sector, such as adopting a centralized process through which researchers could apply to receive access to non-public

[282] *See* McDonald & Cranor, *supra* note 41, at 547-48.

[283] *See, e.g.*, Lorrie Cranor, *Your Research Can Help the FTC Protect Consumers*, FTC (Jan. 17, 2017), https://perma.cc/JPC3-V6US (former FTC Chief Technologist Lorrie Cranor describing her efforts to work with private researchers, and suggesting 9 areas for privacy researchers to pursue); Terrell McSweeny, Commissioner, FTC, & Lorrie Cranor, Chief Technologist, FTC, Address at DEF CON: Research on the Machines (Aug. 5, 2016) (slides available at https://perma.cc/TB9Y-WFVX) (FTC outreach presentation at the security conference DEF CON).

[284] NAT'L SCI. & TECH. COUNCIL, NATIONAL PRIVACY RESEARCH STRATEGY 4 (2016).

government data.[285] The spread of proactive tools like permissions managers and browser extensions can only help improve the ultimate goal of these efforts: to ensure that users are actually informed about their privacy choices and that their privacy preferences are actually respected by the entities collecting and profiting from their information.

V. FURTHER CONSIDERATIONS

The narrow applications of AI discussed in this paper are particularly well-suited to the task of improving our understanding of privacy expectations and norms. Classifiers that can deduce and incorporate granular contextual data can help illustrate the factors behind why an individual makes a given privacy decision, and models that include feedback and auditing mechanisms can adapt as preferences evolve.[286] Natural language processing can be used to automate analysis of privacy policies,[287] demonstrating trends in widely used provisions and terms that may be hampering the ability of users to make the decisions that suit their preferences. Non-AI empirical approaches, such as crowdsourcing, can be used to amplify these techniques and to deduce consensus on broader informational norms.[288]

Not all of the research suggested here requires the use of AI. The educational privacy contextual modeling study conducted by the NYU and Princeton researches illustrated that its results could be *amplified* by using machine learning to predict additional norms, but the study simply relied on crowdsourcing to produce the norms themselves.[289] Other methods could rely on statistical regression, or other analytical methods that prioritize comprehension. The focus of this article is on empirical techniques that can either enhance the value of existing methods of privacy research (fundamentally, crowdsourcing can be understood as a different method of conducting an opinion survey), or support new approaches that will shed light on the discordance between privacy preferences, expectations, and behavior. Not all of these will rely on AI, nor will the particular value of a given study necessarily stem from the use of AI.

[285] *See Evidence-Based Policymaking Report*, *supra* note 257, at 46, 101.

[286] *See, e.g.*, Tsai et al., *supra* note 73, at 147 (describing Turtleguard, and noting that "[r]ecent research on permission models has turned towards using machine learning . . . . One advantage is ML's ability to incorporate nuanced contextual data to predict user preferences") (citations omitted).

[287]*See* Sadeh et al., *supra* note 181, at 3.

[288] *See* Shvartzshnaider et al., *supra* note 158, at 2.

[289] *See* Shvartzshnaider et al., *supra* note 158, at 3.

It will also be enormously important that the limits of this research are carefully considered and understood. In extrapolating any conclusions from quantitative findings, it is crucial to consider what data was collected and interpreted, but also what data was not. Data is not found, but made.[290] A model that analyzes user behavior within a certain context may be idiosyncratic to that context. Not all privacy interests can be quantified or quantified accurately. Beyond the constraints of what evidence policymakers may legally rely on, and what various threads of political philosophy dictate that they should rely on, there are also the factual constraints shaping how AI modeling can be used. Modeled norms will need to be reliable, replicable, and representative of the populations they are intended to model.

In the case of individual modeling, researchers should endeavor to consider how built-in limitations of their experiments might prevent their results from being fully representative. For instance, one of the studies measured user behavior on apps that would not lose functionality if permission settings were denied, so that the researchers could capture user ideal behavior – preferences, as opposed to either expectations or resulting tradeoffs.[291] Further research and modeling would have to be careful to distinguish measuring user expectations from users' likely response when confronted with a conflict, such as loss of functionality, and apply the results of their research accordingly. In each of the studies described, the researchers noted the limits of population size and skew, and how possible it may be to generalize their results. Any research conducted should be scrupulously careful to avoid any bias in sampled groups. Without a careful approach to ensuring that the data used does in fact reflect a range of demographic groups, the objective that such a project will seek to accomplish – providing consumer privacy regulators or judges with a more accurate understanding of primary norms – will fail and create new systematic bias in its wake. In the case of crowdsourcing methods, researchers must be careful to avoid issues of self-selection bias relating to the population of crowdsourced workers used.[292] In particular, Amazon Mechanical Turk, a crowdsourcing platform used by a wide variety of researchers including some discussed here, has been subject to a range of critiques about the population of crowdsourced workers it

[290] *See generally* Nick Diakopoulos, *The Rhetoric of Data*, MUSINGS ON MEDIA (July 25, 2013), https://perma.cc/R6DU-LJN5 ("It's important to recognize and remember that data does not equal truth.").

[291] *See* Wijesekera et al., *supra* note 191, at 1091.

[292] *See* Emami-Naeini et al., *supra* note 46, 408-9.

provides.[293] Most studies of mobile systems are conducted on non-IOS systems, which may limit the applications of those findings.[294]

Different areas of privacy will also present different practical obstacles for experimental design, such as working with children. Researchers have applied the contextual integrity framework to examining how children understand online privacy, but given the legal and ethical concerns of working with children, the survey responses were gathered through in-person interviews as opposed to gathering responses online.[295] The FTC noted in a 2018 Privacy Impact Assessment for FTC Consumer Surveys that the agency does not normally conduct surveys of children under the age of 13, though if it did, COPPA would govern the agency's collection and use of the information.[296] Other contexts may pose similar limitations (or prohibitions) for using crowdsourcing and other methods. The legal regime constraining how certain information can be used, such as health information, will also often restrain research. Any kind of research that relies on collecting and processing data must also anticipate how to protect it. Operative tools that collect contextual data to improve privacy decision-making invite the privacy risks of the very data it collects, which private researchers must consider in how their tools are constructed.[297] Descriptive research conducted by the government must also be held to strong standards of privacy protection of the information gathered and created.[298] Fundamentally, proponents of empirical policy solutions must be carefully attentive that new techniques do not create new problems as they solve old ones: research

[293] *See, e.g.*, Joseph K. Goodman, Cynthia E. Cryder & Amar Cheema, *Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples*, 26 J. BEHAV. DECISION MAKING 213, 1-2 (2012) (noting that Mechanical Turk generally provides valuable and reliable crowd-workers for researchers, but finding "notable differences for [Mechanical Turk] participants that researchers should consider before using [Mechanical Turk] for their own research").

[294] *See, e.g.*, Christopher Buck, Simone Burster & Torsten Eyemann, *Priming App Information Privacy Concerns in Mobile Ecosystems* 14 (Universität Bayreuth Working Paper No. 63, 2017).

[295] *See* Priya Kumar, Shalmali Milind Naik, Utkarsha Ramesh Devkar, Marshini Chetty, Tamara L. Clegg & Jessica Vitak, *"No Telling Passcodes Out Because They're Private": Understanding Children's Mental Models of Privacy and Security Online*, 1 PROC. ASS'N OF COMPUTING MACHINERY ON HUMAN-COMPUTER INTERACTION 64:1, 64:11 (2017).

[296] *See Privacy Impact Assessment*, *supra* note 267, at 10.

[297] *See* Schaub et al., *supra* note 86, at 40 (noting that context-adaptive mechanisms that collect contextual data must have a strategy for protecting it).

[298] *See Evidence-Based Policy Report*, *supra* note 257, at 47 (recommendations for enhanced privacy protections to accompany increased use of evidence in policymaking, including recommendations for technical protocols to adopt, amendments to the Privacy Act and the Confidential Information Protection and Statistical Efficiency Act, and designating senior officials responsible for those efforts).

methodology that produces transparent, reliable results that are capable of replication is absolutely crucial.

CONCLUSION

A quickly evolving world demands reflective attention to how legal standards have withstood the pressure of novel threats. In consumer privacy law, the notice and choice regime rests on the idea that consumers are empowered to make rational privacy decisions that will exert control over their information, and that failure to do so is a rejection of privacy, not a failure of the privacy decision-making process or of the legal mechanisms designed to protect it. In Fourth Amendment law, blunt standards like the third-party doctrine and the content/non-content distinction demonstrate how new contextual factors created by new technology have changed the principles undergirding those rules, while judges fail to glean what the average person's expectation of privacy actually is. New capacities for surveillance and artificial intelligence seem to only further erode a dying structure, making the prognosis all the more grim.

But just as technology may threaten privacy, it may also provide the capacity to illustrate weaknesses in our current thinking, and prevent future mistakes. Proactive privacy tools can nudge consumers into decision-making that better reflects their preferences, and which will provide a better portrait of what privacy norms really are. Contextual modeling can provide policymakers with a clearer portrait of the questions they have struggled to answer since the invention of the personal computer: why, exactly, do people care about privacy, and how should that answer shape the privacy protections offered by the law? Empirical research may not provide that answer definitively anytime soon, if ever – but it may be able to guide policymakers to reach estimations that are more accurate than their own conjecture, and inform the development of more effective and accurate policies around privacy norms and decision-making.