

6-2025

## Can AI, as Such, Invade Your Privacy? An Experimental Study of the Social Element of Surveillance

Aileen Nielsen

Harvard Law School, [ainielsen@law.harvard.edu](mailto:ainielsen@law.harvard.edu)

Follow this and additional works at: <https://www.repository.law.indiana.edu/ilj>



Part of the [Privacy Law Commons](#), and the [Science and Technology Law Commons](#)

### Recommended Citation

Nielsen, Aileen (2025) "Can AI, as Such, Invade Your Privacy? An Experimental Study of the Social Element of Surveillance," *Indiana Law Journal*: Vol. 100: Iss. 4, Article 9.

Available at: <https://www.repository.law.indiana.edu/ilj/vol100/iss4/9>

This Article is brought to you for free and open access by the Maurer Law Journals at Digital Repository @ Maurer Law. It has been accepted for inclusion in Indiana Law Journal by an authorized editor of Digital Repository @ Maurer Law. For more information, please contact [kdcogswe@indiana.edu](mailto:kdcogswe@indiana.edu).



**JEROME HALL LAW LIBRARY**

INDIANA UNIVERSITY  
Maurer School of Law  
Bloomington

# Can AI, as Such, Invade Your Privacy?

## An Experimental Study of the Social Element of Surveillance

AILEEN NIELSEN\*

*The increasing use of AI rather than human surveillance puts pressure on two long-used cultural and (sometimes) legal distinctions: as between human and machine observers and as between content and metadata. Machines do more and more watching through advancing technology, rendering AI a plausible replacement for humans in surveillance tasks. Further, machines can commit to surveil only certain forms of information in a way that humans cannot, rendering the distinction between content and metadata increasingly relevant too for crafting privacy law and policy. Yet despite the increasing importance of these distinctions, their legal importance remains in four key domains of privacy law: Fourth Amendment law, wiretap law, consumer privacy law, and the privacy torts. Given the failure of privacy law to settle conclusively the import of the human/AI and content/metadata distinctions, this Article proposes looking to empirical measures of the judgments of ordinary people to better understand whether and how such distinctions should be made if law is to be responsive to reasonable expectations of privacy.*

*There is incomplete empirical evidence as to whether the AI/human surveillance and content/metadata distinctions hold weight for ordinary people, and if so, how. To address this empirical gap, this Article presents the results of a vignette study carried out on a large ( $N = 1000$ ), demographically representative sample of Americans to elicit their judgments of a state surveillance program that collected either content or metadata and in which potential surveillants could be either human or AI. Unsurprisingly, AI surveillance was judged to be more privacy preserving than human surveillance, empirically buttressing the importance of a human/AI distinction. However, the perceived privacy advantage for an AI surveillant was not a dispositive factor in stated preferences regarding technology use. Accuracy—a factor rarely discussed in defenses of state surveillance—was more influential than privacy in determining participants' preferences for a human or AI surveillant. Further, the scope of information surveilled (content or metadata) strongly influenced accuracy judgments in comparing human and AI systems and shifted surveillance policy preferences as between human and AI surveillants. The empirical data therefore show that the distinction between content and metadata is important to ordinary people, and that this distinction can lead to unexpected outcomes, such as a preference for human rather than AI surveillance when contents of communications are collected.*

---

\* Visiting Assistant Professor, Harvard Law School, [ainielsen@law.harvard.edu](mailto:ainielsen@law.harvard.edu). For helpful comments and discussions, the author thanks Christoph Engel, Nina Grgić-Hlača, Matthew Kugler, and Alexander Stremitzer as well as participants at the following workshops: Experimental Methods in Legal Scholarship (EMLS), Privacy Research Group (PRG) at NYU, Works-in-Progress Roundtable on Law and Computer Science at Penn, GMU Law and Economics Workshop, the Georgetown Law and Economics Workshop, and the Boston University Faculty Workshop. For outstanding research assistance, the author thanks Emily Hua, Christina Lee, Pedro Ribeiro Morais e Silva, Chloe Suzman, and Sydney Veatch.

*These empirical results provide strong evidence for judges and legislators alike to contemplate how various domains of privacy law could be more responsive (and more consistently responsive) to the expectations of ordinary people. Ordinary people have reasonable expectations not only of privacy but also of accuracy, and the tension between these preferences can lead to surprising outcomes. Long held assumptions in state surveillance policy that automated surveillance is not privacy invasive or is otherwise minimally privacy invasive are unresponsive to the nuanced judgments of ordinary people. While law and policy need not always reflect the intuitions or preferences of the citizenry, empirical data such as offered in this Article could prove helpful where judges and lawmakers have otherwise failed to establish a clear and convincing consensus.*

INTRODUCTION .....	1675
I. COMPUTATION AND SURVEILLANCE .....	1681
A. SURVEILLANCE BY HUMANS VERSUS BY MACHINES .....	1683
B. SURVEILLANCE OF CONTENT VERSUS METADATA .....	1686
C. SUMMING UP .....	1689
II. HUMAN VERSUS AI PRIVACY .....	1690
A. THE FOURTH AMENDMENT .....	1692
B. COMMUNICATIONS PRIVACY STATUTES.....	1699
C. THE PRIVACY TORTS.....	1701
1. INTRUSION UPON SECLUSION.....	1702
2. PUBLIC DISCLOSURE OF PRIVATE FACTS.....	1703
D. CONSUMER PROTECTION.....	1706
E. SUMMING UP THE LEGAL DISTINCTIONS AND NONDISTINCTIONS .....	1708
F. EMPIRICAL SCHOLARSHIP .....	1709
III. CONTENT VERSUS METADATA.....	1712
A. FOURTH AMENDMENT.....	1715
B. COMMUNICATIONS PRIVACY .....	1716
C. THE PRIVACY TORTS.....	1719
D. CONSUMER PROTECTION.....	1721
E. SUMMING UP THE LEGAL DISTINCTIONS AND NONDISTINCTIONS .....	1723
F. EMPIRICAL LITERATURE .....	1723
IV. AN EXPERIMENTAL APPROACH.....	1724
A. DESIGN AND PROCEDURE.....	1726
B. METHODS .....	1729
C. RESULTS.....	1729
1. SAMPLE .....	1729
2. PRIVACY ASSESSMENTS OF AI AND HUMAN OBSERVERS .....	1730
3. IMPORTANCE OF PRIVACY IN SURVEILLANCE CHOICES.....	1731
A. SUPPORT FOR THE SURVEILLANCE PROGRAM .....	1731
B. SURVEILLANCE CHOICES .....	1731
C. SHIFTING CHOICES FOR THE STATE VERSUS THE SELF .....	1735
D. DISCUSSION OF RESULTS.....	1736
E. LIMITATIONS OF RESULTS .....	1738
V. POLICY IMPLICATIONS .....	1739
CONCLUSION .....	1741

## INTRODUCTION

With the proliferation of computational surveillance in both governmental and commercial use cases, the status of computers as surveillants has become both increasingly common and also increasingly important to key legal doctrines. Yet, the legal status of human versus AI observation remains unclear in fundamental questions oft posed in public and private law alike. Can observation by a computer impinge on reasonable expectations of privacy in the same way as does observation conducted by a human?<sup>1</sup> It depends. Does acquiescence to surveillance by computer

---

1. Privacy, and reasonable expectations thereof, is a notoriously difficult concept to

implicate a surrender of reasonable expectations of privacy as against a human? It depends. Is the use of human versus AI observation a material element of a process or product that should be disclosed for valid consent? It depends. The first contribution of this paper lies in surveying four key privacy law domains: Fourth Amendment law, federal communications privacy law, the privacy torts, and consumer protection law (henceforth, “the four privacy domains”), each implicated by a human/machine observer distinction.<sup>2</sup> A targeted survey of the four privacy domains shows that courts and legislators alike have been consistently inconsistent in assessing the privacy import of machine, rather than human, observation.

The increasing importance of the human/machine distinction adds new heft to a pre-digital distinction in privacy law: that between content and metadata, a distinction close to those made between content and envelope data under Fourth Amendment law, or between contents and records under federal communications privacy law.<sup>3</sup> Unlike humans—who may promise to honor observational restrictions but cannot be technically guaranteed to do so—machines, including AI, can credibly commit in a technically verifiable manner to collect, use, or retain only certain classes or quantities of information. The longstanding content/metadata distinction is functionally more important when we can grant more credence to selective or granular restrictions on the nature or extent of surveillance, and therefore the importance of the distinction is strongly tied to the increasing capabilities of machine surveillance.<sup>4</sup> And yet, as with the case of the human/machine distinction, a targeted review of relevant law shows that the four privacy domains are also inconsistent in their treatment of the content/metadata distinction.

Inconsistent legal treatment of a non-legal distinction is uninteresting. For example, in a given sample of cases, one could potentially find inconsistent treatment of plaintiffs who wore yellow rather than red at the moment of an alleged privacy incursion. This would be merely a statistical accident. Or one might find a distinction that matters to ordinary people but that has been rejected by law.<sup>5</sup> Or a distinction

---

define. Here, I adopt a broad understanding of privacy that goes well beyond the handling of personal data to also include a subjective sense of being observed and also an understanding that observation can itself impinge on privacy as decisional autonomy.

2. Throughout the work I use AI, computer, and machine interchangeably. One could imagine distinguishing non-digital machines from digital machines and likewise defining AI in a way that could lead to a distinction as between AI and other digital but non-AI applications. Given the inconsistencies documented *infra*, it seems unlikely to be fruitful to search the caselaw with such finer-grained distinctions.

3. Of course, one may identify ways in which these three distinctions are not fully co-extensive. My claim here is that the better-known distinctions drawn from privacy law roughly map on to the content/metadata distinction.

4. The well-informed reader may raise an eyebrow in recognition of the flawed conceptual distinction between content and metadata, which will be discussed *infra*. Nonetheless, the longstanding content/metadata distinction retains a hold on the judicial and legislative imagination, as will be shown *infra*, even if the conceptual clarity of this distinction is not robust and has arguably become increasingly flawed in an era of big data in which distinct sources of metadata can be joined together or in which the extensive recordings of metadata provide a highly informative depiction of a surveillance subject.

5. For example, ordinary people exhibit a per se promising effect, in which they will keep a promise simply because they have promised, independent of economic incentives and

that matters to science but not to law.<sup>6</sup> One might likewise find law making distinctions that do not matter to ordinary people,<sup>7</sup> or for that matter to scientists.<sup>8</sup> Hence, a skeptic could fairly ask whether there is any significance at all to the fact that I identify two distinctions that are inconsistently treated across privacy law. Law is its own system of rules and logic, and that system need not track the science of information or the folk morality of privacy. And yet, at the least, the latter *ought* to be important to privacy law for reasons related to the formalities of privacy law itself. Key areas of privacy law formally and informally operationalize a *legal standard* of reasonable expectations of privacy. Given this *legal formality*, it appears both puzzling and potentially problematic that areas of privacy law claiming to track reasonable expectations of privacy do not consistently treat factors that do indeed matter to ordinary people (as is well known and is further demonstrated in the experimental results presented in this Article).

The importance of reasonable expectations of privacy is firmly established in all four of the privacy domains studied here. In Fourth Amendment law, the legal standard for establishing what is a search relies on expectations of privacy, following *Katz* and particularly following the two-part reasonable expectations of privacy test established in Justice Harlan's concurrence.<sup>9</sup> In communications privacy law,

contrary to the norms promoted by contract law's theory of efficient breach. See Dorothee Mischkowski, Rebecca Stone, & Alexander Stremitzer, *Promises, Expectations, and Social Cooperation*, 62 J.L. & ECON. 687 (2019).

6. One example of such a distinction is the fact that the *Daubert* standard for admission of expert testimony is not rigorous to the same degree as peer review of scientific research. See Louise Marie Jupe & Vincent Denault, *Science or Pseudoscience? A Distinction that Matters for Police Officers, Lawyers, and Judges*, 26 PSYCHIATRY PSYCH. & L. 753 (2019). Another, unrelated example is that the law's notion of physical or tangible for purposes of trespass is unrelated to insights from physics as to the identification or movement of physical particles. For example, the intentional projection of light onto another's property is not recognized as a trespass or as a nuisance, in part because light is held to be "intangible" despite the well-established physics describing photons as physical particles and the impressive empirical physics of counting photons. See Maureen E. Brady, *Property and Projection*, 133 HARV. L. REV. 1144 (2020); see also Liang-Cheng Tu, Jun Luo & George T. Gillies, *The Mass of the Photon*, 68 REPS. PROGRESS PHYSICS 77 (2005) (describing methods and bounding of attempts to measure the rest mass of a photon).

7. A concept that matters to law more than it appears to matter to ordinary people is proper process independent of outcomes. To ordinary people, the way in which an outcome may be influenced by a "technicality" can sometimes undermine the perceived legitimacy of the law; even as to the legal community, it is legitimacy itself that is at issue by assuring appropriate and due process. See *Getting Off on a Technicality: Can It Happen?*, MAY LAW LLP, <https://maylawllp.com/technicalities-in-criminal-cases/> [https://perma.cc/M45E-GP2H].

8. The bifurcation in tort law of the causation element of negligence into actual causation and proximate causation marks a distinction not made in science, where tenuous chains of causation are not only of interest but are the subject of their own study, as in the study of "chaos theory" or "complex systems." See *Complex System*, WIKIPEDIA, [https://en.wikipedia.org/wiki/Complex\\_system](https://en.wikipedia.org/wiki/Complex_system) [https://perma.cc/7KRP-SBDP]; *Chaos Theory*, WIKIPEDIA, [https://en.wikipedia.org/wiki/Chaos\\_theory](https://en.wikipedia.org/wiki/Chaos_theory) [https://perma.cc/9QQH-3FYC].

9. *Katz v. United States*, 389 U.S. 347, 361 (1967) (Harlan, J., concurring).

consider as one example of the importance of expectations of privacy, the definition in the Wiretap Act of an “oral communication,” which defines such protected communications in part contingent on circumstances that would justify an expectation of privacy.<sup>10</sup> The privacy torts, too, routinely lean into an expectation of privacy even though such an expectation constitutes neither a formal element of any privacy tort nor the language used by the Restatement to describe any of the four invasion of privacy torts.<sup>11</sup> In the case of consumer protection and privacy as carried out by the FTC, the relevance of reasonable expectations of privacy is indisputable both with respect to the Agency’s enforcement of cybersecurity practices and with respect to the Agency’s standards for and enforcement actions regarding consumer privacy.<sup>12</sup>

In view of the important but inconsistent treatment of the human/machine and content/metadata distinctions in the four privacy domains, and in view of the importance of reasonable expectations of privacy, this work emphasizes the importance of empirical data about what surveillance implementation factors are important to ordinary people in their judgments of a proposed state surveillance scenario. Such empirical data can offer information rather than guesswork as to the privacy significance and directionality of these distinctions. Most helpfully, the

---

10. 18 U.S.C. § 2510(2) (“[O]ral communication means any oral communication uttered by a person exhibiting an expectation that such communication is not subject to interception under circumstances justifying such expectation . . . .”) (emphasis added).

11. See e.g., *Safari Club Int’l v. Rudolph*, 862 F.3d 1113, 1128 (9th Cir. 2017) (referencing “expectations of privacy” both in assessing state wiretapping and intrusion upon seclusion claims and defenses). See generally Chloe Suzman, *Mapping Fourth Amendment Influence in Privacy Law* (Working Paper, 2024) (on file with author).

12. See e.g., Isabelle Wright & Maia Hamin, “Reasonable” Cybersecurity in Forty-Seven Cases: The Federal Trade Commission’s Enforcement Actions Against Unfair and Deceptive Cyber Practices, ATL. COUNCIL (June 12, 2024), <https://www.atlanticcouncil.org/in-depth-research-reports/report/reasonable-cybersecurity-in-forty-seven-cases-the-federal-trade-commissions-enforcement-actions-against-unfair-and-deceptive-cyber-practices/> [<https://perma.cc/9ZYX-9D8L>] (describing the arguments made by FTC in forty-seven different enforcement actions between 2002 and 2024, each describing how a firm “failed to meet the bar for ‘reasonable’ cybersecurity”). Reasonable cybersecurity is just one subset of reasonable expectations of privacy. The FTC has also invoked “reasonable expectations of privacy” explicitly, for decades. For a description of the policy goals related to the FTC goals, see, for example, Christine A. Varney, *Consumer Privacy in the Information Age: A View From the United States*, FED. TRADE COMM’N (Oct. 9, 1996) <https://www.ftc.gov/news-events/news/speeches/consumer-privacy-information-age-view-united-states> [<https://perma.cc/AX3H-4VAB>] (“First, an individual’s reasonable expectation of privacy regarding access to and use of his or her personal information should be assured.”). For a description of how reasonable expectations of privacy inform enforcement actions under the deceptive practices prong of FTC authority, see Letter from James C. Miller III, Chairman, FTC, to Hon. John D. Dingell, Chairman, House Comm. on Energy & Com. (Oct. 14, 1983) [hereinafter Letter from James C. Miller III to Hon. John D. Dingell] (on file with the *Columbia Law Review*), reprinted in *In re Cliffdale Assocs., Inc.*, 103 F.T.C. 110, 1984 WL 565319 at \*45(1984) (decision & order); *Id.* at \*37 n.4 (“A misleading omission occurs when qualifying information necessary to prevent a practice, claim, representation, or reasonable expectation or belief from being misleading is not disclosed. Not all omissions are deceptive, even if providing the information would benefit consumers.”).

randomized nature of the experimental treatment in the vignette experiment presented here permits us to draw causal conclusions about whether and how the collection of content versus metadata influences preferences and judgments about human versus machine surveillance.

In this work, I present the results of a large ( $N = 1000$  participants), nationally representative vignette study of Americans.<sup>13</sup> The scope of the observation (content versus metadata only) was varied in a randomized controlled experimental treatment, while participants indicated judgments and preferences about human versus AI surveillance. In the vignette study, research participants encountered a state surveillance scenario in which a voluntary monitoring program was offered to families welcoming home previously imprisoned violent criminal offenders. Participants were told that an audio monitoring device could be installed in the home at the option of the family members, and that this device would transmit either full audio information (the Content treatment) or only garbled audio in which tone and volume, but not words, would be discernible (the Metadata treatment).<sup>14</sup> Participants were asked to judge the relative privacy and accuracy of human and AI monitoring systems, where each service was offered by a separate commercial entity, thus mirroring the common scenario in which routine police surveillance capacities are increasingly outsourced to private sector technical capacities.<sup>15</sup> Participants were also asked to indicate which service, if either, they would choose as well as which service they expected the State to choose.

Consistent with widely shared intuitions, AI surveillance was more frequently judged to be more privacy preserving than human surveillance, rather than the other way round. On the other hand, participants more frequently perceived an accuracy advantage for humans than for AI in the case where full audio contents were transmitted (Content case). Putting these findings together, AI surveillance was judged to be more private than human surveillance (in the case of either Content or

---

13. Vignette studies are an increasingly used empirical methodology in legal studies. *See, e.g.*, Tess Wilkinson-Ryan, *Justifying Bad Deals*, 169 U. PA. L. REV. 193 (2020); Meirav Furth-Matzkin & Roseanna Sommers, *Consumer Psychology and the Problem of Fine-Print Fraud*, 72 STAN. L. REV. 503 (2020).

14. Although I am aware of no such Metadata treatment-like product currently on the market, the notion of detecting emotional valence from tone only and not from content is plausible. *See e.g.*, Einat Liebenthal, David A. Silbersweig & Emily Stern, *The Language, Tone, and Prosody of Emotions: Neural Substrates and Dynamics of Spoken-Word Emotion Perception*, 10 FRONTIERS IN NEUROSCIENCE 1, 3 (2016) (“In the auditory system, the voice parallels the face in that it conveys a person’s identity and current emotional status. Some aspects of voice emotions (in particular emotional category, e.g., anger, disgust, fear, sadness, joy) are thought to be perceived quickly based on coarse tone and intensity analysis of brief segments of familiar non-verbal vocalizations (e.g., shriek, cry, laugh etc...) . . .”).

15. Nicol Turner Lee & Caitlin Chin-Rothman, *Police Surveillance and Facial Recognition: Why Data Privacy Is Imperative for Communities of Color*, BROOKINGS (Apr. 12 2022), <https://www.brookings.edu/articles/police-surveillance-and-facial-recognition-why-data-privacy-is-an-imperative-for-communities-of-color/> [https://perma.cc/R9UT-QJUU] (“Federal, state, and local law enforcement agencies often rely upon tools developed within the private sector, and, in certain cases, can access massive amounts of data either stored on private cloud servers or hardware (e.g., smartphones or hard drives) or available in public places like social media or online forums.”).



Metadata), but human surveillance was sometimes judged (in the case of Content but not Metadata) to have a large accuracy advantage. Participants further indicated that in the case that full audio contents were transmitted (Content case), they would prefer human to AI surveillance, although they did not expect the State to have the same judgment.

These results are surprising and informative in problematizing how surveillance policy is often discussed or justified. First, ordinary people care as much (or more) about accuracy as they do about privacy when assessing the implementation of a State surveillance program. Second, ordinary people don't expect the State to agree with them in their surveillance preferences. There are clearly significant policy ramifications that follow from these findings. The results particularly suggest a troubling lack of information or presumptive democratic efficacy in the governance of state surveillance.

Readers who are interested in the social sciences, generally, and in the social sciences of algorithms, in particular, may notice the conceptual relationship between this work and prior work looking to algorithmic aversion,<sup>16</sup> or looking to algorithmic appreciation.<sup>17</sup> This work is certainly related insofar as the vignette experiment here seeks to elicit information regarding a preference for humans or algorithms. However, there are some key differences between those findings and the ones I present here. First, the vignette experiments here leave room for human/AI neutrality, a surprisingly common attitude, as shown *infra*. The vignette experiment here does not ultimately require research participants to choose humans over AI or vice versa, and many in fact choose neutrality. Human/technology neutrality, not previously discussed extensively in the related literature, is a contribution of this work that goes neither to algorithmic appreciation nor algorithmic aversion.

Second, the vignette experiment here goes more towards preferences that do not track some objectively better outcome; the preferences here are not evidence of a lack of optimality or rationality by human judges. Rather the preferences measured here, at least in some cases, go to understanding subjective assessments, such as those about privacy, that do not have a clear right or wrong answer about an objectively verifiable fact in the world. In other cases, the judgments measured here establish what are understood as strong priors, that is priors by research subjects that may behave somewhat like normative judgments rather than factual assessments, insofar as participants can both correctly answer questions about the information provided but nonetheless offer judgments that differ from that information. This latter phenomenon as to strong priors may go some way to explaining algorithmic aversion or algorithmic appreciation.

The remainder of this work proceeds as follows. First, I discuss recent technological, political, and consumer history that has brought us to the current situation, in which machine surveillance is both increasingly common and increasingly performant. Through this overview I show the importance of both the

---

16. Berkeley J. Dietvorst, Joseph P. Simmons & Cade Massey, *Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err*, 144 J. EXPERIMENTAL PSYCH.: GEN. 114 (2014).

17. Jennifer M. Logg, Julia A. Minson & Don A. Moore, *Algorithm Appreciation: People Prefer Algorithmic to Human Judgment*, 151 ORG. BEHAV. & HUM. DECISION PROCESSES 90 (2019).

human/machine distinction and the content/metadata distinction to democratic and market discipline alike for surveillance technology use. Having established that the human/machine and content/metadata distinctions have long been important and widely discussed, I look to the legal import of these two distinctions in the four privacy domains. Not only is there inconsistency across the four privacy domains (an unsurprising result), but the constellation of those inconsistencies is itself consistent in that the correlations between the four privacy domains is different for the two distinctions. For example, until recently the Fourth Amendment and wiretap law offered similar treatments of the content/metadata distinction,<sup>18</sup> but the Fourth Amendment looks more like the privacy torts in its treatment of the human/AI distinction.<sup>19</sup> Given this inconsistency, I present the results of a large-scale vignette experiment to seek empirical clarity where judicial and policy decision makers have failed to establish a clear rule for the privacy import of these distinctions. I conclude with a discussion of the policy implications of the empirical findings.

### I. COMPUTATION AND SURVEILLANCE

The appropriate scope and implementation of surveillance by the State has long been an issue of societal and policy concern,<sup>20</sup> and the specter of ever-increasing technology-driven surveillance capabilities raises particularly pressing questions in this vein. A recent version of the technology-driven subset of such questions, which first arose with vigor at the dawn of the era of mass centralization and digitization of data in government hands, concerns the appropriate and permissible use of scalable surveillance in the form of machine, rather than human, aggregation and analysis of information. In the past two decades these concerns—dating from the 1960s—have grown more pressing in view of the vast surveillance capabilities now in the hands of private actors, whose capacities for surveillance have vastly outstripped those of the State in many cases.<sup>21</sup>

---

18. See *infra* Section III.A.

19. See *infra* Table 2. Fourth Amendment and privacy torts jurisprudence offer similarly structured responses to the human/AI distinction in that they appear to yield clear answers that vary by subdomain. Notably, the Fourth Amendment and privacy torts also offer surprisingly similar treatments of the content/metadata distinction, but only in light of recent Fourth Amendment decisions by the Supreme Court.

20. U.S. DEP'T HEALTH, EDUC. & WELFARE, REPORT OF THE SECRETARY'S ADVISORY COMMITTEE ON AUTOMATED PERSONAL DATA SYSTEMS. (1973).

21. The fact that the private sector enjoys tremendous surveillance capabilities is taken for granted, as in *Carpenter v. United States*, 585 U.S. 296, 311 (2018). Even in emphasizing the government's worrying ability to trace individuals in minute and extensive detail, the Justices in *Carpenter* took for granted that the very data that concerned them in a Fourth Amendment inquiry was held by private parties. *Id.* at 311. There, the Court wrote, "Mapping a cell phone's location over the course of 127 days provides an all-encompassing record of the holder's whereabouts. As with GPS information, the time-stamped data provides an intimate window into a person's life, revealing not only his particular movements, but through them his 'familial, political, professional, religious, and sexual associations.' . . . With just the click of a button, the Government can access *each carrier's* deep repository of historical location information at practically no expense." *Id.* (emphasis added) (quoting *United States v. Jones*, 565 U.S. 400, 415 (2012) (Sotomayor, J., concurring)). Consider also that advocates

Although the initial “computerization” of government and private sector records in the 1960s and 1970s was accompanied with some degree of fanfare about appropriate governance of the awesome powers of compiled information, most notably in the development of the Fair Information Practices,<sup>22</sup> there has not been consistently vigorous or productive activity in lawmaking as surveillance capacities have continued to grow at rapid rates. Most significantly, the far more rapid and socially significant growth of communications and electronic surveillance capacity of recent decades has not resulted in concomitantly rapid innovation in the judicial science of surveillance.<sup>23</sup> Yet despite the lack of consistent legal development to ensure certainty in the application of law to emerging surveillance technologies,<sup>24</sup> policymakers and ordinary people seem to share some basic intuitions about what is or is not privacy invasive, which suggests that some basic surveillance norms and intuitions have developed. Evidence about these norms and intuitions is presented in Part IV.

---

and legislators alike have raised concerns about governmental purchases of data from private entities, highlighting another example of how the private sector is positioned to provide, or deny, surveillance capacity to the government. *See* Fourth Amendment Is Not for Sale Act, H.R. 4639, 118th Cong. (2023); *After House Passes Fourth Amendment Is Not For Sale Act, ACLU Urges Senate to Stop Government from Spying on Americans Without a Warrant*, ACLU (Apr. 17, 2024) <https://www.aclu.org/press-releases/house-passes-fourth-amendment-is-not-for-sale-act> [https://perma.cc/6XLN-PRWV] (“When the government wants to obtain Americans’ private information, like where they live, what doctor’s office they visited, or who they are dating, the Fourth Amendment requires it to go to court and obtain a warrant. But for years now, federal agencies, including the Internal Revenue Service [see Byron Tau, *IRS Used Cellphone Location Data to Try to Find Suspects*, WALL STR. J., (June 19, 2020, 1:46 PM) <https://www.wsj.com/articles/irs-used-cellphone-location-data-to-try-to-find-suspects-11592587815> [https://perma.cc/GN4Q-BVFR]], and the Department of Defense, [see Joseph Cox, *How the U.S. Military Buys Location Data from Ordinary Apps*, VICE (Nov. 16 2020, 11:35 AM), <https://www.vice.com/en/article/us-military-location-data-xmode-locate-x/> [https://perma.cc/BSB2-M2RW]] have been buying their way around this requirement by purchasing Americans’ sensitive information from data brokers. These companies often obtain this information through common applications, like The Weather Channel, [see Suzanne Smalley, *Exploring the Surveillance Partnership Between the Government and Data Brokers*, THE REC. (Mar. 21, 2024), <https://therecord.media/byron-tau-interview-surveillance-government-data-brokers> [https://perma.cc/P6KS-CH2L]] or Tinder, [see Natasha Singer & Aaron Krol, *Grindr and OkCupid Spread Personal Details, Study Says*, N.Y. TIMES, (Jan. 13, 2020, <https://www.nytimes.com/2020/01/13/technology/grindr-apps-dating-data-tracking.html> [https://perma.cc/Q5JA-E64Z]], without users realizing it, and then the government uses it to track people’s location without a warrant or probable cause — or even suspicion that anyone in the dataset had done anything wrong.”).

22. *Fair Information Practice Principles (FIPPs)*, FED. PRIV. COUNCIL, <https://www.fpc.gov/resources/fipps/>.

23. Consider a recent district court’s observation in a prominent geofence warrant case that “Fourth Amendment law develops in a *slow drip*.” *United States v. Chatrie*, 590 F. Supp. 3d 901, 925 (E.D. Va. 2022) (emphasis added).

24. Consider the recent circuit split as to the Fourth Amendment import of geofence warrants. Michael Berman, *Are Geofence Warrants Permissible?—A Circuit Split*, E-DISCOVERY LLC (Aug. 12, 2024) <https://www.ediscoveryllc.com/are-geofence-warrants-permissible-a-circuit-split/> [https://perma.cc/8HKW-PJ4K].

*A. Surveillance by Humans Versus by Machines*

Discussions in the press provide numerous examples of common reactions to both governmental and commercial surveillance. Consider the national outcry resulting from revelations of Stellarwind, the warrantless wiretapping program put in place after the September 11 attacks.<sup>25</sup> There was substantial bipartisan consensus as to the outrageousness of the government's actions,<sup>26</sup> showing the widespread and strong feelings about the perceived privacy-invasive nature of that program. U.S. politics have evolved substantially since that time, but even in the current era of more polarized and fractious politics, skepticism about state surveillance<sup>27</sup> (and, also, dissatisfaction with the current level of private sector surveillance<sup>28</sup>) remain bipartisan issues.

The logic of defending the controversial Stellarwind program demonstrated some commonly shared intuitions, too. Defenders of that program leaned heavily on the *automated, nonhuman* nature of the surveillance in rhetoric justifying the surveillance and arguing that any associated privacy invasion was minimal. Following the Stellarwind revelation, Judge Richard Posner took to the pages of *The Washington Post* to defend the state of post-9/11 surveillance. Though acknowledging the need to protect privacy, Posner defended the implementation of these programs in part through the *computational* nature of the surveillance:

The collection, mainly through electronic means, of vast amounts of personal data is said to invade privacy. But machine collection and processing of data cannot, as such, invade privacy. Because of their volume, the data are first sifted by computers . . . This initial sifting, far from invading privacy (a computer is not a sentient being), keeps most private data from being read by any intelligence officer.<sup>29</sup>

---

25. James Risen & Eric Lichtblau, *Bush Lets U.S. Spy on Callers Without Courts*, N.Y. TIMES, (Dec. 16, 2005), <https://www.nytimes.com/2005/12/16/politics/bush-lets-us-spy-on-callers-without-courts.html> [<https://perma.cc/2AQL-65N9>].

26. Dan Roberts, Spencer Ackerman & Alan Travis, *NSA Surveillance: Anger Mounts in Congress at 'Spying on Americans'*, THE GUARDIAN (June 11, 2013, 8:45 PM), <https://www.theguardian.com/world/2013/jun/12/anger-mounts-congress-telephone-surveillance-programmes> [<https://perma.cc/NPB9-FRZC>].

27. See, e.g., *Despite Bipartisan Outcry, Senate Betrays the Fourth Amendment and Passes Bill to Expand Warrantless Government Surveillance*, ACLU (Apr. 20, 2024, 1:00 PM) <https://www.aclu.org/press-releases/senate-reauthorizes-and-expands-section-702-surveillance> [<https://perma.cc/V3K8-6ERX>] (citing “bipartisan outcry” about FISA § 702 renewal).

28. See, e.g., Andrew Kingman & Willy Martinez, *Data Privacy and Protection in the US: A Sign of Bipartisan Progress*, INT'L ASS'N PRIV. PROS., <https://iapp.org/news/a/data-privacy-and-protection-in-the-us-a-sign-of-bipartisan-progress> [<https://perma.cc/8SD7-HY6A>] (“In an era of persistent political polarization, a noteworthy trend at the state level offers hope for those nostalgic for bipartisan cooperation: data privacy protections.”).

29. Richard A. Posner, *Our Domestic Intelligence Crisis*, WASH. POST (Dec. 21, 2005), <https://www.washingtonpost.com/archive/opinions/2005/12/21/our-domestic-intelligence-crisis/a2b4234d-ba78-4ba1-a350-90e7fbb4e5bb> [<https://perma.cc/DLR9-W5L5>].

Posner suggested that privacy concerns could be addressed or mitigated by computational means, with the implicit assumption that privacy considerations entail a human observer rather than mere data analysis. While this implicit assumption may seem quite simplistic (if not downright wrong) when made explicit, and while Posner wrote in a very different era from the current one—an era with far fewer automated observations and decisions about individuals—Posner’s basic intuition that the human/machine distinction is important seems to be an intuition still in play today across a wide domain of activities. Posner is far from the only surveillance defender to emphasize the privacy differences between being surveilled by a computer versus a human—for example, the same logic was invoked seven years later in the wake of Edward Snowden’s revelations about another form of warrantless government surveillance.<sup>30</sup>

The “computers can’t invade your privacy” intuition is also reflected in some news stories about commercial surveillance. Consider a scandal implicating several large technology companies in 2019. That year, in a wave of revelations, the public learned that several firms marketing in-home voice assistant products were using humans to transcribe recorded conversations from digital assistants. A rapid succession of leaks to the press indicated that human workers were listening to and transcribing recordings from Google’s digital home assistant, Amazon’s digital home assistant, and Apple’s iPhone assistant.<sup>31</sup>

The very wording of news headlines suggested that the sensational element of the leaks and the breadth and depth of the resulting outrage were responsive to the use of *human* transcribers rather than other potentially problematic aspects of these programs. There were many potentially troubling data collection and transmission practices necessarily entailed by humans listening to these conversations, but it seemed to be *the fact of human listening* that drove the strong public response to these news stories. A sample of such headlines included the following:

- “Apple Contractors ‘Regularly Hear Confidential Details’ on Siri Records”<sup>32</sup>

---

30. For a discussion of other examples of distinctions made between humans and computers as observers, see Jay Stanley, *Computers vs. Humans: What Constitutes A Privacy Invasion?*, ACLU (July 2, 2012), <https://www.aclu.org/news/national-security/computers-vs-humans-what-constitutes-privacy> [<https://perma.cc/AS36-DTXW>].

31. For Google Assistant, see, for example, James Vincent, *Yep, Human Workers are Listening to Recordings from Google Assistant, Too*, THE VERGE (July 11, 2019, 5:48 AM), <https://www.theverge.com/2019/7/11/20690020/google-assistant-home-human-contractors-listening-recordings-vrt-nws> [<https://perma.cc/5WRK-PU2L>]. For Amazon Alexa, see, for example, Mariella Moon, *An Amazon Employee Might Have Listened to Your Alexa Recording*, ENGADGET (Apr. 11, 2019), <https://www.engadget.com/2019-04-11-amazon-alexa-voice-recording-human-review.html> [<https://perma.cc/WFZ4-3CME>]. For Apple Siri, see, for example, Alex Hern, *Apple Contractors ‘Regularly Hear Confidential Details’ on Siri Recordings*, THE GUARDIAN (July 26, 2019, 12:34 PM), <https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings> [<https://perma.cc/GXX9-JB2V>].

32. Hern, *supra* note 31.

- “Yep, Human Workers Are Listening to Recordings From Google Assistant, Too”<sup>33</sup>
- “An Amazon Employee Might Have Listened to Your Alexa Recording”<sup>34</sup>

None of the companies implicated in the scandal had explicitly disclosed the use of human transcription in their privacy policies. As the fact of human listening emerged, all three companies responded that the practice of using human listeners was logically entailed by and operationally necessary for their services.<sup>35</sup> But it was clear from the public backlash that consumers didn’t have the same understanding. Ultimately, all three companies added explicit disclosures<sup>36</sup> about human listening as well as product settings allowing consumers to refuse consent for human review.<sup>37</sup> These subsequent changes suggest that the companies identified human transcription—not recording or transmission of audio—as the problematic element of their commercial surveillance programs. Interestingly, as of the time of writing, Amazon has revisited the policy and has announced that it will be phasing out user ability to disable transmission of human voice recordings; it remains to be seen whether this will engender enough consumer backlash to stop the policy, but in any case the policy was again clearly contentious, making national news headlines even in mainstream media.<sup>38</sup>

One might go further and infer that the firms likewise concluded that machine observation was not a problem. Consider Apple’s modification to its privacy policy. Following the 2019 outcry, Apple modified its privacy policy to indicate that *machine transcripts* of user interactions with Siri *would still be used* for product

33. Vincent, *supra* note 31.

34. Moon, *supra* note 31.

35. For Apple not disclosing use of human transcribers in its privacy policy, see Hern, *supra* note 31. For Google Assistant, see Vincent, *supra* note 31.

36. For Amazon, see Natasha Lomas, *Amazon Quietly Adds ‘No Human Review’ Option to Alexa Settings as Voice AIs Face Privacy Scrutiny*, TECHCRUNCH (Aug. 3 2019, 5:46 AM), <https://techcrunch.com/2019/08/03/amazon-quietly-adds-no-human-review-option-to-alexa-as-voice-ais-face-privacy-scrutiny/> [https://perma.cc/EYT4-83SN]. For Apple, see *Improving Siri’s Privacy Protections*, APPLE NEWSROOM (Aug. 28, 2019), <https://www.apple.com/newsroom/2019/08/improving-siris-privacy-protections/> [https://perma.cc/3YZ9-D44N]. For Google and a comparison to Amazon and Apple, see Dieter Bohn, *Google Is Reducing How Much Audio It Saves for Human Review*, THE VERGE (Sept. 23, 2019, 3:01 AM), <https://www.theverge.com/2019/9/23/20878710/google-assistant-audio-recording-policy-hotword-human-review> [https://perma.cc/DZQ2-MNMU].

37. Human audio review was implemented by Google and Apple on an opt-in basis and Amazon on an opt-out basis.

38. Sharon Harding, *Everything You Say to Your Echo Will Be Sent to Amazon Starting on March 28*, ARS TECHNICA (Mar. 14, 2025, 4:59 PM), <https://arstechnica.com/gadgets/2025/03/everything-you-say-to-your-echo-will-be-sent-to-amazon-starting-on-march-28/> [https://perma.cc/4J46-KRJQ]; see also Anthony Robledo, *Amazon Is Removing an Echo Privacy Setting That Keeps Alexa Recordings from the Company, USA TODAY* (Mar. 17, 2025), <https://www.usatoday.com/story/tech/2025/03/17/amazon-echo-alexa-reporting-privacy/82503576007/> [https://perma.cc/4CUN-ZZEL].

improvement even in the case of users who opted out of *human transcription*.<sup>39</sup> There was no option provided to avoid machine transcription. This change suggests that Apple concluded that the use of human listeners offended users' privacy sensibilities in a way that content recording, data transmission, and AI transcription did not.<sup>40</sup> Further, Apple's likely conclusion resonates with behavioral findings (discussed in Section II.F) from human-computer interaction studies showing that a machine interface and a human interface can generate different information sharing choices.<sup>41</sup>

### *B. Surveillance of Content Versus Metadata*

In addition to using the human/machine distinction to justify their surveillance programs, state and commercial actors likewise rely on justifications grounded in the content/metadata distinction when justifying or defending their surveillance implementations.<sup>42</sup> Just as news coverage and public responses to Stellarwind gave Judge Posner the opportunity to air his (seemingly, widely shared) insight that human and machine observation are not equivalent from a privacy perspective, so too the subsequent Edward Snowden leaks provided the opportunity to observe how the content/metadata distinction is operationalized for purposes of defending or justifying surveillance. For example, President Barack Obama leaned heavily on an intuition that observations of content and metadata are not equivalent. Consider President Obama's 2013 defense of state surveillance in the days following Snowden's 2013 revelations about the extent of warrantless telephone record surveillance.<sup>43</sup>

My assessment and my team's assessment about these (monitoring programs) was that they help us prevent terrorist attacks, and the modest encroachments on privacy that are involved—in getting phone numbers

---

39. *Improving Siri's Privacy Protections*, *supra* note 36.

40. It is not clear on what basis the companies drew their conclusion, but the results of the present study could provide further clarification on this point.

41. Cf. Yafit Lev-Aretz, "Nobody Is Watching Me": Towards Human-Centric Privacy and Humanless Information Protection, 26 VA. J.L. & TECH. 1 (2022) (reviewing behavioral literature that shows ways in which humans can be prompted to treat anthropized technology as if it is human).

42. Defining metadata is quite problematic, and it is not clear how to draw a line between content as compared to metadata in many cases. This is particularly so with the advent of (1) widely available joinable datasets that can turn seeming metadata into data that has attributes of content when joined to other data and (2) machine learning techniques from which much content-like information can be deduced from what were previously thought to be metadata. Nonetheless, the distinction, if problematic or possibly close to meaningless, remains relevant both to privacy law and to ordinary people's privacy judgments, as is discussed in this section.

43. Glenn Greenwald, *NSA Collecting Phone Records of Millions of Verizon Customers Daily*, THE GUARDIAN (June 6, 2013, 6:05 PM), <https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order> [<https://perma.cc/K76W-V2D7>]; see also Posner, *supra* note 29.

or data on call duration without a name attached [*and without*] looking at content . . . was worth us doing.<sup>44</sup>

President Obama emphasized that it was not content that was collected,<sup>45</sup> but rather only data *about* communications, that is, metadata. Thus, Obama argued, any potential privacy intrusion was “modest” because content was not collected.<sup>46</sup>

The content/metadata distinction is influential in consumer markets as well as in rhetoric about state surveillance. Consider the January 2021 backlash to a WhatsApp privacy policy update. Many WhatsApp users voiced concerns that WhatsApp’s new policy indicated plans to access the *content* of private conversations (a claim WhatsApp vigorously denied).<sup>47</sup> In the wake of the backlash against WhatsApp, alternative messaging apps were downloaded at record numbers, suggesting an exodus from WhatsApp motivated by reactions to the new privacy policy and attendant news coverage.<sup>48</sup> Regardless of the correct interpretation of WhatsApp’s new privacy policy, the strong market reaction suggests that WhatsApp users made a sharp distinction between surveillance of behavioral metadata (which was already widely known to be collected in abundance by WhatsApp) and scanning of conversational content.<sup>49</sup> As is shown in Figure 1, WhatsApp in fact already

44. Seth Cline, *Obama Defends Sweeping NSA Surveillance Program*, U.S. NEWS (June 7, 2013, 2:19 PM), <https://www.usnews.com/news/articles/2013/06/07/obama-defends-sweeping-nsa-surveillance-program> [<https://perma.cc/5XXN-T5WH>] (emphasis added).

45. It is unclear whether Obama would have understood names to be content or metadata. As will be discussed below, in some cases communications privacy laws may include identity in the protections given to content. On the other hand, the structure of the Stored Communications Act puts subscribed name in the category of records information, which is entitled to less protection than the contents of communications. This difficulty only serves to highlight the conceptual fuzziness, if not deeply problematic nature, of a content/metadata distinction.

46. See Cline, *supra* note 44. Of course, it has long been obvious, and not just to technical thinkers, that a great deal can be learned about content from what is typically viewed as non-content data, and that therefore the distinction is not so clear. Justice Stewart’s dissent in *Smith v. Maryland* argued that the installation of a pen register did constitute a Fourth Amendment search. 442 U.S. 735, 748 (1979) (“The numbers dialed from a private telephone—although certainly more prosaic than the conversation itself—are not without ‘content.’ . . . [A list of numbers called] easily could reveal the identities of the persons and the places called, and thus reveal the most intimate details of a person’s life.”). More recently, empirical researchers have likewise demonstrated this insight. See Jonathan Mayer, Patrick Mutchler & John C. Mitchell, *Evaluating the Privacy Properties of Telephone Metadata*, 113 PROC. NAT’L ACAD. SCIS. 5536 (2016).

47. Abrar Al-Heeti, *WhatsApp Responds to Concerns Over Privacy Policy Update*, CNET (Jan. 12, 2021, 1:55 PM), <https://www.cnet.com/news/whatsapp-responds-to-concerns-over-privacy-policy-update/> [<https://perma.cc/P8H2-RDK2>] (discussing that WhatsApp took pains to emphasize that while it did collect some forms of data to share with Facebook, and had for years, it could not see the content of messages).

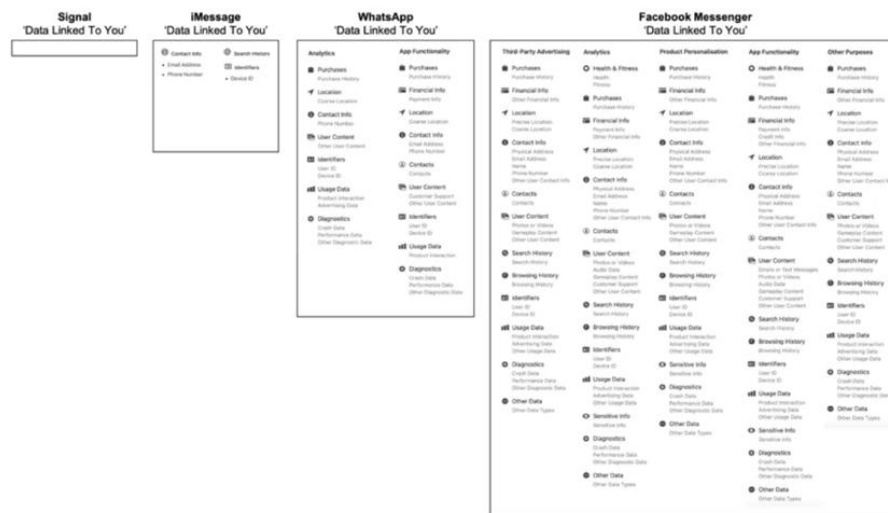
48. Arjun Kharpal, *Signal and Telegram Downloads Surge After WhatsApp Says It Will Share Data with Facebook*, CNBC (Jan. 12, 2021, 12:41 AM) <https://www.cnbc.com/2021/01/12/signal-telegram-downloads-surge-after-update-to-whatsapp-data-policy.html> [<https://perma.cc/EN8E-PLR4>].

49. WhatsApp was already known to collect far more metadata than its competitors, but



collected a substantial amount of information *about* users of the app, but this apparently did not register for consumers as privacy invasive in the same way as the suspected scanning of the contents of conversations. This provides strong circumstantial evidence that consumers were prepared to discount or disregard even substantial amounts of publicly disclosed metadata collection but were nonetheless quite sensitive to the possibility of content scanning.<sup>50</sup>

**Figure 1:** Consumer communications apps report substantially varying metadata collection practices.<sup>51</sup>



this did not cause the public backlash, the likely loss of users from WhatsApp, or the influx to more private messaging platforms (such as Signal) that was apparently caused by the privacy policy update and concerns about privacy of content. *See Ben Lovejoy, App Privacy Labels Show Stark Contrasts Among Messaging Apps*, 9TO5MAC (Jan. 4, 2021, 7:20 AM), <https://9to5mac.com/2021/01/04/app-privacy-labels-messaging-apps/> [https://perma.cc/5YL8-D978]. This blog post predated the WhatsApp privacy policy scandal but pointed out that in privacy labels newly released in the Apple app store in the previous month (December 2020) WhatsApp indicated a wide scale collection of metadata, particularly in comparison to competitors such as Signal and iMessage.

50. There may be and likely are alternative, nonexclusive explanations for this happening. For example, the decision of news media to emphasize or not a particular privacy practice would also render the practice known to or salient to consumers.

51. Figure reproduced from Ben Lovejoy, *App Privacy Labels Show Stark Contrasts Among Messaging Apps*, 9TO5MAC (Jan. 4, 2021, 7:20 AM), <https://9to5mac.com/2021/01/04/app-privacy-labels-messaging-apps/> [https://perma.cc/PY2Z-8DBG].

*C. Summing Up*

We have seen that two key distinctions matter both in classic state surveillance scenarios and also in commercial surveillance:

- (1) the use of a human versus machine observer; and
- (2) the observation of content versus metadata.

These distinctions *are not independent of one another*. Technology adds heftier guarantees to the content/metadata distinction in the form of robust technical safeguards to ensure compliance with privacy promises—promises that are far less tenable and more like useful fictions in the case of human observation. For example, a human can promise to watch a conversation without listening but cannot credibly guarantee that she will do so; computer code can ensure this same outcome (and can likewise be audited to assure compliance). Not only can computer code be audited, but surveillance can be implemented with built in hardware guarantees, such as audio scramblers or limitations to subsets of visual channels.<sup>52</sup>

The recent, rapid advance of digital surveillance technology, in both hardware and software, makes the distinction between human and AI surveillants operational (and thus, relevant). In the past, the degree of state surveillance was cabined by law<sup>53</sup> or otherwise limited by incidental practical constraints (like cost or technical limitations<sup>54</sup>) rather than by strong technological guarantees. Now, there are technological options<sup>55</sup> to perform surveillance *both* pervasively (a topic of rising importance in Fourth Amendment jurisprudence) *but also* in a technologically constrained manner, such as by selecting certain kinds of sensors but not others, or by writing and deploying code<sup>56</sup> that only records certain forms or volumes of information. Technology offers opportunities for systematically blinding observers by choice in ways previously available only via legal rules.

To sum up, we have seen that the distinction between human and machine observation has been a salient one both in defending state surveillance and in tuning

---

52. For example, to block the infrared light found problematic in *Kyllo*. *Kyllo v. United States*, 533 U.S. 27 (2001).

53. For example, warrants that delineated the hours during which a wiretapped line could be audited or rooms in a target's home that could be wired for surveillance. One Massachusetts warrant application specifically stated: "[T]he interception is required to be maintained for a period of 15 calendar days, commencing on the date of installation of the intercepting device, and that the hours of each day during which wire communications may be reasonably expected to occur are those between the hours of 11:00 a.m. to 7:30 p.m.' We find this analysis persuasive." *Commonwealth v. Vitello*, 327 N.E. 2d 819, 846 (1975).

54. Consider Justice Alito's concurrence in *United States v. Jones*: "In the pre-computer age, the greatest protections of privacy were neither constitutional nor statutory, but practical." 565 U.S. 400, 429 (2012).

55. It is worth recognizing, as well, the existence of privacy-enhancing technologies (PETs). PETs are deployed by those who are surveilled to limit the efficacy of that surveillance. Here, the topic of interest is how surveillants might limit the scope of their investigations through technological precommitment mechanisms, but PETs apply the same logic as deployed by surveilled communities.

56. And such code can itself be audited.

consumer surveillance to conform to acceptable bounds. We have likewise seen the importance of the content/metadata distinction both in defenses of state surveillance and in responses to products in consumer markets. Finally, we have discussed the synergies between these two distinctions, as machine surveillance adds new vigor and new plausible possibilities to the content/metadata distinction. These distinctions are widely deployed and seem to matter to voters and consumers alike; further, these distinctions are technically possible and powerful. But what has the law made of these distinctions, if anything?

## II. HUMAN VERSUS AI PRIVACY

Despite a seemingly widespread intuition that the algorithmic gaze is unimportant as to privacy, legal treatment of humans and AI has been inconsistent. This inconsistency is neither new nor specific to digital technologies. Rather, the more general question as to the import of human versus machine observation has surfaced over the course of more than seventy years of legal authority.<sup>57</sup>

The notion that privacy law designates a particular role for human observers intrinsic to the very concept of privacy intrusions, specifically as to the identity of perpetrators, has already been remarked upon in doctrinal legal scholarship. A formal distinction of machine versus human observation dates back at least two decades. Writing in 2001, Daniel Solove observed that the concerns that traditionally justified legal protections for privacy (related to humans censoring themselves due to the observation by other humans) were not fit for purpose in assessing privacy threats created by computational compilation and aggregate description of humans as data objects.<sup>58</sup>

More recently, Yafit Lev-Aretz presented a critique of the implicitly assumed importance of the human gaze in privacy law, demonstrating the ways in which the assumption of a human observer has been a core tenet of privacy—one that, Lev-Aretz argued, has distorted the role of privacy law in regulating modern informational incursions resulting from automated surveillance.<sup>59</sup> Lev-Aretz has proposed limiting privacy to a human-centric understanding and distinguishing privacy in law from information protection, arguing that the latter is effectively a human-less concern.<sup>60</sup> Lev-Aretz argues that creating such a division would enhance the appropriate values germane to the distinct but related human needs and civil rights implicated by (human-centric) privacy and, separately, information protection.

In addition to the broad sweep Solove and Lev-Aretz have applied to information privacy law, scholars of surveillance law have also addressed this issue. Writing on the Fourth Amendment in 2011, Matthew Tokson offered both a novel interpretation

---

57. The earliest authority this work cites for addressing the human/machine distinction is *On Lee v. United States*, 343 U.S. 747 (1952). I do not rule out the possibility that earlier case law, statutes, or other legal sources may also have discussed the distinction in some form.

58. Daniel J. Solove, *Privacy and Power: Computer Databases and Metaphors for Information Privacy*, 53 STAN. L. REV. 1393, 1393 (2001).

59. Lev-Aretz, *supra* note 41.

60. Lev-Aretz's critique mirrors European thinking on this question, in that Europe has traditionally demarcated privacy and data protection as separate if highly related social concerns and rights of the individual.

of third-party doctrine case law as well as empirical evidence drawn from an original survey to argue that the third-party doctrine was not fit for purpose in an emerging digital age and so should not be extended to cover automated transfers of information.<sup>61</sup> His arguments would later be borne out, if perhaps only indirectly, by the Supreme Court's decisions in *United States v. Jones*, decided in 2012, and *Carpenter v. United States*, decided in 2018.<sup>62</sup>

A pair of articles has likewise explored the issue of whether a computer, rather than a human, can intercept a communication under the Wiretap Act. Writing in 2012, Bruce Boyden surveyed federal wiretapping case law that demonstrated situations in which a machine, but perhaps not a human, had undertaken actions that might plausibly constitute an interception.<sup>63</sup> Boyden concluded that a machine alone likely could not intercept a communication. But, writing in 2014, Kevin Bankston and Amie Stepanovich surveyed the same line of cases and reached the opposite conclusion.<sup>64</sup> Unlike the case of new precedent following Tokson's arguments about the Fourth Amendment, the Supreme Court has not offered additional guidance with regard to machine interceptions under the Wiretap Act to address the ambiguity suggested by the opposing conclusions of this line of scholarship.

The scholarly investigation of the human versus machine gaze remains incomplete. For example, scholarship has not yet reckoned with how the human/AI distinction manifests in the privacy torts and other private law causes of actions, but these private law tools have been the most effective, or at least the most relied upon, legal theories in challenging the awesome surveillance capacities of the private sector. This work expands upon existing scholarship in two ways. First, this work *documents the lack of consistency as to the legal importance* of the AI/human distinction across four core areas of privacy law: the Fourth Amendment, communications privacy, the privacy torts, and consumer protection law.<sup>65</sup> If the legal use or nonuse of an AI/human distinction relies on widely shared intuitions, it is troubling that the distinction is not treated uniformly, particularly where the incantation of "reasonable expectations"<sup>66</sup> is pervasive across these domains of privacy law. Second, having identified the uncertainty in the law, this work proposes an empirical methodology grounded in a vignette experiment, permitting us to draw causal conclusions into how ordinary people assess the privacy import of human versus AI surveillance.

---

61. Matthew Tokson, *Automation and the Fourth Amendment*, 96 IOWA L. REV. 581, 581 (2011).

62. *United States v. Jones*, 565 U.S. 400 (2012); *Carpenter v. United States*, 585 U.S. 296 (2018).

63. Bruce E. Boyden, *Can a Computer Intercept Your Email?*, 34 CARDOZO L. REV. 669, 669 (2012).

64. Kevin Bankston and Amie Stepanovich on "When Robot Eyes Are Watching You: The Law & Policy of Automated Communications Surveillance," WE ROBOT, 2014 (Mar. 31, 2014), <https://robots.law.miami.edu/2014/?p=1446> [<https://perma.cc/5JA7-VJYZ>].

65. This work is not the first to recognize and name this inconsistency. See *supra* Section II.

66. See Suzman, *supra* note 11.

*A. The Fourth Amendment*

Arguably, the earliest Supreme Court cases to deal with the human/machine distinction in a privacy context goes back as far as the 1952 case of *On Lee*.<sup>67</sup> In *On Lee*, a government informant elicited incriminating statements from a suspect, and these statements were also heard *via radio* by a federal agent. The agent later offered testimony because the informant could not be located, raising a novel legal issue as to whether the warrantless transmission of the conversation was a Fourth Amendment search. The Supreme Court found that the radio transmission of the conversation was not a Fourth Amendment search. The Court further found that the presence, or use, of a radio did not make a difference to the outcome. The Court found that the plaintiff was speaking carelessly and with (misplaced) confidence to a trusted interlocutor, and that this misplaced trust—and not the device—was the reason the conversation was overheard by a human. “This was due to aid from a transmitter and receiver, to be sure, but with the same effect on his privacy as if agent Lee had been eavesdropping outside an open window.”<sup>68</sup> This language implicitly assumed privacy neutrality of the device; use of the recording device was found to be irrelevant to the privacy considerations invoked by challenging a police practice under the Fourth Amendment.

About ten years later, the Court again had the opportunity to comment on the human/machine distinction in *Lopez*,<sup>69</sup> in which a machine *recorded* audio. Rather than merely enabling a human to hear audio, as was the case in *On Lee*, the machine in *Lopez* arguably *did the hearing and remembering* (in some sense of those words). The warrantless recording was offered in evidence in a criminal proceeding, in addition to human testimony about the same conversation. Once again, the Court found that this warrantless use of a machine was not a Fourth Amendment search, and in doing so gave more explicit commentary on its view of the import of machine use.

Stripped to its essentials, petitioner’s argument [that the recording constituted a Fourth Amendment search] amounts to saying that he has a constitutional right to rely on possible flaws in the agent’s memory, or to challenge the agent’s credibility without being beset by corroborating evidence that is not susceptible of impeachment. For no other argument can justify excluding an accurate version of a conversation that the agent could testify to from memory. We think the risk that petitioner took in offering a bribe to [someone] fairly included the risk that the offer would be accurately reproduced in court, whether by faultless memory or mechanical recording.<sup>70</sup>

The *Lopez* Court viewed a machine as a superior observer as compared to a human because the machine was presumed not to have flaws in its “memory,” in contrast to a human. The *Lopez* court saw veracity or accuracy advantages to the machine.

---

67. *On Lee v. United States*, 343 U.S. 747 (1952).

68. *Id.* at 754.

69. *Lopez v. United States*, 373 U.S. 427 (1963).

70. *Id.* at 439.

Nonetheless, the *Lopez* Court didn't indicate any perceived privacy invasion from the use of a recording device.

Neither the *On Lee* nor the *Lopez* Court found a Fourth Amendment violation where a government was using false friends to obtain evidence. Indeed, both decisions have been criticized for the Court's resistance to the notion that government informants or undercover police officers could elicit false confidences and therefore elicit evidence without any governance via Fourth Amendment discipline. Thus, the fact that the *On Lee* and *Lopez* courts appeared untroubled by any privacy concerns specific to machines does not necessarily go directly to the question of *distinguishing* human and machine observation. In those cases—especially *On Lee*—the human observation was likewise found to be untroubling. The fact that the machine observation was not a Fourth Amendment problem put the machine observation on the same level as the human observation, untroubling (opined the Justices) from a privacy perspective. Nonetheless, the language from these cases suggests that even if neither a human nor a machine was a Fourth Amendment challenge, that the machine was even less concerning to the Court than would be a human listener.

*On Lee* and *Lopez* were decided before the 1967 landmark decision of *Katz v. United States*,<sup>71</sup> which substantially uprooted then established understandings of Fourth Amendment protections. Indeed, *Katz* so uprooted the notion of a Fourth Amendment search that it was thought to have likely overturned the widely cited earlier established trespass theory of Fourth Amendment search from the *Olmstead v. United States*<sup>72</sup> decision.<sup>73</sup> Therefore there is also value in asking whether post-*Katz* jurisprudence might have challenged the earlier no-privacy-import language highlighted in *On Lee* and *Lopez*.

Coming in 1971, *United States v. White* clarified that *On Lee* was still good law and extended the scenario of the machine-assisted human ear or human-memory to include combining radio transmission with a recording device at the receiving end of the transmission.<sup>74</sup>

Concededly, a police agent who conceals his police connections may write down for official use his conversations with a defendant and testify concerning them, without a warrant authorizing his encounters with the defendant and without otherwise violating the latter's Fourth Amendment rights. For constitutional purposes, no different result is required if the agent . . . simultaneously transmits the conversations *either to recording equipment located elsewhere or to other agents monitoring the transmitting frequency*. If the conduct and revelations of an agent operating without electronic equipment do not invade the defendant's constitutionally justifiable expectations of privacy, neither does a simultaneous recording of the same . . .<sup>75</sup>

---

71. 389 U.S. 347 (1967).

72. 277 U.S. 438 (1928).

73. See *United States v. Jones*, 565 U.S. 400 (2012) (explaining that this trespass standard had not been overruled but rather supplemented).

74. 401 U.S. 745 (1971).

75. *Id.* at 751 (emphasis added) (citations omitted).

Even in the post-*Katz* era, the Court treated listening by a human agent or by recording equipment identically—the distinction remained without a constitutional difference. It wasn't that the Court was incapable of seeing difference. Rather the Court saw no difference when it came to the privacy expectations of a person that were constitutionally protected even as the Court recognized differences in the relative abilities or evidentiary value of a machine rather than a human recording.

An electronic recording will many times produce a more reliable rendition of what a defendant has said than will the unaided memory of a police agent . . . [W]e are not prepared to hold that a defendant who has no constitutional right to exclude the informer's unaided testimony nevertheless has a Fourth Amendment privilege against a more accurate version of the events in question.<sup>76</sup>

Thus, a series of cases in the early history of recording or transmitting conversations conducted by government informants or agents gave the Court repeated opportunities to distinguish human versus machine observation, and the Court repeatedly found that machine, rather than human transmission or recording, was a distinction without a constitutional difference. This may have been in part because the use or presence of the human was itself often found to have little or no constitutionally problematic aspect, and therefore where the machine was implicitly the lesser of a privacy problem, the machine too was constitutionally unproblematic. This pattern of reasoning predated and also survived the landmark Fourth Amendment shift under *Katz*. In both pre- and post-*Katz* cases, the Court recognized that machines had salient differences compared to humans (range of hearing, imperfections of memory), but the Court did not find these differences changed the Fourth Amendment outcome.

So, early Fourth Amendment jurisprudence provided some commentary from the Court in which it recognized differences as between humans and machines, but nonetheless did not reach any holdings that turned on such an outcome. In more recent cases, the Court has taken on more examples of super-human surveillance aided by machines, and this has provided a larger sample of case law to understand the Court's views. Scholars have found the Court distinguishing between human and machine observation in cases of more recent vintage. For example, in reviewing cases from recent decades, Orin Kerr proposed to understand Fourth Amendment jurisprudence under an "exposure-based approach"; under this theory, a search occurs "when information from or about the data is exposed to possible human observation."<sup>77</sup> Consistent with Kerr's rationalization of the case law, Matthew Tokson also reviewed more recent precedent to find that "[i]n cases involving new

---

76. *Id.* at 753. Note that the Court's assumption that recordings are necessarily "more accurate" is not clearly correct. For example, recordings may fail to record context that contemporaneous human interlocutors or listeners will access, and therefore recordings alone may mislead rather than necessarily present the most accurate evidence as to a communication. See Clifford S. Fishman, *Recordings, Transcripts, and Translations as Evidence*, 81 WASH. L. REV. 473 (2006).

77. Lev-Aretz, *supra* note 41, at 15 (quoting Orin S. Kerr, *Searches and Seizures in a Digital World*, 119 HARV. L. REV. 531, 549–57 (2005)).

technologies, the [Supreme] Court's holdings support the idea that no Fourth Amendment 'search' occurs until electronic information is exposed to a human being."<sup>78</sup>

Tokson's work is rich in examples demonstrating a Supreme Court logic in which observation by a human is a necessity for a privacy intrusion to rise to the level of a Fourth Amendment search. In *United States v. Karo*,<sup>79</sup> the Supreme Court concluded that the placement of a homing beacon among a suspect's belongings was not a Fourth Amendment search (that is, did not violate the suspect's reasonable expectations of privacy) because the transmission was not monitored by law enforcement agents. The Court, Tokson argued, thus clarified that it is not the mere existence or presence of a device but its exploitation by law enforcement that brings a physical measuring device within the ambit of the Fourth Amendment.<sup>80</sup> *Karo* continues to be cited for this proposition.<sup>81</sup>

Later, in *Kyllo v. United States*,<sup>82</sup> the Supreme Court found that using a thermal imaging device from a public location to monitor the heat radiation from a private home was a Fourth Amendment search because the device allowed government agents to infer activity occurring inside the home. The Court recognized that heat radiation was not itself protected by the Fourth Amendment, but, Tokson argued, the Court found that *human exploitation* of these waves violated reasonable expectations of privacy and was thus a Fourth Amendment regulated activity.<sup>83</sup>

*Kyllo* and *Karo* both establish that the human/machine distinction matters for purposes of assessing whether a particular form of information acquisition constitutes a search. In *Karo*, the fact that a device is used is inadequate to establish a search, rather it is the fact that the device produces information consumed by a human. In *Kyllo*, the fact that a machine can accomplish something that a human could not has constitutional relevance. The human/machine distinction matters both for constraining *what is and is not a search* (based on human capacities) and also establishing *when a search takes place* (when the human sees it).

Tokson also discussed the human/machine distinction with respect to the third-party doctrine,<sup>84</sup> and here found that the distinction was without a constitutional difference under current jurisprudence. Under case law articulating or adhering to the third-party doctrine, the human/machine distinction is *unimportant*. As Tokson showed, under the Supreme Court's third-party doctrine line of cases—as of 2011—

---

78. Matthew Tokson, *Automation and the Fourth Amendment*, 96 IOWA L. REV. 581, 615 (2011).

79. 468 U.S. 705 (1984).

80. Tokson, *supra* note 78, at 615.

81. See *Carias v. Harrison*, No. 5:13-CT-3264-FL 2017 WL 1155749 (E.D.N.C. Mar. 27, 2017) (holding that privacy infringement from installation of a GPS tracker occurs not when the tracker is installed but when the information is disclosed to the United States Drug Enforcement Administration or another entity).

82. 533 U.S. 27 (2001).

83. Tokson, *supra* note 78, at 596.

84. The third-party doctrine renders information voluntarily turned over to third parties unprotected by the Fourth Amendment. The rationale for this doctrine is that parties who voluntarily turn information over to a third-party have no reasonable expectation of privacy due to their decision to render the data less private by turning it over.



“any information disclosed to a third party’s equipment is likely to be unprotected by the Fourth Amendment,” even where this information has merely been disclosed via an automated process in which a human would have played no role.<sup>85</sup> This suggests that surrender of information to a machine is already enough to establish a voluntary surrender of privacy inferred under the third-party doctrine. *Smith v. Maryland* even made this explicit; in the Court’s view it was the right outcome that the human/machine distinction *should make no difference*.<sup>86</sup> This line of cases has continued into our current era of digital information technology, such as in the much-discussed Ninth Circuit case of *United States v. Forrester*, which discussed automated Internet Protocol address scanning.<sup>87</sup>

Looking at cases about obtaining information only obtainable by machine (*Karo*, *Kyllo*) and cases about the third-party doctrine, Tokson identified that in the former case the presence of the machine made a difference while in the latter case it did not. But Tokson went further. Gathering original survey evidence as to the privacy judgments of laypeople depending on whether surveillance was conducted by a human or by a computer, Tokson showed that laypeople indeed found inspection by a human more privacy invasive than inspection by a computer.<sup>88</sup> Tokson argued that the human/machine (or, more specifically, human/computer) distinction should matter and that observation by a machine should not be adequate for inferring that someone had voluntarily surrendered their expectation of privacy in information.<sup>89</sup>

To sum up the state of affairs until recently, two lines of cases—(1) government informant transmissions/recordings of conversations and (2) the third-party doctrine—made no constitutional distinction between machine and human observation. A third line of cases (machine-enabled superhuman<sup>90</sup> surveillance, as in *Kyllo* and *Karo*) indicated that it was the *human use* of information produced by a machine, rather than observation by the machine itself, which made a constitutional difference.

That the more recent cases Tokson surveyed when writing in 2011 began to make the difference is not surprising if we contemplate how various new technologies—particularly digital technologies—were beginning to make their way into use, rendering machine surveillance far more pervasive and potent than had previously

---

85. Tokson, *supra* note 78, at 601.

86. 442 U.S. 735, 744–45 (1979) (“Petitioner concedes that, if he had placed his calls through an operator, he could claim no legitimate expectation of privacy. We are not inclined to hold that a different constitutional result is required because the telephone company has decided to automate.”).

87. 495 F.3d 1041 (9th Cir. 2007) (analyzing the computer surveillance of IP addresses as governed by substantially earlier precedents, like *Smith v. Maryland*).

88. Tokson points out that under a *Katz* analysis, which presumptively undergirds the third-party doctrine, there is not a subjective or objective basis for arguing that people have surrendered their privacy purely because an AI inspection of data has taken place. Tokson, *supra* note 78, at 615.

89. Tokson, *supra* note 78, at 615.

90. Of course, in these and other cases that are likely to come, the whole point is that— from a practical perspective due to cost and logistics—a human observer could not possibly offer such pervasive surveillance capabilities even if humans are legally entitled to make the attempt without a warrant.

been the case. Tokson and others likely had the sense that something would have to give, and give something did, with two watershed cases coming in 2012, *United States v. Jones*, and in 2018, *Carpenter v. United States*. In both cases the Court implicitly distinguished the surveillance that had actually taken place—assisted by sophisticated communications and surveillance technologies (in *Jones* a GPS device and in *Carpenter* cell site location information)—from what could have reasonably taken place with surveillance conducted by humans. In both cases the surveillance at issue could not realistically be performed by humans at all, and in both cases the Justices found that despite seemingly on point and contravening precedent that there was no privacy interest in one’s locations on a public way (*Jones*) or in third-party records (*Carpenter*)—such precedents would not control in the face of machine-enabled pervasive surveillance.


Tokson’s 2011 observation when writing about *Karo* and *Kyllo*—that machine observation alone could constitute a Fourth Amendment search when accompanied by human consumption of information—was carried forward in these more recent cases of *Jones* and *Carpenter*.<sup>91</sup> More recent Supreme Court jurisprudence has indicated an understanding that machine surveillance can be more invasive than human surveillance and that certain forms of surveillance may, surprisingly, be a Fourth Amendment search when conducted by a machine but not when conducted by a human (such as tracking a vehicle on a public road or tracking a person’s location in public places). This adds quite a new wrinkle indeed, where the human/machine distinction matters but goes *against the usual intuition* (that machine observation is less privacy invasive).

To sum up, a review of Fourth Amendment cases reveals the following inconsistent conclusions about machine versus human observation in three lines of Fourth Amendment cases.

---

91. Tokson, *supra* note 61, at 615.

**Table 1:** Supreme Court precedent is inconsistent in the constitutional importance of human versus machine observation.

Line of cases		Constitutional distinction, if any, between human and machine surveillance
Oldest  Newest	Machine-transmission/recording of conversations with government informants ( <i>On Lee, Lopez, White</i> )	None
	Third party collection of information ( <i>Smith v. Maryland</i> )	None
	Machine-enabled surveillance ( <i>Karo, Kyllo, Jones, Carpenter</i> )	<i>Less pervasive surveillance:</i> Constitutes a search only after human sees the information ( <i>Karo, Kyllo</i> )  <i>More pervasive surveillance:</i> Constitutes a search when conducted by a machine; likely does not constitute a search when conducted by a human (but this latter also likely not possible) ( <i>Jones, Carpenter</i> )

The more recent cases—even if consistent in making a difference and finding that machine surveillance can be more privacy invasive than human surveillance (*Jones, Carpenter*)—do not necessarily represent a widespread repudiation of the widely shared intuitive notion described *supra* that machines as such cannot invade privacy. Consider a transcript from a recent oral argument in the Colorado Supreme Court, in a case in which the parties readily recognized the relevance of recent Supreme Court jurisprudence, such as that in *Jones* or *Carpenter*. The case concerned the automated scanning of a search history, akin to a fact pattern of more pervasive surveillance. As discussed *supra* such cases have represented a point of departure for the Supreme Court to sometimes recognize that machines can be *more invasive* of privacy than humans, and yet this very recent case highlighted the longstanding intuition that there is something particularly invasive about the human gaze.

[I]n the briefing, we start hearing about the colonists . . . having their own [belongings] rummage[d] through. And that seems to be *a far cry* from [a Google search history]. I realize in a very conceptual sense that there's a similarity . . . . But we're not rummaging through. We're not even looking at a ledger where you might along the way sort of spy personal information that you might use in some other way as a

government agent. *You know, it's just the computer plucking these, you know, the X's and the O's are the ones in the end result, right?*<sup>92</sup>

Another recent example showing robust support for the intuition that machine surveillance is *less* rather than *more* invasive than equivalent human surveillance is new litigation challenging an arrest made on the basis of an identification with facial recognition technology in *Reid v. Bartholomew*.<sup>93</sup> In that case, an arrest is being challenged as constitutionally inadequate, with the plaintiff arguing that “an identification based solely on facial recognition technology does not provide probable cause for an arrest.”<sup>94</sup> The case is not on all fours with our present concern as to privacy because that case goes to the probable cause for the issuance of an arrest warrant, but the sentiments and high stakes are clearly similar to those in the Fourth Amendment context.

A skeptical reader may point out that even if the distinction seems to matter in some sense, it's not necessarily a problem of any kind that the distinction does not have consistent results throughout time or throughout related areas of case law. Fourth Amendment jurisprudence may be inconsistent on the question of machine versus human surveillance because this is not a lens the Supreme Court has made legally dispositive. In each case discussed here, the human versus machine distinction has been discussed but has never clearly been part of the holding. The inconsistency described here may therefore simply reflect the Supreme Court's deeply fact and context intensive analysis in Fourth Amendment cases. Nonetheless, given that the distinction has proven so important in other surveillance contexts (defense of national security surveillance, design of privacy settings in consumer markets) as shown *supra*, and given that the Supreme Court itself has sometimes recognized the distinction as important, it is unfortunate and even disappointing not to have greater clarity or consistency from the Court on this matter.

### *B. Communications Privacy Statutes*

Federal communications privacy law famously grew out of the nation's reaction—outrage—to the *Olmstead* decision, in which the Supreme Court found that wiretapping a phone line, when not trespassing upon a defendant's private property, did not constitute a Fourth Amendment search and therefore did not require a warrant. In passing the Wiretap Act six years later, Congress created a “warrant plus” requirement for interceptions of the contents of communications (more on this later in our discussion of the content/metadata distinction). This statutory regime, applicable in similar situations to the Fourth Amendment, has created a source of judicial discussion of the human/AI distinction: Courts must consider whether a

---

92. *Colorado Supreme Court Colorado v. Seymore, Case No. 23SA12*, AUDIO ARGUENDO at 04:21 (May 6, 2023), <https://podcasts.apple.com/us/podcast/audio-arguendo/id1067649051?i=1000612007059> [<https://perma.cc/H3LR-E6PK>] (quote extracted from automatically generated transcript accompanying podcast) (emphasis added).

93. Complaint for Damages, *Reid v. Bartholomew*, No. 1:23-cv-04035-JPB (N.D. Ga. Sept. 8, 2023) (transferred to E.D. La., No. 2:24-cv-02844).

94. *Id.* at \*3.

machine's interposition in the chain along which a communication was passed could constitute an interception of a communication, as covered by the Wiretap Act.

The few scholars that have written on the human/machine distinction and the Wiretap Act have come to conflicting conclusions. Writing in 2012, Bruce Boyden found that judges have interpreted the Wiretap Act such that machine observation is largely judged not to be an interception where identical human actions are an interception.<sup>95</sup> On the other hand, in 2014, Kevin Bankston and Amie Stepanovich found that any judicial distinction is minimal and rather that machine observation is fully adequate for statutory violations of the Wiretap Act. Here, I briefly review and compare the three cases that are the basis of both pieces of previous scholarship.

In *United States v. Turk*, officers stopped a car that matched a tip regarding transport of cocaine.<sup>96</sup> In the car was a box with two cassette tapes, to which the officers subsequently listened, discovering a recorded phone conversation through which they were able to locate another defendant and charge him. That defendant, Turk, argued that the act by the officers of listening to the recording constituted an impermissible "interception" under the Electronic Communications Privacy Act (ECPA),<sup>97</sup> which defines a procedure through which law enforcement can be authorized to intercept wire or oral communications.<sup>98</sup> The Fifth Circuit rejected this theory, finding that an ECPA violation necessarily entailed *the person who caused the recording to be made* and not parties who merely listened.

From the Fifth Circuit's decision, Boyden concluded that a human listener is necessary to an interception and that a machine can perpetrate an interception only when preserving communication for listening by the human party who instigated the recording.<sup>99</sup> Bankston and Stepanovich drew the opposite conclusion, emphasizing that it was the *act of recording*—a machine act—and not the act of human listening that would have constituted the offense.<sup>100</sup> *United States v. Turk* settles that mere listeners to a recording cannot be guilty of an interception but provides unclear guidance as to whether an interception occurs when a machine captures a communication or only occurs with some necessary subsequent action by a human perpetrator.

In *United States v. Rodriguez*,<sup>101</sup> a physical device was put in place to redirect calls arriving at a New York payphone to a New Jersey number. The difference of states mattered because the use of the physical device required a warrant in the jurisdiction in which the interception took place. The law enforcement officers had obtained a warrant in New Jersey but not in New York. If the physical device (rather than its human operators) performed an interception, then the interception took place

---

95. Boyden, *supra* note 63, at 669.

96. 526 F.2d 654 (5th Cir. 1976).

97. 18 U.S.C. §§ 2510–20 (2023).

98. Turk argued that the officers had "aurally acquired" the communications at the time of listening, a logic that would suggest that each incident of a human listening constituted a separate interception and thus a separate count of violating the statute. *Turk*, 526 F.2d at 657.

99. "If a machine can intercept a communication at all, the implication is that it occurs only when the machine preserves a communication for the purpose of later review by a . . . person." Boyden, *supra* note 63, at 669.

100. Bankston & Stepanovich, *supra* note 64.

101. 968 F.2d 130 (2d Cir. 1992).

in New York. The Second Circuit held that an interception had taken place in New York and that a warrant therefore should have been obtained in that district.<sup>102</sup> Bankston and Stepanovich took this as proof that a machine can perform an interception without any involvement by a human.<sup>103</sup> Boyden argued that the case was limited to a jurisdictional question rather than to a substantive violation of federal wiretapping laws and therefore did not provide insight as to whether a machine alone can intercept a communication.<sup>104</sup>

In *People v. Bialostok*,<sup>105</sup> a district court addressed whether a pen register—generally having the ability to record audio but not actually recording audio in the challenged instance—had intercepted communications.<sup>106</sup> The court concluded that the device acquired communications from the moment it was installed even if it was not physically able to record at the time (due to a lack of necessary accessories). From this, Bankston and Stepanovich found another example of a machine presence adequate to serve as an interception.<sup>107</sup> Boyden discounted *Bialostok*, this time in arguing that the *Bialostok* court misread *Smith v. Maryland*, a case the *Bialostok* court found to be controlling.<sup>108</sup>

In summary, if we assume cases could be taken on their merits and not taken to have misread precedent, then two cases establish that the presence of a machine itself is enough to constitute an interception (*Bialostok*) and the location of that interception (*Rodriguez*). What seems less clear is whether the person who instigated the recording must be later involved for a recording to be an interception (*Turk*). Reasonable legal scholars can disagree on whether computers can—on their own—violate federal communications privacy laws. To date, the Supreme Court has been silent on whether there may be a categorical distinction to draw in the federal wiretapping laws, as between machine and human actions, that could constitute an interception.

### C. The Privacy Torts

We next consider what, if anything, can be learned about the machine versus human privacy distinction in the context of the privacy torts. We consider two torts traditionally falling under the heading of invasion of privacy (intrusion upon seclusion and public disclosure of private facts).<sup>109</sup>

---

102. The court held that an interception occurred “when the contents of a wire communication are captured or redirected in any way.” *Id.* at 136.

103. Bankston & Stepanovich, *supra* note 64.

104. Boyden, *supra* note 63, at 717.

105. 610 N.E.2d 374 (N.Y. 1993).

106. A pen register is a physical device that records phone numbers dialed from a phone but that does not typically record other information, such as audio content of a call.

107. Bankston & Stepanovich, *supra* note 64.

108. Boyden, *supra* note 63, at 695.

109. I leave to the side discussions of defamation James Grimmelmman has posed the question of whether machines can commit defamation. See James Grimmelmman, The Defamation Machine (Aug. 19, 2024), <https://ssrn.com/abstract=4914458> [<https://perma.cc/S8PL-XJ6D>] (edited version of the 38th Annual Silha Lecture delivered Oct. 23, 2023). Further, arguably the most headline grabbing human/machine privacy litigation

### 1. Intrusion upon Seclusion

The tort of intrusion upon seclusion describes situations of illicit or inappropriate information acquisition. Courts have found that machine observation can constitute the intrusion necessary to establish such a tort.

One who intentionally intrudes, physically or otherwise, upon the solitude or seclusion of another or his private affairs or concerns, is subject to liability to the other for invasion of his privacy, if the intrusion would be highly offensive to a reasonable person.<sup>110</sup>

It is not difficult to identify cases of alleged intrusion upon seclusion in which it is not the bare discovery of a fact or listening by a human that has triggered circumstances for a lawsuit. Rather it is straightforward to find cases *in which the additional, undisclosed, use of a machine recording device* has triggered circumstances to bring suit.

Consider the case of *Safari Club International v. Rudolph*.<sup>111</sup> Two people met for lunch. They were formerly friends but formally in an adversarial relationship due to impending employment-related litigation. The friend who had initiated the lunch invitation surreptitiously recorded his interlocutor and was later sued for doing so under an intrusion upon seclusion claim. As the court described:

[Plaintiff] avers [defendant's] *surreptitious recording* of their lunchtime discussion intruded unlawfully into his private conversation. [Plaintiff] maintains the occurrence was objectively offensive because [defendant] used friendship to lure him to lunch, then secretly recorded their conversation and shared it widely with members of the public . . . . Though the question is close, we think plaintiffs' proffered evidence, taken as whole, could support a reasonable jury finding that [defendant's] actions constituted a "highly offensive" intrusion into [plaintiff's] privacy.<sup>112</sup>

Note the logic, and how this logic appears to differ from that of *Lee* and *Lopez*. It was, at least in part, the use of the recording device that rendered the conduct offensive. The *Safari Club* decision went on to give other examples of cases in which surreptitious recording had been problematic. The court found a series of cases that supported the understanding that the surreptitious use of a recording device—machine observation—could suffice both to satisfy the intrusion element of intrusion upon seclusion *and also* the highly offensive element. The court noted that the State

---

ever to take place invokes defamation, and so the human/machine distinction is of skyrocketing importance in part due to defamation theories. *See, e.g.*, Complaint, *Walters v. OpenAI, L.L.C.*, No. 1:23-cv-03122-MLB (N.D. Ga. July 14, 2023). This emerging line of scholarship and case law so far looks primarily at whether machines can commit defamation as speakers, not whether their gaze is legally adequate, and there the question is related to but distinct from that addressed in this work.

110. RESTATEMENT (SECOND) OF TORTS, § 652B (AM. L. INST. 1977).

111. 862 F.3d 1113 (9th Cir. 2017).

112. *Id.* at 1127.

lacked a bright line rule to govern all uses of machine observation but nonetheless found that a range of factual circumstances in using machine observation could rise to the level of intrusion upon seclusion.

Consumer privacy litigation also includes a theory of intrusion upon seclusion predicated upon machine activity alone. Consider recent litigation brought against both Google<sup>113</sup> and OpenAI.<sup>114</sup> In the Google litigation, the plaintiffs allege that the training and normal operation of these AI systems constitute intrusion upon seclusion. In the OpenAI case, the plaintiffs allege that the OpenAI system (a machine) is intruding upon their communications with websites (machines)<sup>115</sup> upon which they have a reasonable expectation of privacy. That's right, it's turtles all the way down; it may be that human users have an expectation of privacy in their communications with a machine that may be infringed upon by yet another machine.

It is not just the tort of intrusion upon seclusion that understands machines to intrude upon seclusion. Other intrusive actions implicated by expansive notions of privacy encompassing the right to be let alone are also widely recognized in the context of machine use. For example, the CAN-SPAM Act and the Telephone Consumer Protection Act both prohibit precisely the kinds of intrusions that occur as a result of the use of machines disturbing humans through unsolicited communications. The rapid development of these laws and their relatively stringent and consistent enforcement over time suggests that, in addition to the direct evidence provided by intrusion upon seclusion case law, society more widely recognizes the intrusive nature of a machine interlocutor vis-à-vis the human right to be left alone.

## 2. Public Disclosure of Private Facts

The tort of public disclosure recognizes circumstances in which private information has been inappropriately shared or conveyed beyond limited circumstances.

One who gives publicity to a matter concerning the private life of another is subject to liability to the other for invasion of his privacy, if the matter

---

113. Class Action Complaint at 73, *In re Google Generative AI Copyright Litig.*, No. 3:23-cv-03440 (N.D. Cal. July 11, 2023) (“Defendants intentionally intruded on and into Plaintiffs’ and Class Members’ solitude, seclusion, or private affairs by constructing a system which collects, stores, and uses Personal Information of millions of individuals (both users/nonusers of Google products). This information includes personal, medical, financial information, and copyrighted materials.”).

114. Complaint at 108, *A.T. v. OpenAI LP*, No. 3:23-cv-04557 (N.D. Cal. Sept. 5, 2023) (“In carrying out their scheme to track and intercept Plaintiffs’ and Class members’ communications and other private data, Defendants intentionally intruded upon the Plaintiffs’ and Class members’ solitude or seclusion in that Defendants effectively placed themselves in the middle of conversations to which they were not authorized parties.”).

115. Of course, one could rejoin that the communication is with the firm—the legal person—with which plaintiffs intend to communicate. Perhaps that could legally be constructed as the case, but it is not the language plaintiffs use. Rather plaintiffs refer to “the [w]ebsites with which they were communicating.” *Id.*



publicized is of a kind that (a) would be highly offensive to a reasonable person, and (b) is not of legitimate concern to the public.<sup>116</sup>

Courts have not found machine observation to be clearly adequate to establish this tort, in contrast to their adequacy for purposes of establishing intrusion upon seclusion. Consider for example a case of consumer privacy litigation against Facebook that challenged Facebook's practices of making information about an app user's contacts (or "friends") available to third-party "apps," even in contradiction to Facebook's own representations of their privacy practices. Plaintiffs brought suit challenging these practices on a number of theories, including the tort of public disclosure of private facts. Rejecting Facebook's motion to dismiss vis-à-vis claims that (1) the firm had permitted disclosure to whitelisted third parties of information (2) contrary to its privacy policy, (3) that the firm engaged in data reciprocity with business partners (also undisclosed in its privacy policy), and (4) that Facebook had failed to supervise or restrict third parties' use of the data, Judge Chhabria of the Northern District of California found that Facebook had not offered adequate reason to dismiss plaintiff's public disclosure of private facts claims based on Facebook's own arguments. The judge, however, did suggest that Facebook had missed the opportunity to make a more effective argument. As the judge saw it, a better argument than Facebook's arguments, which leaned heavily on consent and also on the data at issue not being private, would have been an argument that emphasized the human/machine distinction.

Perhaps Facebook could have made a better argument . . . . [T]here's a difference between publicizing your sensitive information for actual human beings to scrutinize (like, in a newspaper) and allowing your information to be added to the vast sea of "big data" that computers rather than humans analyze for the purpose of sending targeted advertising on behalf of companies. Perhaps there is an argument that the former is the "public disclosure" of information within the meaning of California law while the latter is not.<sup>117</sup>

Since Facebook did not make such an argument, we do not have a substantive result to justify or disprove the judge's speculation. Judge Chhabria's initiative in suggesting this is strong evidence that he or other judges would be unlikely to view disclosure among machines as adequate for the disclosure tort.

Prominent Supreme Court precedent also indirectly provides evidence as to the inadequacy of a public composed of machines for purposes of the public disclosure of private facts tort (and the related tort of defamation). Consider the landmark Article III standing case of *TransUnion v. Ramirez*.<sup>118</sup> The Supreme Court considered whether plaintiffs alleging violations of the Fair Credit Reporting Act could establish standing where the allegation of harm concerned false reports included in plaintiffs' credit reports that stated the plaintiffs were on the Department

---

116. RESTATEMENT (SECOND) OF TORTS § 652D (AM. L. INST. 1977).

117. *In re Facebook, Inc., Consumer Priv. User Profile Litig.*, 402 F. Supp. 3d 767, 796 (N.D. Cal. 2019).

118. 594 U.S. 413 (2021).

of the Treasury's Office of Foreign Assets Control (OFAC) watch list for terrorism and other serious crimes and had been distributed to those requesting a credit report. The plaintiffs were then divided into two groups: one whose credit reports had been distributed to third parties with the false report and one whose credit reports had not been distributed while containing the false report.

The Supreme Court found that those whose credit reports had been sent to third parties could establish standing.<sup>119</sup> In doing so, the Court built on new guidance from another recent precedent, *Spokeo v. Robins*.<sup>120</sup> (indicating that those who sought standing for violation of statutory rights had to show that they had suffered a concrete and particularized injury), to announce that the adequacy of an intangible harm necessary to establish Article III standing would require an inquiry into analogs in history and tradition, as evidenced by connections between a posited harm and legal theories long recognized at common law.<sup>121</sup> In *TransUnion*, those plaintiffs whose credit reports were distributed to third parties with false information were found to have standing, as justified under the new analysis required by *TransUnion* via analogies between the harm those plaintiffs suffered and the tort of defamation. But the plaintiffs whose credit reports had not been distributed while containing the false report did not have standing.<sup>122</sup> The *TransUnion* Court did not explicitly single out the importance of a human inspection of the defamatory-like content, although this may have been in part because the named plaintiff did in fact report adverse consequences in a sequence of events that did entail human inspection of the credit report.<sup>123</sup>

---

119. *Id.* at 432 (“The 1,853 class members therefore suffered a harm with a ‘close relationship’ to the harm associated with the tort of defamation.”)

120. *Spokeo, Inc. v. Robins*, 578 U.S. 330 (2016).

121. *TransUnion v. Ramirez*, 594 U.S. 413, 424 (2021) (stating the relevant inquiry is whether “plaintiffs have identified a close historical or common-law analogue for their asserted injury”)

122. *Id.* at 427–8 (“Even if Congress affords both hypothetical plaintiffs a cause of action (with statutory damages available) to sue over the defendant's legal violation, Article III standing doctrine sharply distinguishes between those two scenarios. The first lawsuit may of course proceed in federal court because the plaintiff has suffered concrete harm to her property. But the second lawsuit may not proceed because that plaintiff has not suffered any physical, monetary, or cognizable intangible harm traditionally recognized as providing a basis for a lawsuit in American courts. An uninjured plaintiff who sues in those circumstances is, by definition, not seeking to remedy any harm to herself but instead is merely seeking to ensure a defendant's ‘compliance with regulatory law’ (and, of course, to obtain some money via the statutory damages).”). The *TransUnion* Court did not explicitly engage with the likelihood that transfer of the credit reports to third parties entailed inspection of those reports by human third parties. It seems plausible that in many such cases no human did in fact inspect the reports and defamation-like content therein. See Cecilia Petit & Stephen Gregg, *Automation: Powering More Efficient and Effective Credit Decisions*, DUN & BRADSTREET, (Oct. 7, 2020) <https://www.dnb.com/perspectives/finance-credit-risk/credit-decision-automation.html> [<https://perma.cc/A2F4-APUD>].

123. *TransUnion*, 594 U.S. at 420. (“A Nissan salesman told Ramirez that Nissan would not sell the car to him because his name was on a ‘terrorist list.’ . . . Ramirez’s wife had to purchase the car in her own name.”).

Nonetheless, later cases making use of *TransUnion* have called out this distinction. For example, in a recent data breach case where plaintiffs brought claims, *inter alia*, for public disclosure of private facts, the court found that they could not establish standing under the requisite *TransUnion* analysis. In *In re Practicefirst Data Breach Litigation*, the court found that there was no injury in fact in a data breach, explicitly rejecting the plaintiffs' contention that they were the victims of the tort of public disclosure of private facts.<sup>124</sup> The plaintiffs consequently failed both to establish the injury required for the tort but also the (possibly less searching) injury in fact required for standing.

The complaint does not allege that Practicefirst directly disclosed plaintiffs' confidential information to the public at large. Indeed, the breach involved the PHI/PII of over 1.2 million people and there are no allegations in the complaint to indicate that plaintiffs' particular private information was ever specifically viewed by any one person, let alone that it was disclosed 'publicly.'<sup>125</sup>

Thus, in contrast to *TransUnion*, where the Supreme Court did not go further to inquire how and by whom (or what) the information in the problematic credit reports had been seen, other courts—for example looking to data breaches—have looked at exactly *who* or *what* has seen or processed the data, finding a lack of standing where there has been no clear disclosure to “any one person.”<sup>126</sup>

#### D. Consumer Protection

Consumer protection enforcement sometimes mirrors market trends and reactions, and this was somewhat the case with regard to the 2019 wave of consumer privacy scandals at large tech firms in which a number of voice-assistant products were found to transcribe or otherwise process audio recordings of customer conversations with the assistance of employees as well as automated methods.<sup>127</sup> As was described *supra*, the resulting backlash prompted several firms to update their privacy policies to disclose these practices and also prompted the firms to enable customers to opt out of human audio processing. Perhaps related to this news coverage, the FTC began its own investigation of Amazon's Alexa voice assistant, and in 2023 the FTC filed a consent decree with Amazon. The FTC alleged that

---

124. *In re Practicefirst Data Breach Litigation*, No. 1:21-CV-00790(JLS/MJR), 2022 WL 354544 (W.D.N.Y. Feb. 2, 2022). I thank Mary Noh for this example.

125. *Id.* at \*7.

126. Note that this disclosure—seemingly required to humans not just to machines—is not as high a threshold as is other jurisprudence that has emerged, in some cases requiring documented misuse of data or something similar in order for data breach plaintiffs to establish that any mitigation efforts they have undertaken have been reasonable. *See e.g.*, *Tsao v. Captiva MVP Rest. Partners, LLC* 986 F.3d 1332 (11th Cir. 2021) (requiring plaintiffs to provide evidence of misuse, implicitly rejecting the notion that proof of a human gaze was adequate).

127. *See supra* notes 16–21 and accompanying text.

Amazon's data practices violated both the FTC Act<sup>128</sup> and the Children's Online Privacy Protection Act (COPPA).<sup>129</sup>

The FTC complaint identified both human and machine-related data practices as problems. In the complaint, the FTC cited as evidence of poor privacy protection the fact that over 30,000 Amazon employees had access to Alexa users' voice recordings, over half of whom "lacked any business need" for such access.<sup>130</sup> The FTC, however, also evinced concern about algorithmic access to voice recordings. For example, the FTC expressed particular concern about the use and retention of children's voices as "a valuable data bank for training the Alexa to understand children."<sup>131</sup> The FTC interpretation of its privacy protection responsibilities implicates both human and machine access to information. Both humans and machines can trigger privacy incursions addressed under the FTC Act.

Further evidence of FTC attention to both human and machine access to information comes in the same enforcement action. Another detail the FTC called out in its complaint was Amazon's practice of deleting only voice recordings but not written transcripts of conversations when users requested the deletion of audio recordings.<sup>132</sup> The FTC argued that this practice was misleading because the interface could reasonably lead users to believe that both the audio and written transcripts had been deleted. The FTC called out the possible use of such written transcripts only with respect to machine use of these transcripts. While, of course, human employees could presumably read the transcripts too, the language of the complaint emphasized as problematic algorithmic access of the written transcripts.<sup>133</sup>

In the Amazon Alexa enforcement action, the FTC separately called out human access to voice recordings and machine access to voice recordings and written transcripts. If humans and machines were one and the same (members of the Amazon corporation), it would not be necessary or interesting to make these finer distinctions. This is in stark contrast to the Fourth Amendment jurisprudence discussed earlier (*On Lee, Lopez*) in which the Supreme Court explicitly rejected machines as constituting a privacy infringement at all, let alone one that could layer atop human privacy infringements.

A review of other recent FTC enforcement action shows a consistent logic by the FTC that either humans or machines can invade privacy. The FTC enforcement action against Ring shows an approach in which the privacy incursions of interest

---

128. 15 U.S.C. § 45(5)(a).

129. Complaint for Permanent Injunction, Civil Penalties, and Other Relief, United States v. Amazon.com, Inc., No. 2:23-cv-00811 (W.D. Wash. May 31, 2023).

130. *Id.* at 8.

131. *Id.* at 6.

132. *Id.* at 8.

133. *Id.* ("Those transcripts remained available for Amazon's benefit and use for product improvement. . . . Amazon used children's recordings—both audio files and transcripts—for purposes such as refining Alexa's voice recognition and natural language processing capabilities."). While the language does not clearly rule out the use of algorithms, the use of algorithms seems far more likely to be what was referred to in understanding how such recordings or transcripts could be used to improve the product. There is also an explicit reference to two algorithm use cases: voice recognition and natural language processing.

were perpetrated nearly exclusively by humans.<sup>134</sup> Even the security lapses that were cited in the complaint related to creating the risk of human access and to concrete scenarios demonstrating the realization of that risk, circumstances in which hackers gained access to video feeds. The complaint particularly emphasized privacy incursions where hackers gained access to video feeds in intimate spaces. But the FTC has also brought enforcement actions where the privacy incursion is purely by machine. For example, the FTC's enforcement action against BetterHelp alleged that the firm failed to obtain consent prior to disclosing health information to online platforms.<sup>135</sup> There was no allegation of human involvement or human consumption of this information. Likewise, in the enforcement action against Everalbum, there was similarly no allegation of privacy incursions by humans; rather, the offending events related to algorithmic use of data (facial recognition technology applied to user uploaded content).<sup>136</sup> Thus a review of some recent FTC enforcement shows that (1) human observation is adequate to establish a privacy harm, (2) machine observation is adequate to establish a privacy violation, and (3) when both forms of observation coincide, they can be understood as layers of distinct harm rather than redundant.<sup>137</sup>

#### *E. Summing Up the Legal Distinctions and Nondistinctions*

An overview of the findings so far is summarized in the table below. These findings show that machine versus human observation is treated inconsistently both within certain domains and also across domains. Given that these four privacy domains all evince an interest in the reasonable expectations of ordinary people,<sup>138</sup> it is surprising that they have come to such a wide variety of conclusions.

---

134. Complaint for Permanent Injunction and Other Relief, *FTC v. Ring LLC*, No. 1:23-cv-01549 (D.D.C. May 31, 2023).

135. Complaint, *In re BetterHelp, Inc.*, No. C-4796 (July 7, 2023).

136. Complaint, *In re Everalbum*, No. C-4743 (May 6, 2021).

137. Those well versed in the privacy law and conceptual privacy literature will of course be aware that different allegations can relate to different privacy interests. For example, one might argue that the privacy interest protected against facial recognition technology could relate to an economic right not to have one's likeness appropriated without consent or a dignitary right not to have one's unique appearance reduced to a vector representation. On the other hand, with regard to the human incursions described in *Ring*, one could argue for special protections for one's home or a right to be free from intrusion, that is the classic right to be let alone. In looking across broad areas of privacy law, this work is necessarily looking at the human versus machine distinction in privacy across many understandings of privacy. Importantly, I look to those understandings of privacy that have some legal recognition and recourse, but this still leaves quite a wide tent implementation of privacy.

138. See *supra* note 11.

**Table 2:** Overview of inconsistencies within and across the four privacy domains.

Domain	Is machine observation itself invasive of privacy?
Fourth Amendment law	<p>No in the case of government informant cases (undisclosed/additional machine observation changes nothing in privacy calculus as compared to mere fact of human observation)</p> <p>No in the case of technology and observations of public movement (human use of that machine observation is needed for invasion of privacy per <i>Karo</i> and <i>Kyllo</i>)</p> <p>Yes in the case of the third party doctrine (machine observation defeats expectations of privacy with regard to humans in the same way as human observation)</p>
Communications privacy law	Unclear / debatable
Privacy torts	<p>Yes in the case of intrusion upon seclusion</p> <p>No in the case of public disclosure of private facts</p>
Consumer protection law	Yes

#### F. Empirical Scholarship

A brief review of cases from the four privacy domains has revealed ambiguity as to the privacy implications of machine observation (communications privacy laws), clear inconsistencies in the inferred privacy implications of machine observation (Fourth Amendment and the privacy torts), and clear consistency in the legal weight given to both human and machine observation (consumer privacy law as reflected in FTC enforcement). The four privacy domains are inconsistent with one another, and sometimes inconsistent internally. At this point, one plausible critique of the exercise is that we are looking at law on the wrong terms—insisting on making a distinction that is legally unimportant and therefore likely to generate different outcomes not because the law is poorly designed or poorly developed but rather because the question itself is poorly designed. As discussed *supra* one rebuttal to this critique is that the four privacy domains all integrate expectations of privacy—often through the *Katzian* incantation of reasonable expectations of privacy—into legal analysis. But that's not the only reason the distinction is justified.

To further address any such concerns—that the importance of the machine versus human observation distinction is a nondistinction upon which I foolishly insist—I next turn to some empirical social science results that likewise establish a behaviorally grounded understanding of the importance of the human/machine distinction when it comes to privacy. These results show that the distinction is one that also matters to ordinary people—the very people upon whose reasonable

expectations of privacy so much turns.<sup>139</sup> Here I briefly discuss a few examples of such findings to give the reader a flavor of what is understood about how humans distinguish between human and machine observation.

The human-computer interaction literature provides one channel to understand how people judge humans versus AIs. This literature has long focused on elicitation of information, rather than observation as such, likely reflecting the technological realities that for quite some time it has been possible for computers to directly elicit information from humans (a specific form of observation), but only more recently has it become practicable for computers to observe humans and generate information about those observations. Results from these earlier experiments on the role of AIs in disclosure exercises showed that often people feel more comfortable sharing sensitive information with a computer than with a person, showing (1) that humans distinguish human and machine observation and (2) that at least in some cases humans find machine observation less privacy intrusive than they do human observation.

Consider a 2014 study in which Lucas et al. measure information disclosure behaviors with respect to a machine or human interviewer.<sup>140</sup> Participants reported lower subjective evaluation fears and disclosed more information when told they were interacting with a machine rather than a human. This reflects the common scenario one would expect, and is consistent with the widespread intuitions discussed *supra*, that machine observation is often experienced as less privacy invasive than human observation.<sup>141</sup> However, this trend observed by Lucas et al. (and also more recently by others)<sup>142</sup> is not absolute; humans don't always disclose more information to machines. In a 2016 study, Pickard et al. found that people prefer to answer questions on more sensitive topics with a machine interviewer, but that people disclose *more to humans* when responding to questions on less sensitive topics.<sup>143</sup> Of course, the reason for disclosing more need not be about privacy but

---

139. The human/machine distinction is so consistently important to social science research participants and consequently interesting to social scientists that Cass Sunstein and Jared Gaffe have even put together a taxonomy of the scholarship studying the documented phenomena associated with differential treatment for humans versus algorithms. *See* Cass R. Sunstein & Jared H. Gaffe, *An Anatomy of Algorithm Aversion*, 26 COLUM. SCI. & TECH. L. REV. 290 (2025).

140. Gale M. Lucas, Jonathan Gratch, Aisha King & Louis-Philippe Morency, *It's Only a Computer: Virtual Humans Increase Willingness to Disclose*, 37 COMPUTS. HUM. BEHAV. 94 (2014).

141. Note that this claim is only that machine observation is *intuited* or *experienced* as less privacy invasive. This claim does not contravene that there can be privacy related risks that are greater due to observation by computer than by person. Further, this claim does not contravene that ordinary people could be aware of the specific and increased risks created by computer observation.

142. *See e.g.*, Jasper Holthöwer & Jenny van Doorn, *Robots Do Not Judge: Service Robots Can Alleviate Embarrassment in Service Encounters*, 51 J. ACAD. MKTG. SCI. 767 (2022); Valentina Pitardi et al., *Service Robots, Agency and Embarrassing Service Encounters*, 33 J. SERV. MGMT. 389 (2021).

143. Matthew D. Pickard, Catherine A. Roster & Yixing Chen, *Revealing Sensitive Information in Personal Interviews: Is Self-Disclosure Easier with Humans or Avatars and Under What Conditions?*, 65 COMPUTS. HUM. BEHAV. 23 (2016).

could relate to other social, contextual, or instrumental factors. For example, if a machine asks me how I am doing, I would understand this question to be even more gratuitous than it is widely understood to be (in American culture) when coming from a human.<sup>144</sup> Thus, behavioral results show that, at least sometimes, behavior suggests that machine observers are perceived to be less privacy invasive but that this is not the only factor at play in the relative degree of disclosure to a human versus machine interlocutor.

As the observational capacities of computers have evolved, the topic of surveillance acceptance and specifically AI surveillance acceptance—rather than information disclosure—has emerged as a related but distinct question. In this new genre of studies consider for example a 2021 experiment by Raveendhran and Fast,<sup>145</sup> who studied acceptance of workplace surveillance. They found that potential job applicants were more willing to apply to jobs when behavioral surveillance would be performed by AI rather than by humans, a finding consistent in both a vignette experiment and a laboratory study. In a series of follow-on studies to understand the reason for this preference, Raveendhran and Fast found that the difference in preference was driven by concerns about potential negative judgments by humans and separately about autonomy, each of which research participants expected to be more impacted by human rather than AI observation. Participants expected that they would feel freer to make choices when facing AI rather than human observation of their work. The authors' findings as to a human/machine distinction with regard to observation were consistent with the earlier literature on information disclosure: Computers are generally perceived or otherwise experienced as less privacy invasive, or are otherwise preferred in privacy-relevant situations (that is, situations where one would prefer not to be observed by everyone) for reasons related or unrelated to privacy.

These findings suggest that when ordinary people are surveilled, we expect and avoid judgment from fellow humans, but we do not clearly have the same response with respect to machines.<sup>146</sup> Such experimental evidence tends to affirm the speculation found in doctrinal works, like those discussed earlier by Solove and Lev-Aretz,<sup>147</sup> that certain concerns associated with a broad notion of privacy are concerns *specific to how human observation* will distort behavior rather than about information flows as such. Prior doctrinal work and this Article's insistence on the importance of the human/machine distinction are consistent with a host of social

---

144. u/Drecher\_91, *Why Do Americans Ask "How Are You Doing Today" Without Actually Meaning It?*, Reddit, (Oct. 16, 2018), [https://www.reddit.com/r/AskAnAmerican/comments/9ov3qy/why\\_do\\_americans\\_ask\\_how\\_are\\_you\\_doing\\_today/](https://www.reddit.com/r/AskAnAmerican/comments/9ov3qy/why_do_americans_ask_how_are_you_doing_today/) [https://perma.cc/4PAX-R9UA].

145. Roshni Raveendhran & Nathanael J. Fast, *Humans Judge, Algorithms Nudge: The Psychology of Behavior Tracking Acceptance*, 164 *ORG. BEHAV. HUM. DECISION PROCESSES* 11 (2021).

146. These results suggest that people do not factor in "downstream" judgment from humans, such as judgment from someone looking at the data after it has been collected and initially processed by an AI. This has been a consistent result—that data subjects do not evince long term concerns about the possibility of later human viewing but tend to focus on the point of data's "first impression" by an AI.

147. *See supra* notes 58–59 and accompanying text.



science findings that the distinction is indeed one with a difference in human behavior. Machine versus human observation matters to humans, and therefore there is good reason that it should matter to law, and, further, that law should move towards offering a consistent and predictable account as to the privacy import of the machine gaze.

### III. CONTENT VERSUS METADATA

Let's now turn our attention to the second distinction of interest in this Article, that of content versus metadata. In the Fourth Amendment context, this distinction has sometimes been articulated as the difference between contents and envelope data,<sup>148</sup> and in communications privacy law as the distinction between content and record data.<sup>149</sup> The importance, if any, of this distinction is not so formally apparent as an organizing theme in either the privacy torts or in consumer privacy law, but as we shall see below these areas of law nonetheless sometimes give importance to the distinction too.

Before we embark upon a review of relevant law, it is worth noting that in law and in the statistical sciences alike, it is well acknowledged that the distinguishing of content and metadata is a conceptually and empirically problematic exercise.<sup>150</sup> Consider once more *Smith v. Maryland*, discussed previously as to the third-party doctrine. *Smith* rejected the possibility that there could be different outcomes regarding the expectation of privacy depending on whether a telephone switch was automated or operated by a human. While implicitly rejecting any privacy importance of the human/machine distinction, the case provides some helpful if non-controlling language from Justice Stewart's dissent when it comes to understanding what was appreciated, if anything, about the content/metadata distinction. Justice Stewart authored a dissent in that case not premised on the human versus machine distinction but rather on the majority's assumption that information about phone numbers dialed was not covered by a reasonable expectation of privacy:

The numbers dialed from a private telephone—although certainly more prosaic than the conversation itself—are not without “content.” Most private telephone subscribers may have their own numbers listed in a publicly distributed directory, but I doubt there are any who would be happy to have broadcast to the world a list of the local or long distance

---

148. Matthew J. Tokson, *The Content/Envelope Distinction in Internet Law*, 50 WM. & MARY L. REV. 2105 (2009).

149. The Stored Communications Act distinguishes the governance regime applied to covered entities depending on whether it is contents or record data that are disclosed. See 18 U.S.C. § 2703.

150. For a reasonable summing up of examples of the distinction between content and metadata and the fluidity or lack of robustness of that distinction, I refer the reader to a Perplexity.ai output. Perplexity, *Statistical Sciences Content Versus Metadata Distinction is Illusory*, PERPLEXITY (Mar. 19, 2025), [https://www.perplexity.ai/search/statistical-sciences-content-v-F\\_abaUmQQMOQsLXCn2nOTg](https://www.perplexity.ai/search/statistical-sciences-content-v-F_abaUmQQMOQsLXCn2nOTg) [https://perma.cc/YNY5-8TEW].

(AI generated output that I have reviewed and endorse). A particularly excellent example output by Perplexity was the following: timestamp on a data point is metadata when describing when that data was collected but becomes content when used in time series analysis.

numbers they have called. This is not because such a list might in some sense be incriminating, but because it easily could reveal the identities of the persons and the places called, and thus reveal the most intimate details of a person's life.<sup>151</sup>

Unsurprisingly, Justice Stewart's common-sense intuition in 1979 has subsequently been substantiated with empirical research. Consider a 2016 study authored by lawyer and computer scientist Jonathan Mayer and his team.<sup>152</sup> Studying a population of over 800 volunteer research participants and looking to the telephone numbers of outgoing and incoming calls on the participants' device logs, the authors demonstrated the ease of determining the geographic location of the volunteers, the structure of their social networks, and their romantic relationship status. Further, Mayer et al. showed examples of how various sensitive or constitutionally protected behaviors could be identified, such as calls to religious organizations, health services, firearms sales, and political campaigns. Justice Stewart anticipated these findings decades earlier, but Mayer et al. empirically demonstrated how informative telephone metadata could be.<sup>153</sup>

Implicitly in the discussion above, I have made an assumption about what constitutes content and what constitutes metadata. I have assumed that inferences that a call is about a religious, political, or health topic in some sense hews closer to content than to metadata. Beyond the realm of communications privacy law, it is of course not entirely clear what "content" is, but inferring with a reasonable degree of certainty would appear reasonably to fall within the Wiretap Act's understanding of the contents of a communication:

"[C]ontents", when used with respect to any wire, oral, or electronic communication, includes any information concerning the substance, purport, or meaning of that communication"<sup>154</sup>

The illusory nature of the content/metadata distinction likely explains two phenomena. First, in the area of Fourth Amendment law, the content/metadata (or content/envelope) distinction is one that is understood to be increasingly tenuous. Courts are increasingly recognizing that much information that would traditionally be understood as metadata, or record information, is highly informative in ways that intuitively call out for privacy protection. This may be why, recently, Fourth Amendment jurisprudence is shifting away from a stance of offering no protection

---

151. *Smith v. Maryland*, 442 U.S. 735, 748 (1979) (Stewart, J., dissenting).

152. Jonathan Mayer, Patrick Mutchler & John C. Mitchell, *Evaluating the Privacy Properties of Telephone Metadata*, 113 PROC. NAT'L ACAD. SCI. 5536 (2016) (conducting a study of telephone metadata with more than 800 participants, which revealed that the phone record metadata from these participants was sufficient to identify in a majority of cases the romantic relationship status of a participant, their geographic location, and in some cases, their religious affiliations or medical service providers).

153. In principle such research would have been possible decades earlier from landline telephone bills, but of course modern mobile devices and the full informational infrastructure of the internet have made this research both more informative and also easier to carry out.

154. 18 U.S.C. § 2510(8).

to metadata in the hands of third parties toward a stance of looking at the privacy import of the metadata, as described *infra*.

Second, the modern litigation of consumer privacy in the modern, surveillance-driven digital ecosystem has traditionally not emphasized a distinction between content versus metadata, perhaps itself an implicit admission of the misleading nature of such a distinction. This may be in part because, particularly with respect to consumers, the distinction between information provided *by* consumers and the information observed *about* consumers is not what is concerning about firms' privacy practices. Either content or metadata may be used to influence or "profile" consumers, a major concern intrinsic to reactions to data-driven consumer markets.<sup>155</sup> Further, to the extent distinctions are made, the distinctions are made as between sensitive data or information versus other information. But sensitive information is unrelated to what a consumer communicates and is instead related to attributes about a consumer that are judged to be particularly worthy of legal protection.

So, it may be that law is slowly and implicitly moving away from any previous endorsement of a content/metadata distinction. And yet, the distinction—though illusory—remains important to the extent it matters to ordinary people and explains behaviors in the consumer-facing digital markets.<sup>156</sup> Further, the distinction matters because—whatever one makes of it—the distinction can be to some extent operationalized with technology. That is, as machines become increasingly credible and reliable observers, they can, at least in theory, offer increasingly robust and credible promises about the nature and extent of information that will be collected. A human cannot promise to see only metadata, but a machine can.<sup>157</sup> Therefore, the privacy import of different forms of observation becomes more important when different forms of observation can, in fact, be conducted in a credible manner.

Having conceded the fuzziness and problematic nature of the content/metadata distinction but nonetheless having protested its importance on the basis of the popular imagination and the behavior of digital markets, I now propose a loose

---

155. Consider, for example, the much discussed and highly controversial Cambridge Analytica scandal. Nicholas Confessore, *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far*, N.Y. TIMES (Apr. 4, 2018), <https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html> [<https://perma.cc/6VQP-LYD2>] (making repeated references to "profiles" developed for purposes of political campaigning).

156. See, e.g., *supra* notes 47–49 and accompanying text (discussing the WhatsApp example).

157. For example, flight recorders continue to track only audio and not visual data from the cockpit and record only the most recent 30 minutes of data, providing a built-in and technically guaranteed privacy protection to personnel in the cockpits and to the firm operating the plane. See Claire Suddath, *Black Boxes*, TIME (July 2, 2009), <https://content.time.com/time/subscriber/article/0,33009,1909619,00.html> [<https://perma.cc/SW8A-GE6L>]. But see ACLU OF CALIF., METADATA PIECING TOGETHER A PRIVACY SOLUTION 2 (2014), [https://www.aclunc.org/sites/default/files/Metadata%20report%20FINAL%202%2021%2014%20cover%20+%20inside%20for%20web%20\(3\).pdf](https://www.aclunc.org/sites/default/files/Metadata%20report%20FINAL%202%2021%2014%20cover%20+%20inside%20for%20web%20(3).pdf) [<https://perma.cc/8ZDP-HKNW>] (arguing that advancing technology makes the content/metadata distinction less rather than more meaningful).

definition that should suffice for purposes of this discussion and experiment. In the case of legal regimes that include some notion of “content”—typically those legal regimes that concern explicitly communications—“metadata” will be taken to be everything that is not “content.” In categories of law for which content is not a clearly relevant or formally defined category, metadata will concern that which is observed *about* a surveillance subject whereas content will be what is directly elicited *from* a surveillance subject. For example, in the case of a social media platform, a surveillance subject’s words in a posting would be content whereas the timestamp of when those words were posted or the amount of time the surveillance subject was recording as having spent composing the posting would be understood as metadata. With this loose definition in place, I move on to surveying how the content/metadata distinction has been understood in the four privacy domains.

#### A. Fourth Amendment

The Fourth Amendment is again relevant here. Some believe that a new Fourth Amendment standard, more protective of metadata collected by automated AI processes, may be emerging. For more than a century, the Fourth Amendment has been interpreted to protect a privacy interest in contents but not exteriors, and so a privacy interest exists in the contents of packages and correspondence but not in external signifiers, such as indicating the sender and recipient of information.<sup>158</sup> However, this distinction did not remain an easy-to-implement one, particularly in the rise of digital communications and information technologies, which led to substantial confusion as to which structural portions of digital packets of information were protected by a *Katz* reasonable expectation of privacy and which were more apposite to envelope data.<sup>159</sup>

However, over time, it became increasingly apparent—particularly at the high rate of granularity that metadata was collected—that the distinction between content and metadata that had traditionally governed the lack of Fourth Amendment protection for metadata in third party hands was “increasingly untenable.”<sup>160</sup> In other words, despite the long-standing third-party doctrine (discussed *supra*), courts found themselves increasingly resistant to the notion that metadata in third-party hands could ineluctably be accessed warrantlessly by the government. This discomfort came to a head in the Supreme Court’s 2018 *Carpenter v. United States* decision, in which the Court declined to extend the third-party doctrine to cover warrantless access to cell-site location information.<sup>161</sup> *Carpenter* has been understood by some as a recognition that the traditional inquiry of whether metadata in the hands of a third-party is insufficient in light of enormous levels of automated data collection that otherwise creates highly revealing sets of information.<sup>162</sup> At the least, *Carpenter* has been understood to raise obstacles to the application of the third-party doctrine in cases where the transfer of the metadata to the third party is not clearly voluntary,

---

158. Tokson, *supra* note 148, at 2112–13.

159. *Id.* at 2114, 2116.

160. *United States v. Moalin*, 973 F.3d 977, 991 (9th Cir. 2020).

161. 585 U.S. 296 (2018).

162. See, e.g., Mary-Kathryn Takeuchi, *A New Third-Party Doctrine: The Telephone Metadata Program and Carpenter v. United States*, 94 NOTRE DAME L. REV. 2243 (2019).

not only in a formal sense (electing to use a particular service) but also with regards to knowledge about the transfer and with regards to the necessity of using a form of technology to meaningfully participate in society.

In short, for nearly a hundred years, it seemed clear that envelope data, or metadata, largely, or even universally, did not benefit from Fourth Amendment protection. That convention, however, has been substantially curtailed since the 2018 *Carpenter* decision. The impact of *Carpenter* continues to develop, as in the recent case of a circuit split regarding whether geofencing is a Fourth Amendment search at all<sup>163</sup> and, if so, whether the standard geofence warrants that have come into use in the past decade are unconstitutional general warrants.<sup>164</sup> Given the importance of evolving technology and the shifting Fourth Amendment jurisprudence on this topic, it is particularly important to contemplate the possibility that the privacy valence of machine versus human observation could itself have a complex interplay with the observation of content versus metadata. Recent Fourth Amendment jurisprudence suggests that machine observation may render the privacy distinction between content and metadata irrelevant, whereas human observation may maintain the importance of this distinction.

### *B. Communications Privacy*

These examples from news stories and behavioral studies suggest that lay perceptions of technological surveillance are influenced by the perceived amount or qualitative nature of information thought to be conveyed by metadata or content. Such examples are consistent with the history and structure of the Electronic Communications Privacy Act (ECPA).<sup>165</sup> Title I of ECPA exclusively focuses on protecting communications, and in this context offers a warrant plus regime, in which

---

163. *Compare* United States v. Smith, 110 F.4th 817 (5th Cir. 2024), *with* United States v. Chartrie, 107 F.4th 319 (4th Cir. 2024).

164. 110 F.4th 817.

165. “The principle behind the data / metadata distinction — that people have a greater interest in the content of transmitted information than they do in ‘everything else’ surrounding the transmission — has long informed U.S. statutory law. The Electronic Communications Privacy Act (ECPA), 18 U.S.C. §§ 2510–22, ratified in 1986, distinguishes the content of electronic communications (18 U.S.C. § 2510(8)), whose interception generally requires a probable cause warrant, from dialing, routing, and address information (18 U.S.C. §§ 3127(3)–(4)), whose collection only requires a court order based on relevance to an ongoing investigation. The Foreign Intelligence Surveillance Act (FISA), 50 U.S.C. §§ 1801[–11, 1821–29, 1841–46, 1861–62, 1871], first ratified in 1978 and continually amended since (especially after September 11th), makes a similar distinction. FISA requires a probable cause warrant for electronic surveillance (see 50 U.S.C. § 1805(a)(2)(A), 50 U.S.C. § 1801(a)(4)), at least when there is ‘substantial likelihood that the surveillance will acquire the contents of any communication to which a United States person is a party’ (50 U.S.C. § 1802(a)(B)). By contrast, to obtain a court order for dialing, routing, and address information, only relevance to an authorized foreign intelligence investigation that is not ‘conducted solely upon the basis of activities protected by the first amendment’ is required (50 U.S.C. § 1842(a)(1)).” Helen Nissenbaum, Katherine Strandburg, Kiel Brennan-Marquez & Paula Kift, *Metadata Project*, NYU INFO. L. INST. (2023), <https://www.law.nyu.edu/centers/ili/metadataproject> [<https://perma.cc/HXA3-G7E3>].

communications are protected from interception by law enforcement not only by the warrant requirement under *Katz* but by statutory requirements that require a showing beyond the probable cause requirement for a warrant.

The Wiretap Act defines contents of communications: “contents, when used with respect to any wire, oral, or electronic communication, includes any information concerning the substance, purport, or meaning of that communication.”<sup>166</sup> The contents of communications are entitled to a probable cause plus standard, in that law enforcement must show not only probable cause (the traditional standard for a warrant required under the Fourth Amendment) but must also show that other methods of seeking the information are not readily available.<sup>167</sup>

There is some protection for records, but it is starkly different from that for contents of communications.<sup>168</sup> In the case of stored communications, contents and records are distinguished in the Stored Communications Act at 18 U.S. Code section 2703(a), (b), and (c). Records may be shared freely with any party other than the government,<sup>169</sup> while communications may not be freely shared in this way. Likewise, in the case of disclosure to the government, there are more protections in place when the government would seek access to the content of communications than when the government would seek access to provider records.<sup>170</sup>

Finally, the Pen Register Act is also worth considering. If we accept the basis of *Smith v. Maryland* that numbers dialed are similar to the envelope data (rather than content data) that has for a hundred years been seen as not having a Fourth Amendment privacy protection, this also rationalizes the minimal protections provided by the Pen Register Act. This Act, which protects only a form of metadata rather than content, offers minimal protection both as to scope (it applies only to government, not to private parties) and as to barriers to information collection (the Act has been described as a “rubber stamp”).<sup>171</sup>

The distinction between content and metadata (the latter being everything other than content) also plays out indirectly in state wiretapping laws. Consider a recent decision, *Commonwealth v. Thanh Du*,<sup>172</sup> by the Supreme Judicial Court of Massachusetts, which contemplated whether the suppression remedy in the state wiretapping laws covered only oral recordings in a case where a police officer made a warrantless audiovisual recording during a drug purchase. The defendant moved to

166. 18 U.S.C. § 2510.

167. 18 U.S.C. § 2518(3)(c) (“normal investigative procedures have been tried and have failed or reasonably appear to be unlikely to succeed if tried or to be too dangerous”).

168. 18 U.S.C. § 2703.

169. 18 U.S.C. § 2702(c)(6).

170. Compare 18 U.S.C. § 2703(c)(3) (government access to records does not require notice to the subscriber), with 18 U.S.C. § 2703(b) (government access to contents will require notice to subscribers except in special circumstances).

171. Jim Dempsey, *CDT’s Analysis of S. 2092: Amending the Pen Register and Trap and Trace Statute in Response to Recent Internet Denial of Service Attacks and to Establish Meaningful Privacy Protections*, CTR. DEMOCRACY & TECH., (Apr. 4, 2000), <https://cdt.org/wp-content/uploads/security/000404amending.shtml> [<https://perma.cc/WP5R-HRML>] (“[T]he standard is that of a rubber stamp.”).

172. 245 N.E.3d 1046 (Mass. 2024). I thank Regan Hawkins for bringing this case to my attention.

suppress the recording, but the trial court and the appellate court both ruled to suppress only the audio recording and to admit the visual recording. The Supreme Judicial Court concluded that the visual component of the recording met the broad definition as to the “contents” of a communication because the recording contained information related to the identity of a party to the communication and information relating to the existence of a communication.<sup>173</sup> Under the broad language of the statute, therefore, the court concluded that the visual component of the recording also contained information.

In this state law case, we see that the content/metadata distinction comes into play under communications privacy laws in ways that go beyond the content/records distinction common in this law, and we see how courts often interpret these statutes in highly privacy-protective ways, including by importing a broad meaning of “content” into application of the statute. In short, there is a distinction between content and metadata, but the division is pushed toward a generous inclusion of content to encompass much beyond actual word choice.<sup>174</sup>

---

173. “Under the statute’s broad definition of the ‘contents’ of the oral communication to be suppressed, such video footage is ‘contents’ of the communication for two reasons. First, the footage shows one of the speakers -- here, the defendant -- meaning that the footage contains ‘information concerning the identity’ of a party to the communication. Second, the footage shows the person engaging in the unlawfully intercepted oral communication and therefore contains ‘information concerning . . . the existence . . . of that communication.’ . . . The wiretap act’s [sic] plain language thus requires suppression of the video footage as ‘contents’ of the oral communication; ‘[i]t is not our function to craft unwarranted judicial exceptions to a statute that is unambiguous on its face.’” *Id.* at 1051 (citations omitted).

174. Law may not draw the distinction between content and metadata in quite the same way as common use of recording technologies has been judged in other situations. Many surveillance devices are designed to record visual but not audio information, perhaps on the assumption that such a setup will not violate state wiretapping laws or other electronic surveillance laws. However, there is not a clear consensus that this is the primary reason surveillance devices often record only visual and not audio data. For example, another possibility is that common surveillance devices record only sound to limit file size or due to privacy concerns apart from electronic surveillance statutes. See e.g., *How to Tell if Security Cameras Have Audio: 5 Easy Methods*, ALFREDCAMERA, (June 17, 2024), <https://alfred.camera/blog/how-to-tell-if-security-cameras-have-audio/> [<https://perma.cc/PW9P-8AW2>]. Consider likewise and relatedly an editorial by Angella Foster, a nanny working in Massachusetts. Foster described a sense of violation upon learning that the nanny cams in use at her place of employment were—unbeknownst to her—recording audio as well as visual information. Foster found the collection of audio information far more intrusive than visual information. Foster’s distinction seems to map closely to the distinction made by WhatsApp users (discussed *supra*), who may have accepted that they were “watched” via behavioral metadata recording but did not accept that their conversations would be “listened to” via content scanning. See Angella Foster, *When Parents Eavesdrop on Nannies*, N.Y. TIMES (Aug. 19, 2019), <https://www.nytimes.com/2019/08/19/opinion/nanny-cams-privacy.html> [<https://perma.cc/QDD8-8UKN>]. There are of course distinctions between the situation Foster describes, in which she knew about the visual recording and so perhaps implicitly consented at least to the visual recording, in contrast to *Du*, in which the defendant presumably did not know about the recording and so could not be understood to have consented to even the visual recording.

*C. The Privacy Torts*

There is not any readily identifiable pattern of distinguishing content from metadata in the privacy torts. This would be most likely to arise in scenarios in which both were present but in some way distinguished, but research did not reveal any such cases.

A rare example where a firm was known to perpetrate privacy incursions into contents rather than into metadata occurred when Google began scanning customer emails for purposes of generating targeted advertising. In that case, the firm was rapidly sued under both wiretapping laws and under invasion of privacy torts theories. In any case, the legal response to the firm's decision was so aggressive and widespread that the firm quickly withdrew this email scanning program.<sup>175</sup> In that case, the firm returned to targeting advertising based on consumer metadata but not on the basis of the contents of emails. This offers additional evidence, in addition to that discussed *supra*, that there is a privacy distinction *for consumers* when it comes to whether firms use metadata (that is observations about their behavior) or contents of communications for purposes of targeted advertising.

On the other hand, it's not clear that this content/metadata distinction matters for legal purposes even if it does matter for consumer judgments. The same causes of action (intrusion upon seclusion) are regularly pleaded in cases of consumer privacy litigation related to firms collecting or sharing metadata in ways that plaintiffs allege are tortious. Of course, consumer privacy litigation rarely reaches a decision on the merits but more often is determined on the basis of procedural hurdles. Where plaintiffs overcome those hurdles, cases settle.

It is therefore instructive to contemplate how the distinction of content versus metadata might map onto the elements of the privacy torts, and specifically of intrusion upon seclusion and public disclosure of private facts. Of course, one demonstration of the problematic nature of the distinction of content versus metadata is understanding how to translate this outside the context of communications privacy. For our purposes here, content will be limited to communications, whereas metadata will be understood to be that which can be observed about someone from a formally public location rather than directly communicated from them.

In the case of intrusion upon seclusion, an analysis of the formal elements would seem to cut both ways. To the extent that information typically understood to come into the realm of metadata, such as where someone is, what they look like, or how they behave, might be understood to be publicly accessible, such information would likely fail to be private in the sense expected by the tort. This is also true for the public disclosure tort.

On the other hand, these torts recognize that not all metadata is necessarily of the kind that is public. Some of the intrusion upon seclusion examples of private concerns include examples of what is metadata. In the Second Restatement of Torts, Illustration 4 in intrusion upon seclusion discusses that accessing a party's private

---

175. Alex Hern, *Google Will Stop Scanning Content of Personal Emails*, THE GUARDIAN, (June 26, 2017, 5:39 AM), <https://www.theguardian.com/technology/2017/jun/26/google-will-stop-scanning-content-of-personal-emails> [<https://perma.cc/7P2L-LZKM>].



financial records could constitute an intrusion upon seclusion,<sup>176</sup> an observation backed up by the outcome in *Nader v. General Motors*,<sup>177</sup> where a court likewise found intrusion upon seclusion where thugs crowding Ralph Nader at a public location (the bank) could nonetheless plausibly have intruded upon Nader's private affairs through observation of behavioral financial metadata. This provides some evidence that the privacy torts would protect some kinds of metadata as much as the content of a conversation that was expected to be private.

On the other hand, observational data, such as who someone lives with, is not necessarily entitled to privacy protection. In the Restatement's discussion of public disclosure of private information, there is some basis to find a content versus metadata distinction. In Illustration 20, the Restatement contemplates whether the fact of a social relationship between two people is private information and finds that it is not necessarily private.<sup>178</sup> Where someone can be found, through public observation, to be carrying on a romantic relationship at odds with public morality, this information is not necessarily private. This provides some evidence that the privacy torts will not always protect some kinds of metadata even where that metadata relates to sensitive information.

Finally, another way in which the content versus metadata distinction might correlate to different outcomes is under the "highly offensive" prong, which is present in both the intrusion<sup>179</sup> and the disclosure<sup>180</sup> tort. It seems clear—given the widespread condemnation and rapid legislative response to warrantless wiretapping—that accessing communications in which parties appear to have a reasonable expectation of privacy would be highly offensive to a reasonable person. In the case of observations that could be made in public, it is not clear that they would always meet this high threshold. This is an area, therefore, where one could expect different outcomes under the invasion of privacy torts with respect to content versus metadata (where the latter is broadly understood to relate to public observations). In short, it seems clear that contents of communications would likely benefit from robust privacy protections in more cases than would metadata, but the distinctions

---

176. RESTATEMENT (SECOND) OF TORTS: INTRUSION UPON SECLUSION, § 652B cmt. B, illus. 4 (AM. L. INST. 1977) ("A is seeking evidence for use in a civil action he is bringing against B. He goes to the bank in which B has his personal account, exhibits a forged court order, and demands to be allowed to examine the bank's records of the account. The bank submits to the order and permits him to do so. A has invaded B's privacy.").

177. 255 N.E.2d 765 (1970).

178. RESTATEMENT (SECOND) OF TORTS: PUBLICITY GIVEN TO PRIVATE LIFE § 652D cmt. h, illus. 20 (AM. L. INST. 1977) ("A is run down and injured in an automobile accident. A reporter from B Newspaper, investigating the accident, discovers that A is living with a man who is not her husband. B Newspaper publishes that fact in its account of the accident. This may be but is not definitely an invasion of A's privacy.").

179. "One who intentionally intrudes, physically or otherwise, upon the solitude or seclusion of another or his private affairs or concerns, is subject to liability to the other for invasion of his privacy, if the intrusion would be highly offensive to a reasonable person." RESTATEMENT (SECOND) OF TORTS 110, § 652B (AM. L. INST. 1977) (emphasis added).

180. "One who gives publicity to a matter concerning the private life of another is subject to liability to the other for invasion of his privacy, if the matter publicized is of a kind that (a) *would be highly offensive* to a reasonable person, and (b) is not of legitimate concern to the public." RESTATEMENT (SECOND) OF TORTS § 652D (AM. L. INST. 1977) (emphasis added).

are not clear and not absolute. Further, there seems to be a good case that the privacy torts are so context dependent and fact specific that there is no per se rule or even strong trend that would likely correlate to the content versus metadata distinction.

#### *D. Consumer Protection*

Consumer protection law does not parse distinctions as between contents of communications and everything else. Rather, both FTC enforcement and recent data privacy reform suggest that it is the *sensitivity* of information rather than the content/metadata nature of information that is of importance.

First, consider recent FTC enforcement. Enforcement has focused heavily on the use of observational data, whether it be finding that selling sensitive location information is an unfair trade practice<sup>181</sup> or finding that incorrect observational or inferential conclusions premised on public observations could likewise constitute unfair trade practices.<sup>182</sup> As in the case of consumer privacy litigation, the FTC has not had reason or opportunity to involve itself in cases involving direct observations or misuse of the content of consumer communications. Nonetheless, the agency has been extremely active because it has found so many legally cognizable harms stemming from the use of metadata—that is, stemming from the use of observational data about consumers. This suggests, at the least, that metadata is itself entitled to all the legal protections in both scope and extent that the FTC is poised to offer, in turn suggesting that the content/metadata distinction is likely unimportant for FTC enforcement.

Further to the point that the content/metadata distinction is unimportant in the consumer protection context, we find more evidence in this same vein when contemplating recent consumer privacy statutory reform. The California Consumer Protection Act grants to consumers various controls over any metadata connected to their identities as well as to more traditional categories of personal data, with the latter seeming more content-like.<sup>183</sup> Likewise, Virginia’s new Consumer Data law

---

181. See Fed. Trade Comm’n, *FTC Finalizes Order with X-Mode and Successor Outlogic Prohibiting it from Sharing or Selling Sensitive Location Data* (Apr. 12, 2024), <https://www.ftc.gov/news-events/news/press-releases/2024/04/ftc-finalizes-order-x-mode-successor-outlogic-prohibiting-it-sharing-or-selling-sensitive-location> [<https://perma.cc/KBS5-NLLZ>].

182. See, e.g., Fed. Trade Comm’n, *Rite Aid Corporation, FTC v.*, (March 8, 2024), <https://www.ftc.gov/legal-library/browse/cases-proceedings/2023190-rite-aid-corporation-ftc-v> [<https://perma.cc/M3B9-TJTW>].

183. Malcolm Chisholm, *California Consumer Privacy Act of 2018 vs. GDPR*, FIRST SAN FRANCISCO PARTNERS (June 29, 2018), <https://www.firstsanfranciscopartners.com/blog/california-consumer-privacy-act-of-2018-vs-gdpr/> [<https://perma.cc/E8XM-MMNJ>] (“One interesting difference between the CCPA and the GDPR is a difference between metadata and data. The CCPA explicitly states that a consumer has the right to be informed of the categories of personal data, categories of sources of data and categories of third parties that a business shares personal data with. The GDPR really only speaks about data and the need for plain language in terms of disclosures to data subjects. The emphasis on categories in the CCPA raises some interesting metadata concerns, like what the categories are, how they are defined, and how information, sources and third parties are actually categorized. All of this is metadata. Of course, as noted above, consumers

would reasonably include behavioral metadata.<sup>184</sup> Further, the sensitive categories that receive additional protection under these laws and other recent state data privacy do not readily reflect the distinction between content and metadata.<sup>185</sup> Yet the protection of sensitive data, rather than of content of communications, appears to enjoy a broad consensus. Consider that of the nineteen states that have new data privacy statutes on the books, sixteen prohibit firms from processing sensitive data (which is usually enumerated) unless a consumer *opts in*.<sup>186</sup> In other words, such processing is prohibited by default.

In the most relevant consumer privacy statutes, there is no clear statutory text that singles out privacy protections specific to metadata, nor are there discussions that propose to differentiate explicitly between content and metadata in the context of consumer protection. There are a few potential reasons for this. First, it may be that communications are so sensitive that access to the contents of consumer communications is so unusual that it clearly goes beyond any standard set to protect consumer privacy. In other words, given that this area of law is routinely prepared to recognize privacy incursions related only to metadata—if we are to assume that contents have a more significant privacy valence—this could offer one explanation as to why the distinction need not even be discussed.

A related possibility is that—at least until recently—most (though not all<sup>187</sup>) data harvesting of interest to the commercial sector has been of customer metadata, that is, relating to observations about customers. Further, the unification of consumer privacy protections offered to content and metadata may very well be all the more complete in the case of consumer data. For example, when a customer types “ski socks” into a search engine, this is not only content about her thoughts and not only metadata describing the kind of customer she is. It is ineluctably both, and this too may be a reason that consumer protection law has not grappled with any distinctions along the lines of content versus metadata. This may be changing in the wake of large language models (LLMs). These models *are* reliant on content rather than metadata and have, famously, been trained on various online content, including content produced by ordinary consumers.<sup>188</sup>

---

also have rights to actual data in the CCPA.”).

184. Virginia’s Consumer Data Protection Act, VA. CODE ANN. § 59.1-571 (2025).

185. See Sydney Veatch, *State Privacy Discrimination Law Categories*, (Jan. 10, 2024), [https://docs.google.com/spreadsheets/d/1v9uu8Jzp3AFTYuKw8xB2SYXr3KflpE\\_a4BSbebKk1pQ/edit?gid=0#gid=0](https://docs.google.com/spreadsheets/d/1v9uu8Jzp3AFTYuKw8xB2SYXr3KflpE_a4BSbebKk1pQ/edit?gid=0#gid=0) [https://perma.cc/9A84-92YW].

186. International Association of Privacy Professionals, *US State Privacy Legislation Tracker 2025: Comprehensive Consumer Privacy Bills*, [https://iapp.org/media/pdf/resource\\_center/State\\_Comp\\_Privacy\\_Law\\_Chart.pdf](https://iapp.org/media/pdf/resource_center/State_Comp_Privacy_Law_Chart.pdf) [https://perma.cc/2YG7-5CTH].

187. Consider the example of Google scanning emails, a very contentious practice that the firm ultimately ended following several rounds of consumer protection and wiretapping lawsuits. *Google Message Scanning*, LIEFF CABRASER HEIMANN & BERNSTEIN (2018), <https://www.lieffcabraser.com/privacy/google-message-scanning/#:~:text=The%20complaint%20alleged%20that%20Google,the%20California%20I nvasion%20of%20Privacy> [https://perma.cc/YJ9D-H6SU].

188. Aileen Nielsen, *Whose Data, Whose Value? Simple Exercises in Data and Modeling Evaluation with Implications for Tech Law and Policy*, 99 N.Y.U. L. REV. ONLINE 1, 20 (2024).

*E. Summing Up the Legal Distinctions and Nondistinctions*

In summary, a distinction between content and metadata is quite important to the digital consumer marketplace and is occasionally but not uniformly present in diverse branches of law that govern privacy rights and protections. Evidence from both privacy statutes and Supreme Court jurisprudence suggest that law is increasingly moving towards a regime in which metadata is more protected than it was in the past.<sup>189</sup> However, there are broad areas of law where this continues not to be the case (those covered by federal communications statutes). These observations suggest the utility of experimental evidence to inform the direction of data protection laws and also to assess whether existing statutes (such as federal communications privacy laws) are in need of updating.

**Table 3:** Overview of inconsistencies within and across the four privacy domains in treatment of whether content versus metadata observation is distinct from a privacy perspective.

Domain	Is privacy protection different for content versus metadata?
Fourth Amendment law	Debatable but likely no (historically yes in effect under the third-party doctrine, but likely evolving towards no in many new cases such that the degree of privacy incursion rather than the third-party doctrine will govern metadata)
Communications privacy law	Yes (built into the structure of the Electronic Communications Privacy Act)
Privacy torts	Debatable, but likely no under the fact and context intensive evaluations of the elements of invasion of privacy torts
Consumer protection law	No; it is the sensitivity of information rather than its content/metadata valence that governs likelihood of FTC enforcement or likelihood of special protections under statutory privacy regimes

*F. Empirical Literature*

To date, the human-computer interaction literature has not placed an emphasis on behavioral responses to content versus metadata.<sup>190</sup> Some incidental findings of one

189. *See supra* Sections III.A & III.D.

190. *But see* Anne E. Boustead & Matthew B. Kugler, *Juror Interpretations of Metadata and Content Information: Implications for the Going Dark Debate*, 9 J. CYBERSECURITY 1, 1 (2023) (presenting empirical evidence regarding how jurors are likely to judge content versus metadata evidence of a crime and finding that “the rise of encryption will have a highly heterogeneous effect on criminal cases, with the direction of the effect depending on the exact nature of the facts that the prosecution must prove”).

study provide suggestive evidence regarding the impact of this distinction. Lee et al. studied human interactions with and judgments about an ambulatory SnackBot.<sup>191</sup> They found that people did not “make the distinction between data and information.”<sup>192</sup> For example, participants did not distinguish, in qualitative interviews, between a pure sound recording and a processed speech recording. Yet, when it came to security concerns, participants did make distinctions between different kinds of data, and more often evinced concerns related to the processing of metadata, such as the potential inference that one was arriving late to work based on viewing timestamps. Interestingly, these findings seem to go against the received wisdom that metadata collection would be perceived as less intrusive, both because the distinction isn’t always clear to people, and to the extent the distinction is made by participants, the distinction may surface more concerns about metadata than about content.<sup>193</sup>

The impact of content versus metadata has also been studied empirically in an explicitly legal application. Boustead and Kugler examined how potential jurors might interpret evidence presented to support equivalent assertions. They compared evidence based on the content of a communication—accessible only if the communication was not encrypted—with metadata that could lead to the same inference as the content itself.<sup>194</sup> Boustead and Kugler find a heterogeneous effect. In the case of either content or metadata used to establish equivalent facts necessary for a conviction (such as evidence placing a defendant at the scene of the crime), Boustead and Kugler found no statistically significant rates in the probability of conviction as between participants who saw metadata evidence and those who saw content evidence.

However, there were some differences in Boustead and Kugler’s study. Participants were more likely to find guilt with regard to mens rea when presented with content rather than metadata evidence; on the other hand, they were more likely to find a pattern of behavior when this was established with metadata rather than content evidence. In short, sometimes the difference matters, sometimes it does not, and it will depend on context. This latter finding suggests the importance not only of understanding the privacy valence of content versus metadata but also suggests that the setting (whether one encounters a human versus machine surveillant) can matter to the direction of the effect.

#### IV. AN EXPERIMENTAL APPROACH

To this point, we have seen that a variety of privacy-sensitive behaviors highlight the importance of human/machine and content/metadata distinctions, but that domains of law that are, in theory, responsive to reasonable expectations of privacy have not clearly established a consensus as to the legal import of these distinctions.

---

191. Min Kyung Lee, Karen P. Tang, Jodi Forlizzi & Sara Kiesler, *Understanding Users’ Perception of Privacy in Human-Robot Interaction*, in 2011 6TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI), LAUSANNE, SWITZERLAND 181 (2011).

192. *Id.* at 182.

193. *Id.* Of course, these results could likely be driven in part by the fact that people would probably not expect to have sensitive discussions in front of a SnackBot.

194. Boustead & Kugler, *supra* note 190.

Considering the importance of expectations of privacy, it seems reasonable to seek out empirical guidance: What do people actually make of these distinctions and might that offer any insights to lawmakers? This Article's study looks to one specific surveillance scenario (a state government surveillance program associated with re-entry into family life for previously imprisoned violent criminal offenders), setting out an example of how empirical data can inform previously inconsistent privacy jurisprudence.

Here, we elicit reactions from research participants via a vignette study, in which participants read about a situation and then respond to questions about that situation. A vignette methodology lends itself to studying the topic of surveillance for a number of reasons. For starters, a vignette methodology invites participants to take on perspectives or to experience situations that are outside of their typical experience.<sup>195</sup> Given the many ways in which AI or human surveillance is deployed, it is helpful to have a flexible experimental methodology that allows exploration of diverse perspectives and of experiences that may be atypical or otherwise rare. Also, a vignette methodology relies on a high-level, stylized account of a potential scenario; a vignette, therefore, typically abstracts away from concrete implementation details. This stylization is a strength in that, sometimes, minor details or elements of design can prove dispositive,<sup>196</sup> even though such details are usually unimportant to the central legal questions of interest.<sup>197</sup> Further, a field study would be ethically and logistically challenging in many surveillance scenarios. Ultimately, it is valuable to

---

195. See generally Herman Aguinis & Kyle J. Bradley, *Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies*, 17 *ORG. RSCH. METHODS* 351 (2014).

196. Consider, for example, a surveillance robot designed to look like a puppy as compared to one designed to be a box. Such a detail can make a significant distinction to how humans interact with the robot. See, e.g., KATE DARLING, *THE NEW BREED: WHAT OUR HISTORY WITH ANIMALS REVEALS ABOUT OUR FUTURE WITH ROBOTS* (2021). Or, in the case of "dark patterns," consider the many examples of how seemingly minor design details can change the rate of a legally important behavior (consent by button click) at very high rates. See, e.g., Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub & Thorsten Holz, *(Un)informed Consent: Studying GDPR Consent Notices in the Field*, in *PROCEEDINGS OF THE 2019 ACM SIGSAC CONFERENCE ON COMPUTER AND COMMUNICATIONS SECURITY (CCS '19)* 980 (2019) (showing rates of consent varying by nearly a factor of ten, from 2.4% to 18.4%, depending on whether the placement of a pop-up banner soliciting consent was in the upper right hand corner or lower left hand corner of a screen in a field-experiment on a commercial website).

197. The fact that such details matter, even when law gives no effect (such as to the placement of a banner or the color of a consent button), points to a broader class of questions, of which this study raises just one: What should the law make of distinctions that affect, sometimes drastically, human behavior that have previously been given no effect in law? Often such distinctions are given no importance in law where they are perceived to have no *ex ante* moral valence. The ongoing difficulties regulators and scholars face in defining and potentially mitigating "dark patterns" represent another, more applied struggle with this question; some dark patterns map onto previously existing legal concepts, such as deception or hard sales. Some tactics posited to be potential dark patterns (such as strategic, behaviorally-informed (usually through A/B testing) banner placement or coloring of a button) do not clearly do so and cannot be deduced *ex ante* to be problematic until empirical data reveal them to be so.

have an experiment precisely because a randomized controlled experiment can allow *causal* conclusions about the impact of a factor, and a vignette methodology permits such an experiment without the ethical challenges that would be implicated with a field experiment.

The scenario in the experiment concerns a proposed use of in-home audio monitoring for convicted violent offenders who have been released from prison and who will reside with vulnerable family members. While such a scenario may seem quite dystopian, the program could likely pass legal review as a non-punitive monitoring measure.<sup>198</sup> Moreover, the scenario is technically plausible as well as legally plausible.<sup>199</sup> There are numerous examples of police and prosecutors using audio monitoring, not only as a public safety measure but even as evidence put forward to support an accusation of criminal conduct.<sup>200</sup> Thus, this scenario provides an opportunity to study attitudes in the case of plausible routinized government surveillance applied to a compelling safety purpose, much like domestic surveillance programs already in place at the national level in the service of identifying potential terrorists.<sup>201</sup>

#### A. Design and Procedure

Participants read about two companies bidding to provide an audio surveillance service, one of which used human observers and the other of which used AI. Both companies were said to be reliable and to offer services with the same level of accuracy for the same price.<sup>202</sup> The companies' practices were described to make the degree of data collection and transmission the same, apart from the presence of a human or AI observer.

---

198. See, e.g., *State v. Muldrow*, 912 N.W.2d 74, 88–89 (Wis. 2018).

199. See, e.g., Rebecca Torrence, *Amazon Launches New Elder Care Subscription Service Alexa Together with Emergency Assistance, Fall Detection*, FIERCE HEALTHCARE (Dec. 8, 2021, 11:18 AM), <https://www.fiercehealthcare.com/tech/amazon-launches-new-elder-care-subscription-service-alexa-together-emergency-assistance-fall> [https://perma.cc/QJ9L-KEUL].

200. Consider the case of a Chicago man jailed for a year on the strength of a silent surveillance video showing his car passing through an intersection, combined with AI analysis of data from ambient noise sensors. The proprietary AIs were those of a highly controversial company, ShotSpotter. Garance Burke, Martha Mendoza, Juliet Linderman & Michael Tarm *How AI-powered Tech Landed Man in Jail with Scant Evidence*, AP NEWS (Mar. 5, 2022, 1:23 PM), <https://apnews.com/article/artificial-intelligence-AI-technology-police-crime-7e3345485aa668c97606d4b54f9b6220> [https://perma.cc/K3FQ-7UQJ]. Burke et al. also cite other uses of the AI in criminal proceedings.

201. Of course, many have questioned whether the stated purpose is the purpose actually intended or served by such programs. Such discussions are out of the scope of this study.

202. To the best of my knowledge, there are no existing studies able to directly compare the cost or accuracy of human versus AI surveillance. However, there are indications that equal cost and accuracy are surprisingly plausible in current AI surveillance products used by police. With respect to AI versus human *accuracy* in policing products, consider that the widely used COMPAS recidivism scoring AI performs no better than untrained Amazon Mechanical Turk workers at predicting recidivism. Julie Dressel & Hany Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, 4 SCI. ADVANCES 1, 1 (2018).

Below are the descriptions of the companies<sup>203</sup> presented to participants. The sentences with gray highlighting indicate language only present in the Metadata treatment. The Content treatment descriptions included the same language as shown below but without the gray highlighted portions. The gray highlighting was not present in the experiment but is only used here for brevity of presentation.

- Digi-Ears provides technology for specialized monitoring. They program computers to monitor live audio feeds for sounds of violence. Digi-Ears uses specialized hardware to scramble the audio feed before transmission. In this way, no word-level information is transmitted. Thus, when the audio information is transmitted, an AI analyzes tone and sound but not content, to detect potential violence. **Digi-Ears has strict controls to ensure that only AIs access the audio feeds.**
- Assist-Ears provides human workers for specialized monitoring. They train human workers to monitor live audio feeds for sounds of violence. Assist-Ears uses specialized hardware to scramble the audio feed before transmission. In this way, no word-level information is transmitted. Thus, when the audio information is transmitted, a human worker analyzes tone and sound but not content, to detect potential violence. **Assist-Ears has strict controls to ensure that workers follow proper privacy protocols at all times.**

The experimental manipulation in the vignette experiment should be apparent in the gray highlighted text, which only appeared in one of the experimental treatments. There was a single experimental manipulation: whether content was surveilled or only metadata. In the Content treatment, all audio information was transmitted for analysis. In the Metadata treatment, no word-level information was available.<sup>204</sup> By construction, therefore, the information transmitted in the Metadata treatment is strictly a subset of the information transmitted in the Content treatment. The Metadata treatment involves transmission of scrambled audio such that tone or affect may be preserved but not word-level information. The gray highlighted text only appeared in the Metadata treatment.

After reading about the option of a surveillance program and the two options presented above, participants were asked which program they expected the state to choose. The question wording and possible responses were as follows:

Which company will the state likely select to provide this monitoring service?<sup>205</sup>

---

203. Ordering of the two descriptions was counterbalanced.

204. One way in which this can be achieved is through time-domain scrambling of audio. See Daniel P.W. Ellis & Keansub Lee, *Minimal-Impact Audio-Based Personal Archives*, in CARPE'04: PROCEEDINGS OF THE FIRST ACM WORKSHOP ON CONTINUOUS ARCHIVING AND RECORDING OF PERSONAL EXPERIENCES 39, 39–47 (Association for Computing Machinery ed., 2004).

205. Crowdsourcing may well be the best way to predict the future of government activities. See, e.g., Lyle Ungar, Barb Mellors, Ville Satopää, Jon Baron, Phil Tetlock, Jaime



- Assist-Ears (the human workers)
- Digi-Ears (the AI solution)
- Neither company; this program should not be offered.
- It doesn't matter; both companies sound like acceptable options.<sup>206</sup>

Notice that four response choices were available for this question (and all others) to participants. The question responses were designed to give participants an option to indicate technology neutrality ("It doesn't matter") or to indicate a refusal to participate in the exercise altogether ("Neither company"). Thus, participants were not forced to make distinctions they didn't find meaningful between humans and AI, and they likewise were not forced to choose between least evil options to the extent that they found this surveillance exercise too problematic to tolerate.

The question about the state's choice was counterbalanced with a question about participants' own preferences: "If you were a family member eligible for this program, which monitoring service would you likely select?"<sup>207</sup> This question also gave the same menu of options (human surveillance, machine surveillance, either, or neither). After these two prediction questions, participants indicated their level of support for the surveillance program itself, using a five-point Likert scale in response to the prompt, "Do you support the creation of this monitoring program?"

Participants next responded to counterbalanced questions about subjective privacy and accuracy assessments of the proposed surveillance services. The wording of the subjective privacy question and responses is below. The subjective accuracy<sup>208</sup> query was worded similarly:

In your opinion, how does the monitoring service privacy compare for Assist-Ears and Digi-Ears?

- Assist-Ears protects privacy better than Digi-Ears.
- Digi-Ears protects privacy better than Assist-Ears.
- The services protect privacy equally well.
- Neither company's product protects privacy enough for this sensitive task.

---

Ramos & Sam Swift, *The Good Judgment Project: A Large Scale Test of Different Methods of Combining Expert Predictions*, in AAAI FALL SYMP.: MACH. AGGREGATION OF HUM. JUDGMENT 37 (2012) (finding that crowdsourcing outperformed experts in predicting future government decisions and political events).

206. In all questions with categorical answer choices, the ordering of possible responses was counterbalanced.

207. The pairing of this first-person question with the question about state preference was motivated by existing work showing that a shift in perspective between third person and first person has been associated with strong reversals of technology-related attitudes. See, e.g., Armin Granulo, Christoph Fuchs & Stefano Puntoni, *Psychological Reactions to Human Versus Robotic Job Replacement*, 3 NATURE HUM. BEHAV. 1062 (2019).

208. The wording of the question ("In your opinion") was designed so as to not conflict with information given earlier to participants that the AI and human service had equal accuracy, by emphasizing that a subjective assessment was requested.

Finally, participants passed three attention checks related to the content of the vignette and also provided basic demographic information about themselves. The full experimental text is included in the online Appendix.

### B. Methods

All reported statistical analyses were pre-registered<sup>209</sup> unless explicitly labeled as post hoc. All analysis code and experimental data are available online.<sup>210</sup> The chi-square test (binary outcomes) was used to test for differences; any p-values reported in pairwise comparisons are the result of this statistical test unless otherwise stated. All chi-square test reporting was post hoc (and labeled as such) due to an error in the pre-registration document. Two one-sided t-tests (TOST test) with a medium effect size (Cohen's  $d = 0.5$ ) were used to test for statistical equivalence. Logistic regressions were run with the `glm()` function from the stats package. For comparisons between human and AI monitoring, the comparison was conducted on the sample of those participants who chose a non-neutral technology preference. The statistical comparisons included only those who chose either the human or AI monitoring system and not those who chose the technology-neutral option ("either") or who refused to entertain either option ("neither"). All analysis was conducted with the R statistical language. The chi-square test is included in base R, while equivalence testing was carried out via functions from the TOSTER package.<sup>211</sup> Reported p-values for pairwise comparisons were corrected for multiple comparisons.

### C. Results

#### 1. Sample

Data was collected in November 2021 on the Prolific.com polling platform from a representative<sup>212</sup> sample of 1019 U.S. adults. Only participants who passed all three attention checks (83.3% of participants) were included in the data analysis. The included subject pool was 49.9% female, 47.9% male, and 1.5% non-binary. The average age was 45.2 years, with a standard deviation of sixteen years. White ethnicity was indicated by 76.1% of participants, Hispanic or Latino ethnicity by 5%

---

209. Aileen Nielsen, *An Experimental Matchup of Algorithmic and Human Surveillance*, ASPREDICTED, <https://aspredicted.org/vg3q-26sn.pdf> [https://perma.cc/6NLG-ZM4M] (AsPredicted #78656).

210. Aileen Nielson, *HumanVersusAlgoSurveillance*, OPEN SCI. FRAMEWORK [hereinafter Appendix], <https://osf.io/6f8qp/> [https://perma.cc/9N8A-QBPA] (April 19, 2025, 5:55 PM).

211. See generally Daniel Lakens, *Introduction to Equivalence Testing with TOSTER*, TOSTER (Mar. 28, 2025), <https://aaroncaldwell.us/TOSTERpkg/articles/IntroductionToTOSTER.html> [https://perma.cc/7FNN-YZ42].

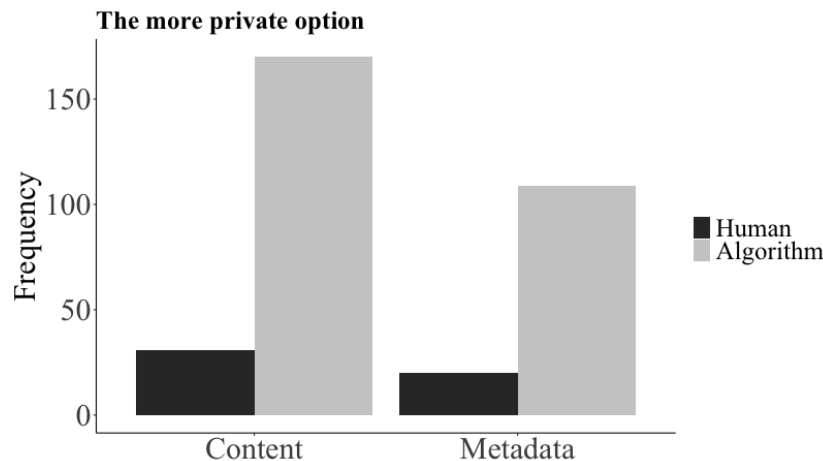
212. The sample was representative by race, age, and gender.

of participants,<sup>213</sup> Black or African American ethnicity by 14.3%, and Asian ethnicity by 6.2%.<sup>214</sup>

## 2. Privacy Assessments of AI and Human Observers

Consistent with the pre-registered hypothesis, the rate of finding the AI service more private (85% for Content, 84% for Metadata)<sup>215</sup> was substantially higher than the rate of finding the human service more private ( $p < .0001$ , post hoc). The distribution of these privacy judgments is shown in Figure 2.

**Figure 2:** AI surveillance was more frequently chosen as the more private option than human surveillance.



A perceived privacy advantage for AI observation did not dominate across the full response space but only in the human/AI match-up. In the Content treatment, 29.9% of participants indicated that the two services protected privacy equally well, and in the case of Metadata, a majority of participants (52.2%) believed the two services adequately protected their privacy.<sup>216</sup> Thus, technology neutrality was a surprisingly common stance among research participants.

213. This proportion of Hispanic or Latino participants does not correspond to U.S. demographics. There appears to be a discrepancy between the survey vendor's labels of Hispanic or Latino participants and how those participants self-identified at the end of this experiment.

214. Participants could select multiple ethnic identities.

215. This refers to the rate among non-technology-neutral choices.

216. See Appendix, *supra* note 210, for the full distribution of privacy judgments.

### 3. Importance of Privacy in Surveillance Choices

The importance of privacy in responding to surveillance is examined in two ways. First, if privacy is a key assessment factor, surveillance programs that provide more privacy should be more strongly supported than those that provide less privacy, all else equal (as it is by construction in this scenario). Second, if privacy is a dominant factor for evaluating surveillance, privacy should be more influential than other factors, such as accuracy, in predicting surveillance choices.

#### a. Support for the Surveillance Program

The mean level of support for the surveillance program was 3.4 in the Content treatment and 3.3 in the Metadata treatment (on a 5-point Likert scale). Consistent with the pre-registered hypothesis, the level of support was statistically equivalent in the two treatments ( $p < .0001$ , 90% CI =  $(-.24, .04)$ ). This provides strong evidence against the importance of the content/metadata distinction in overall support of a state surveillance program, in that the level of support for the program is independent of whether content is included.<sup>217</sup> Circumstantially, this also tends to weaken the case that privacy is the dominant consideration for ordinary people when assessing surveillance.

#### b. Surveillance Choices

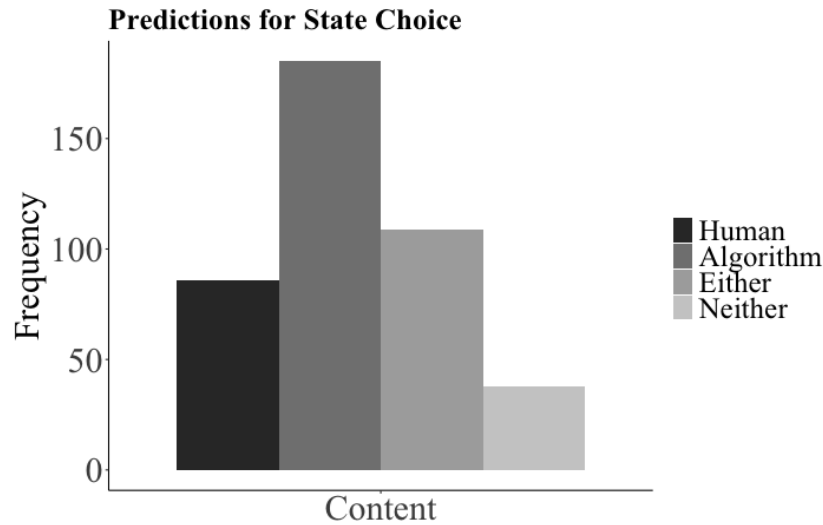
Predictions for the state's choice of surveillance service were spread among all four categorical options, and the distribution of responses is shown in Figure 3. In the Content treatment, 20.6% of participants predicted a choice of the human service, 44.3% of participants predicted the AI service, 26.1% of participants indicated that either service was equally acceptable, and 9.0% of participants indicated that neither service was acceptable. The rates were similar, though not identical, in the Metadata treatment.<sup>218</sup> Overall, the distributions suggest a modal expectation that the state would choose the AI option.

---

217. In lieu of a theory under which privacy is not influential on support for the program, one might speculate that as participants weighed accuracy against privacy, they found that privacy was reduced in the Content treatment but that accuracy was simultaneously increased, meaning that the net desirability of the program was the same. A follow-on experiment was conducted, in which participants were exposed to only one kind of surveillance (Content or Metadata, and computer or AI), and in which numeric elicitations of privacy and accuracy were collected. Privacy and accuracy were found to be positively correlated in all treatment samples, negating the possibility that the equivalent support in the two information treatments could be a function of compensating (inverse) changes in accuracy and privacy as a difference between Content and Metadata that would nonetheless produce the same overall support.

218. The rate of predicting a human system was reduced for the Metadata treatment relative to the Content treatment ( $p < .05$ ), consistent with the pre-registered hypothesis. The change in the rate may have been driven in part by an increase in a technology neutral prediction in the case of the Metadata treatment as compared to Content, although this change was not statistically significant ( $p = .16$ ). The proportional changes from one treatment to the other were closely matched. The rate of predicting the human service decreased from 20.6% to 15.3%, a 5.3% decrease, while the rate of indicating that either service was acceptable

**Figure 3:** AI surveillance was the most commonly predicted choice for the state.<sup>219</sup>



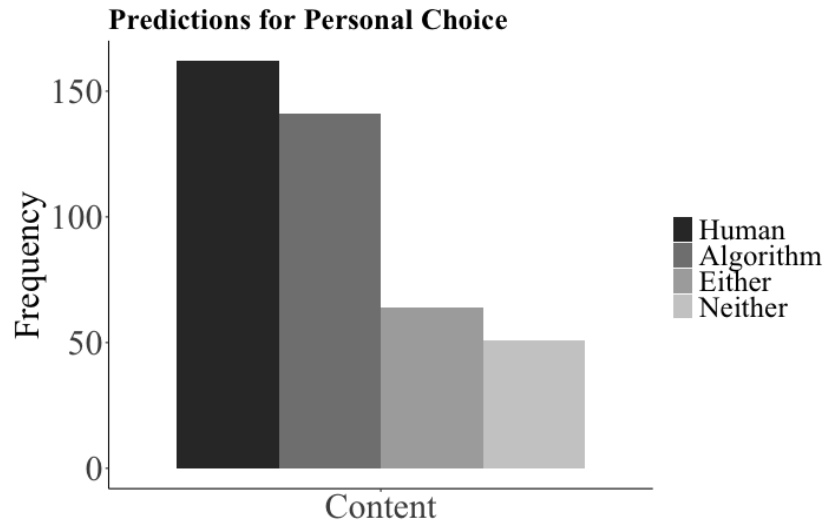
This expectation about the state's preference for an AI solution did not carry over in the case of participants' own preferences. Contrary to the pre-registered hypothesis, participants selected the human system at a far higher rate in the case of their personal preference as compared to their predictions for the state's choice ( $p < .0001$ , post hoc). In the case of personal preference in the Content treatment, the human monitoring system was the modal option, as shown in Figure 4. In the Metadata treatment, AI surveillance was the modal option, but human surveillance remained a popular option, as indicated in the Appendix.

---

increased correspondingly by 4.4%.

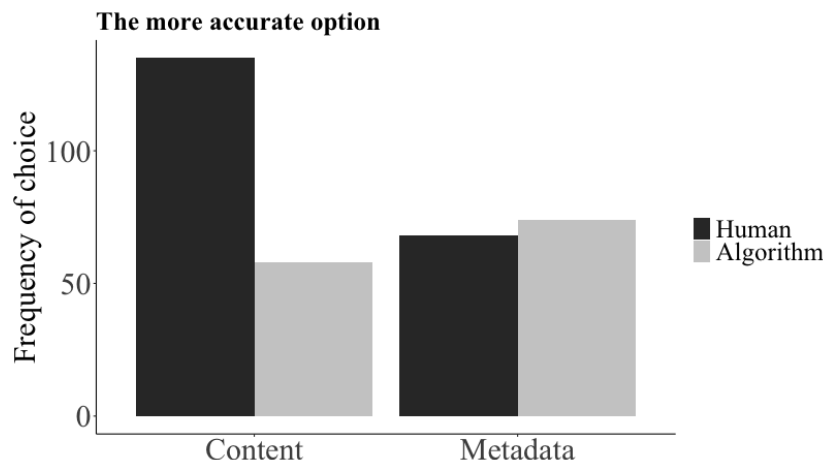
219. Figure 3 presents the distribution of responses for the Content treatment; the distribution was substantially similar for the Metadata treatment.

**Figure 4:** Human surveillance was the most commonly chosen prediction for one's own choice of technology in the Content treatment.



Consider next how the monitoring systems were judged with respect to accuracy, as shown in Figure 5. The distribution of accuracy judgments shifted substantially in response to content. In the Content treatment, the rate of finding the human service more accurate was higher than the rate of finding the AI service more accurate ( $p < .0001$ , post hoc). However, in the case of Metadata, these services were described as the more accurate option equally often ( $p < .0001$ , 90% CI =  $(-.02, .05)$ , post hoc). This distribution of the subjective accuracy judgments suggests a role of accuracy in explaining participants' preferences for surveillance options because it introduces variation of participants that can be regressed against participants' surveillance choices.

**Figure 5:** Accuracy judgments differed for the Content and Metadata treatments.



A logistic regression analysis was conducted to understand the influences of privacy and accuracy on surveillance judgments. The regression variables were expressed as binary indicators of preference for the human service<sup>220</sup> with respect to privacy and accuracy. These, alongside the content treatment and all two-variable interaction terms, were regressed on a personal surveillance technology preference for the human service. The log odds are reported in Table 4. The judgment that the human service was more private influenced the likelihood of choosing the human monitoring system, but so too did the judgment that the human service was more accurate. There was also an interaction effect between privacy and accuracy judgments.

---

220. The results are similar in the case of binary indicators for preference of AIs.

**Table 4:** Logit regression for prediction of self-choice

Dependent Variable: Predicting Personal Choice of Human Surveillance	Log odds
Content	1.29
Human more private	<b>4.92*</b>
Human more accurate	<b>7.42***</b>
Human more private * Human more accurate	<b>.09*</b>
Human more private * Content	.97
Human more accurate * Content	.68
Human more private * Human more accurate * Content	22e6
Intercept	<b>.24***</b>
* p < .01, ** p < .001, *** p < .0001	

The same logistic regression analysis was performed for predictions of the state's choice, as reported in Table 2. Participants who believed the human service to be more accurate were more likely to expect the state to choose the human service, but there was no effect of subjective privacy judgments ( $p = .36$ ).

**Table 5:** Logit regression for prediction of state choosing human service

Dependent Variable: Predicting State Choice of Human Surveillance	Log odds
Content	1.14
Human more private	2.08
Human more accurate	<b>7.37***</b>
Human more private * Human more accurate	.49
Human more private * Content	.99
Human more accurate * Content	.71
Human more private * Human more accurate * Content	1.58
Intercept	<b>.107***</b>
* p < .01, ** p < .001, *** p < .0001	

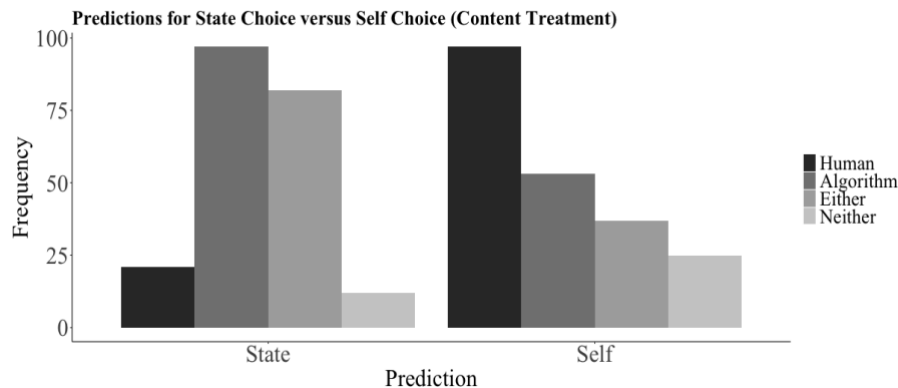
### c. Shifting Choices for the State Versus the Self

Finally, consider the within-subjects difference of how participants predicted the choice of the state as compared to predicting their own preferences. More than 45% of participants gave different predictions for the state's choice versus their own



choice in the imagined place of a vulnerable family member. The distribution of choices for this subpopulation of participants is shown in Figure 6.

**Figure 6:** Shifting choices for the state versus the self



Consistent with the between-subjects results reported above, the rate of selecting the human service increased substantially in the choice for self as compared to the prediction for the state ( $p < .0001$ , post hoc). This appears to be driven by a decrease in the rate of choosing the AI ( $p < .0001$ , post hoc) but also by a decrease in choosing technology neutrality ( $p < .01$ , post hoc). The tendency to prefer human surveillance when choosing for oneself occurred even in the case of participants who indicated a judgment of equal accuracy of the services and even in participants who indicated equal judgments in both accuracy and privacy.<sup>221</sup>

#### *D. Discussion of Results*

The intuition that machine observation is less privacy-intrusive than human observation was strongly supported in this study. But the notion that privacy alone should or does govern the preferences or judgments by ordinary people about state surveillance choices or policy was strongly rebutted. Surprisingly, there are situations in which people would rather be watched by a human than by a computer in a domestic government surveillance scenario—despite strong judgments in favor of AIs as less privacy-invasive.

It's especially surprising that the preference for a human over an AI emerges in the case of the Content treatment. Privacy concerns would intuitively seem to be heightened when more information is transmitted, and so one would expect a preference *not* to be watched by a human to be the strongest case of monitoring for conversational content. But there is a reason for this unintuitive finding. People hold a strong subjective accuracy belief that favors human monitoring in the Content condition, and this belief pushes them to choose human rather than AI monitoring.

221. See Appendix, *supra* note 210.

This finding that privacy considerations are not the driving force in judging humans versus AIs is consistent with the results of studies looking at judgments of humans and AIs in decision-making. For example, in consumer contexts, Bambauer and Risch found that cost, speed, and accuracy—but not privacy—were important in choices between a human or AI decision-making process.<sup>222</sup>

Consider also the result that support for the surveillance program was the same regardless of whether participants were in the Content or Metadata treatment. Though objectively more information was transmitted in the Content treatment, the inclusion of content did not diminish support for the surveillance program overall. This equality of support is not explained, say, by counteracting expectations with respect to accuracy and privacy but rather suggests insensitivity of support to variation of the degree of information transmitted. The lack of influence of degree of privacy on judgments about surveillance is consistent with the work of Kugler and Strahilevitz, who showed that Americans' expectation of privacy was not responsive to the length of tracking time of a surveillance action, in contrast to the proposed "mosaic theory" of the Fourth Amendment.<sup>223</sup>

These two findings together—the relative insensitivity of support for surveillance to the content/metadata distinction (a relative insensitivity to privacy of surveillance support) combined with the sensitivity of surveillance choices to accuracy—raise questions of the legitimacy and efficacy of government communications about domestic surveillance. The government has tried to show that its surveillance programs could be consistent with reasonable expectations of privacy (a common standard in privacy law), but the government has never attempted to show that its surveillance programs meet *reasonable expectations of accuracy*. More concerning still, it seems possible, and even plausible, that in many cases of state surveillance, the state would not be able to meet a standard of reasonable expectations of accuracy.<sup>224</sup>

Throughout the results, there were also interesting patterns with respect to technology neutrality. Technology-neutral choices were more common in the case of the Metadata treatment than in the case of Content treatment. Technology-neutral choices were also more common in the case of predicting the state's choice rather than participants' self-choice. More investigation is needed to understand where lay

---

222. Derek E. Bambauer & Michael Risch, *Worse Than Human?*, 53 ARIZ. ST. L.J. 1091 (2021).

223. Matthew B. Kugler & Lior Jacob Strahilevitz, *Actual Expectations of Privacy, Fourth Amendment Doctrine, and the Mosaic Theory*, 2015 SUP. CT. REV. 205, 209 (2016) ("[O]nly a very small proportion of the respondents in our representative (census-weighted) national sample said that the duration of the surveillance affected whether they would expect privacy in their geolocation information . . .").

224. Consider, for example, recent responses from the New York City Police Department (NYPD) in response to Freedom of Information Law (FOIL) requests by the S.T.O.P. project, in which the NYPD maintained that it had no records regarding the accuracy of the facial recognition products it uses. The NYPD likewise indicated that it had no record of representations by vendors of the products regarding the accuracy of these products. Documents on file with the author. For public records of the legal complaint and responses, see STOP: SURVEILLANCE TECH. OVERSIGHT PROJECT, <https://www.stopspying.org/nypd-facial-rec> [<https://perma.cc/L6EQ-GK3V>].

expectations and judgments of privacy-relevant issues are technology-neutral and where they are not. Perhaps, over time, the distinction between humans and AIs may disappear, at least in some domains.<sup>225</sup>

### *E. Limitations of Results*

There are necessarily limits to the conclusions that can be drawn from this study. First, due to the categorical nature of the data elicited from participants, this experiment sheds light on relative preferences regarding surveillance but not on the strength of those preferences. It could be that those who prefer human surveillance have only a slight preference, while those who prefer AIs have a strong preference. This seems unlikely to be the case. Pilot experiment results looking to numerical expressions of preferences did not identify any such differences in degree of preference.<sup>226</sup> A mismatch in the strength of preferences for different categories of expressed preference seems unlikely to influence the implications of this study.

Another limitation of this study is that the situation portrayed in the vignette is one most likely to affect some historically marginalized communities. It is possible that participants do not strongly identify with the position of a vulnerable family member even when invited to do so. If this were the case, the results of the choice for self should look like the results for the state's choice. Yet, at least for the 45% of participants who diverged in their predictions of state versus self preference for surveillance, it was clear that participants understood their expectations and preferences to differ from those of the state. This suggests that participants were able to empathize or, in other ways, put themselves into the imagined position of the populations most likely to be affected by the kind of state surveillance studied here. Further evidence that participants of all backgrounds gamely entered into the imagined circumstances is shown through the robustness of the trends identified here. The pattern of favoring a human surveillant over an AI surveillant for personal choice was robust when examined in subpopulations of the participant populations demographically more likely to be victims of violence in their homes (females)<sup>227</sup> or more likely to have an incarcerated family member (non-white, non-Asian participants),<sup>228</sup> again suggesting that the pattern of results was not strongly influenced by an inability to empathize with the putative vulnerable family members.

---

225. It is also important to note that the human/AI distinction in the real world will typically be one of immediacy rather than an absolute distinction. That is, it can be more or less probable that a human versus a machine will inspect certain information. It seems unlikely or otherwise implausible that there will commonly be scenarios in which a real guarantee of no human gaze would be promised or enforced. Therefore, the distinction studied here relates to which entity is likely to see information and need not relate to an absolute guarantee.

226. Pilot data available upon request.

227. See, e.g., *Why Do We Say Domestic Abuse is Gendered?*, WOMEN'S AID, <https://www.womensaid.org.uk/information-support/what-is-domestic-abuse/domestic-abuse-is-a-gendered-crime/>, [https://perma.cc/58R6-R5B2].

228. See, e.g., Wendy Sawyer, *U.S. Incarceration Rates by Race and Ethnicity*, PRISON POL'Y INITIATIVE (2020), <https://www.prisonpolicy.org/graphs/raceinc.html>, [https://perma.cc/DE8K-NFXV].

Finally, any potential perceived social distance<sup>229</sup> between the participants and the subjects of the vignette likely reproduces the same dynamic at play when Americans judge domestic surveillance aimed at terrorism,<sup>230</sup> in that they would likely perceive themselves as unlikely to fall victim to terrorist acts or to become terrorists. To the extent that participants did not fully relate to the surveillance subjects, this mirrors likely real-world decision-making conditions.

## V. POLICY IMPLICATIONS

We live in an era when state surveillance is particularly contentious and up for democratic debate, thanks in part to the particularly troubling considerations related to the surveillance at scale that has existed ever since the 2008 adoption of Section 702 of the Foreign Intelligence Surveillance Act (FISA). Most recently, the saga of the December 2023 near-sunset of that provision and subsequent April 2024 time-limited and contentious renewal illustrate strong, bipartisan concerns related to scalable state surveillance.

The 702 program has been contentious in part due to the ideological battle about the contours of exactly how and when Americans can get swept up in this surveillance, but it has also been contentious precisely because it has represented *surveillance at scale*, as can only occur in the case of machine rather than human observation. Historically, surveillance apologists and consumers alike have judged machine surveillance to be less invasive than human surveillance. And yet the results of the experiment show that a judgment that a particular form of surveillance is less invasive does not mean that the particular form of surveillance will be preferred in all cases. Rather, ordinary people also want to know about the accuracy of that surveillance, a figure that continues to be entirely unknown to the American public, even in the face of a deepening crisis regarding the democratic legitimacy of state surveillance.<sup>231</sup> A conversation that needs to be had but has not yet taken place, at least in the sunshine of democratic debate, is whether the FISA 702 surveillance at scale is adequately accurate in identifying threats to American security.<sup>232</sup>

---

229. In freeform responses in pilot studies, multiple participants mentioned that they had witnessed or been victims of violence in their own homes. This serves as a reminder that violence in the home may be a tragically widespread experience.

230. Even as Americans might politically weigh in on domestic surveillance justified as an anti-terrorism matter, that they likely do not see themselves as future victims of terrorism given how (mercifully) low the probability is of being present at a terrorist attack. Likewise, they likely also do not see themselves as possibly falsely accused victims of such a surveillance system. To the extent that participants failed to project themselves into the perspective of the vulnerable family member, they would likewise fail to project themselves into the perspective of a victim of crime or a victim of a false accusation of terrorism. Yet they would still weigh in on these policy issues as constituents whose opinions necessarily shape politics and (indirectly) domestic surveillance policy.

231. For a discussion of the lack of substance in public discussions of FISA 702, see Cindy Cohn, *Word Games: What the NSA Means by "Targeted" Surveillance Under Section 702*, ELEC. FRONTIER FOUND. (Aug. 24, 2016), <https://www.eff.org/deeplinks/2016/08/nsa-word-games-mass-v-targeted-surveillance-under-section-702>.

232. See Louise Matsakis, *Congress Is Debating Warrantless Surveillance in the Dark*, WIRED (Dec. 23, 2017, 7:00 AM), <https://www.wired.com/story/section-702-warrantless->

Indeed, accuracy information is hard to come by even in the case of less sensitive examples of state surveillance where more information is available.<sup>233</sup> Consider the highly contentious ShotSpotter program. It has been difficult to have a public debate about this program in part because the program's accuracy itself is a highly contested figure, with various audits at various times finding wildly different accuracies, including 25% in 2013 in Newark, New Jersey,<sup>234</sup> and 13% in 2024 in New York City,<sup>235</sup> but also 97% according to the proprietary firm itself in 2024 in an unspecified time and geographic range.<sup>236</sup> There are a variety of reasons that the reporting could indicate such different accuracies, some of which are discussed at length in the ShotSpotter firm's response to the NYPD report of 13%.<sup>237</sup> Regardless of the explanation for the inconsistent values, many debates, even about domestic, non-national security-related examples of state surveillance, are taking place in the absence of information, or in the absence of reliable information, about surveillance accuracy.

Alas, there is reason to think these programs probably don't live up to expectations. Consider that the widely used COMPAS program has an accuracy that is shockingly low—no higher than an untrained online survey respondent can replicate.<sup>238</sup> Surely the accuracy to be justified by incursions of privacy and the potential dignitary or privacy harms implicated by use of machine surveillance should *at the least* be better than those made possible by an unskilled human decider.

surveillance-debate/ [https://perma.cc/UC95-RSF2] (“‘There’s been no meaningful assessment, no data-driven cost-benefit analysis,’ says Sascha Meinrath, the founder of the Open Technology Institute at the New America Foundation and the founder of technology policy think tank X-Lab. ‘It’s a massive experiment with no checks, no scientific methodology. We have no idea if this is causing more harm than good, we have no way to know.’”).

233. In some cases of state surveillance, the government has even resisted disclosing information about accuracy of AI surveillance in part on the theory that such information is proprietary and belongs to the firm that supplied the surveillance to the government. This illustrates yet another aspect of why understanding public-private surveillance cooperation is so important given the realities of state surveillance, which is often powered by private sector capacities. *See, e.g.*, Hous. Fed’n of Tchrs., Loc. 2415 v. Hous. Indep. Sch. Dist., 251 F. Supp. 3d 1158 (S.D. Tex. 2017).

234. *See* Sarah Gonzalez, *In Newark, Gunshot Detection System Falls Short of Booker’s Claims*, WNYC (Aug. 9, 2013), <https://www.wnyc.org/story/311533-gunshot-detection-sensors-newark-result-17-arrests-over-three-years/> [https://perma.cc/C9C9-SABC] (“Since 2010, 75 percent of the gunshot alerts have been false alarms. But police are often deployed to the location anyway, just in case there is a shooter.”).

235. *See* Jenna DeAngelis, *ShotSpotter Technology Sends NYPD After False Alarms 87% of the Time, Report Finds*, CBS NEWS (June 20, 2024, 8:31 PM), <https://www.cbsnews.com/newyork/news/nypd-shotspotter-report/>.

236. Jacqueline Berkman, *Optimal Gunshot Tracking for Understaffed Law Enforcement*, SOUNDTHINKING (May 2, 2024), <https://www.soundthinking.com/blog/optimal-gunshot-tracking-for-understaffed-law-enforcement/> [https://perma.cc/KF9X-7L4C].

237. *See Response Letter to NYC Comptroller*, SOUNDTHINKING (July 11, 2024), [https://www.soundthinking.com/wp-content/uploads/2024/07/SoundThinking-Reponse-letter-to-NYC-Comptroller\\_11JUL2024\\_FINAL.pdf](https://www.soundthinking.com/wp-content/uploads/2024/07/SoundThinking-Reponse-letter-to-NYC-Comptroller_11JUL2024_FINAL.pdf) [https://perma.cc/J524-DFRN].

238. *See, e.g.*, Julia Dressel & Hany Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, 4 SCI. ADVANCES 1 (2018).

One of the legal and policy realities made particularly salient and problematic by the experimental results is how rarely any information related to surveillance accuracy or efficacy is put forward for the public. The normative implications of this study are clear: Just as *Katz* taught us that the Constitution protects reasonable expectations of privacy, empirical study shows that ordinary people have *reasonable expectations of accuracy*. It is time for surveillance policymakers in the public and private sector alike to honor such reasonable expectations and provide both the information necessary for appropriate accountability and further, one hopes, for the higher levels of accuracy that would match up to reasonable expectations of accuracy.

In an era in which the accuracy and efficacy of surveillance should be made far more legible and otherwise amenable to governance using machine rather than human surveillance, it is high time to make surveillance law and policy itself more responsive to human judgments. The experimental results show that both the human/machine observer distinction and the content/metadata distinction matter to ordinary people, but this work's survey of the four privacy domains shows that the salience and direction of these distinctions are inconsistent. This state of the law is untenable. As firms and government grow (and merge) in their surveillance capacities, it makes sense that the laws applying to these capacities should be both (1) uniform and (2) responsive to the reasonable expectations of ordinary people, both with regards to their reasonable expectations of privacy and also with regards to their reasonable expectations of accuracy.

#### CONCLUSION

The empirical results are in: AIs are less invasive of privacy than are human observers in a state surveillance scenario. But, this is far from dispositive. Sometimes we prefer to be surveilled by other humans, even when we view this as a less privacy-respecting surveillance program. Law and policy relating to both public and private sector surveillance are unresponsive to these empirical realities. Current policy justifications for domestic surveillance likely reflect an overemphasis on privacy (or, rather, putative privacy) to the detriment of other important factors, such as accuracy. Going forward, policymakers and jurists should do more to learn about and be responsive to reasonable expectations of privacy and accuracy alike.