Idaho Law Review

Volume 51 | Number 3 Article 6

March 2019

Getting to October: Why Understanding Technology Is Essential for Privacy Law

Aaron Massey

GETTING TO OCTOBER: WHY UNDERSTANDING TECHNOLOGY IS ESSENTIAL FOR PRIVACY LAW

AARON MASSEY*

TABLE OF CONTENTS

I. INTRODUCTION

Internet technologies were originally built, deployed, and used by universities and research labs. Usenet is one of these early Internet technologies. It is a distributed discussion forum and file transfer protocol that has been continually operated since 1979. Each fall, new college freshmen would discover Usenet and frustrate experienced Usenet users for a month or two with basic questions about how to use the technology and how it worked. In September 1993, America Online began allowing their users access to Usenet.[1] This triggered a continual influx of new users, and it became known as the "September that never ended."[2] Since then the phrase "Eternal September" has become a demarcation point for technologists referring to problems that result primarily from a lack of technical knowledge.[3]

Technology policy remains trapped in an Eternal September despite the fact that information technologies are both critical and ubiquitous. Marc Andreessen famously announced that software was "eating the world" and that traditional busi-

* Aaron Massey is a Postdoctoral Fellow at Georgia Tech's School of Interactive Computing and the Associate Director of ThePrivacyPlace.org. He earned a PhD and MS in Computer Science from North Carolina State University and a BS in Computer Engineering from Purdue University. He is a member of the ACM, IEEE, IAPP, and the USACM Public Policy Council.

1. WENDY M. GROSSMAN, NET.WARS 11 (1998), *available at* http://www.nyupress.org/netwars/contents/contents.html.

2. *September That Never Ended*, THE JARGON FILE, http://catb.org/~esr/jargon/html/S/September-that-never-ended.html (last visited Apr. 7. 2015).

3. This Article is not the first to appear in legal academic writing to reference "Eternal September" as it relates to technology policy. *See* Paul Ohm, *The Rise and Fall of Invasive ISP Surveillance*, 2009 U. of Ill. L. Rev. 1417, 1430 (2009).

nesses needed to transform themselves into software businesses to remain successful,[4] but many of the laws, regulations, and policies governing these businesses are either outdated or irrelevant in this new technology-dominated world. Some of this is due to advancements in technology that have greatly outpaced laws and regulations.[5] Some of it is due to a failure to react to technological developments. Law schools do not typically cover relevant technologies[6] and often encourage students with STEM backgrounds to specialize in patent law, leaving other areas of law with even less technical expertise.

Individual privacy has been greatly affected by the gap between the reality of information technologies and the mitigations afforded by technology policy. As database technologies developed over the decades, they created privacy concerns with each new development. Computer scientists were concerned with so-called "data banks" and the gap between technical developments and legal awareness of these developments as early as 1969.[7] Today's technology dwarfs the power of those ancient "data banks." The average cell phone is more powerful than all of NASA at the time of the moon landing in 1969.[8] By the end of the century, the privacy concerns had only escalated.[9]

Around this same time, the development of the Internet provided businesses, governments, and other organizations with a new avenue to collect information. Tracking technologies are critical to basic functions of the Internet, but their development held obvious and important implications for privacy, resulting in the first

4. Marc Andreessen, *Why Software is Eating the World*, WALL ST. J., Aug. 11, 2011, http://www.wsj.com/articles/SB10001424053111903480904576512250915629460.

5. *See* JONATHAN ZITTRAIN, THE FUTURE OF THE INTERNET AND HOW TO STOP IT 11–35 (2008) (discussing the rapid advancements in computers themselves and in computer networks). *See also* Orin Kerr, *The Next Generation Communications Privacy Act*, 162 U. PA. L. REV. 373, 390 (2014) (discussing how the Electronic Communications Privacy Act was crafted into law to regulate technologies that have been obsolete for years).

6. *See* Gene Koo, *New Skills, New Learning: Legal Education and the Promise of New Technology*, Berkman Ctr. Res. Publ'n No. 2007-4 1, 2 (2007) (discussing the failure of the law school curriculum to prepare lawyers for a workplace that demands technology-related skills).

7. Lance J. Hoffman, *Computers and Privacy: A Survey*, 1 no. 2 Computing Surveys 1, 85 (1996). Hoffman describes the concern this way: "Most states, however, lag seriously in awareness of contemporary data processing capabilities and techniques. A few of the more highly computerized areas are, however, trying to approach the idea of regional data banks in a rational manner." *Id.* at 87.

8. MICHIO. KAKU. PHYSICS OF THE FUTURE: HOW SCIENCE WILL SHAPE HUMAN DESTINY AND OUR DAILY LIVES BY THE YEAR 2100 21 (1st ed. 2011).

9. *See* Daniel. J. Solove, *Privacy and Power: Computer Databases and Metaphors for Information Privacy*, 53 STAN. L. REV. 1393, 1394–99 (2001) (detailing how legal approaches to privacy have failed to appropriately address databases). *See also* SIMSON GARFINKEL, DATABASE NATION: THE DEATH OF PRIVACY IN THE 21ST CENTURY 2–7 (Deborah Russell ed., 1st ed. 2000) (summarizing the threat to privacy by databanks and surveillance technology).

investigations of the Federal Trade Commission on tracking individuals with technology.[10]

The continuing development of tracking technologies has prevented the establishment of reliable, well-understood privacy norms, leaving privacy conceptually vague and challenging to describe. Users are mostly unaware of tracking technologies, and the routinely fail to take steps to protect their privacy. Regulators tread lightly both to avoid curtailing innovation and to ensure an appropriate response. In the absence of privacy norms, businesses pursue technologies that improve their services and their profitability.

The remainder of this paper is organized as follows: In Section I, I introduce privacy frameworks and summarize why they are too generalized to provide guidance to engineers seeking to build new technologies. In Section II, I examine current tracking technologies and use technical details to create a four-element classification of those technologies that may be useful for policy makers. In Section III, I describe how engineers approach ambiguity in legal texts. Finally, I conclude with a brief summary and a call for active efforts to improve communication between the legal and technical fields.

II. PROVIDING PRIVACY GUIDANCE FOR TECHNOLOGISTS

Privacy by Design (PbD) is an excellent example of a policy initiative that does not provide enough specific guidance to be practical for software engineers. As originally defined by Ann Cavoukian, the Information & Privacy Commissioner of Canada, PbD has seven foundational principles:

1. Proactive not Reactive; Preventative not Remedial
2. Privacy as the Default Setting
3. Privacy Embedded into Design
4. Full Functionality–Positive-Sum, not Zero-Sum
5. End-to-End Security
6. Visibility and Transparency–Keep it Open
7. Respect for User Privacy–Keep it User-Centric[11]

At a glance, these principles are unobjectionable, but they are simply too general to be useful. Rubinstein and Good critique these principles as follows:

Principles 1-3 provide useful, if somewhat repetitive, guidance about the

---

10. *See* FTC, *Online Profiling: A Report to Congress*, (2000), https://www.ftc.gov/sites/default/files/documents/reports/online-profiling-federal-trade-commission-report-congress/onlineprofilingreportjune2000.pdf ; *See also* FTC, *Online Profiling: A Report to Congress Part 2: Recommendations*, (2000), http://www.steptoe.com/assets/attachments/934.pdf.

11. Ann Cavoukian, *Privacy by Design in Law, Policy and Practice: A White Paper for Regulators, Decision-makers and Policy-makers*, OFFICE OF THE INFO. AND PRIVACY COMM'R/ONT., 27–29 (2011).

> *importance of considering privacy issues early in the design process and setting defaults accordingly, but they stop far short of offering any design guidance. Granted, Cavoukian offers more practical advice in several of her technology-specific papers, but she makes little effort to systematize or even summarize the design principles found therein. Principle 4 seems unrealistic in an era when some view personal data as the "new oil" of the Internet and privacy controls only tend to limit the exploitation of this valuable commodity. Principle 5 emphasizes lifecycle management, which is a key aspect of privacy engineering. Principle 6 resembles the familiar transparency principle found in all versions of FIPs, while Principle 7 functions primarily as a summing up of the earlier principles.*[12]

Other high-level privacy frameworks are similarly generic, limiting their utility for engineers. The Organisation for Economic Co-operation and Development's (OECD's) *Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data* serve as a good example because of their statement of the "Fair Information Practice" Principles (FIPPs).[13] The OECD's FIPPs influenced the privacy frameworks developed by the Federal Trade Commission, and both approaches have been criticized extensively for their generality.[14] These generalized frameworks remain disconnected from the approaches taken to address privacy in industry.[15]

III. TRACKING TECHNOLOGIES

More specific guidance is required for engineers building technologies that may affect consumer privacy, and the only way to provide that guidance is to seek to understand the technologies involved. This is a matter of both lawyers seeking technical expertise[16] and technologists seeking to understand the law.[17] The technologies to be examined must be understood well to craft appropriate standards and

---

12. Ira S. Rubinstein & Nathaniel Good, *Privacy by Design: A Counterfactual Analysis of Google and Facebook Privacy Incidents*, 28 BERKELEY TECH. L.J. 1333, 1338 (2013).

13. OECD, *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, ORG. FOR ECON. COOPERATION AND DEV., http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm (last visited Apr. 24, 2015).

14. *See* Kenneth A. Bamberger & Deirdre K. Mulligan, *Privacy on the Books and on the Ground*, 63 STAN. L. REV. 247, 255–60 (2011) (enumerating common critiques of FIPPs-based privacy regulation and enforcement).

15. *Id.*

16. *See* Peter P. Swire, *Trustwrap: The Importance of Legal Rules to Electronic Commerce and Internet Privacy*, 54 HASTINGS L.J. 847, 873–75 (2003) (arguing that for legal privacy protections to be successful, policy makers must carefully study the technologies involved).

17. Aaron K. Massey & Annie I. Antón, *Behavioral Advertising Ethics*, *in* INFORMATION ASSURANCE AND SECURITY ETHICS IN COMPLEX SYSTEMS: INTERDISCIPLINARY PERSPECTIVES 162, (Melissa Jane Dark ed., 2011).

rules. Similarly, technologists must seek to understand the laws and regulations that govern their industries.

Two approaches to technology ethics are useful when considering tracking technologies. In the first approach, technologies are considered to be ethically value-free.[18] That is, until they are used for a particular purpose, they are ethically irrelevant.[19] For example, a hammer can be used to build a tree house or to attack someone, and the use entirely determines the ethical evaluation of the scenario.[20] The designer of the hammer may not have considered how it could be used for negative purposes, or they may not have been able to design a hammer that could function for its intended purposes without also functioning for negative, unintended purposes.[21]

In contrast, the second approach to technology ethics considers the design process.[22] Under this view, tools are imbued with the ethical decisions made by the engineers that designed them.[23] For example, imagine a hammer that could automatically be hard or soft based on whether it was being used to make a tree house or to attack a person. The design of this hammer makes negative uses more difficult to perform and less successful than positive uses. Of course, this design comes at a cost. Even if it were possible to build a hammer that could make these adjustments, it would almost certainly be prohibitively expensive in real-world scenarios.

With these two perspectives in mind, we now examine four categories of tracking technologies around which coherent policy can be formed. First, we examine tracking technologies that are designed defensively. These are technologies that attempt to solve a problem efficiently while maintaining user privacy. Second, we examine technologies designed transparently. These technologies may not be designed to maintain user privacy at all costs, but they do take measures to actively inform users about their operation and provide users with a measure of control over their information. Third, we examine tracking technologies designed to solve a problem without regard to user privacy. Finally, we examine surreptitious tracking technologies, which are designed to track individuals without their knowledge or control. By classifying tracking technologies using these four categories, all of which are based on technical design decisions, policy analysts could craft a meaningful privacy framework.

---

18. *Id.* at 169.
19. *Id.*
20. *Id.*
21. *See id.*
22. *See* Arvind Narayanan, Assistant Professor of computer science at Princeton, Opening Comments at the Markkula Center for Applied Ethics at Santa Clara University panel discussion: Privacy by Design (Jan. 23, 2013), *available at* http://www.scu.edu/ethics/practicing/focusareas/technology/internet/privacy/privacy-by-design.html.
23. *Id.*

A. Defensive Tracking Technologies

Cookies are an excellent example of a defensively designed tracking technology, but it is necessary to explain the problem they were designed to solve before we can detail the defensive measures built into their design. The World Wide Web is a collection of hypertext pages accessed over Internet protocols, including the Hyper Text Transfer Protocol (HTTP).[24] HTTP is a stateless protocol, which means that an HTTP server simply responds to a series of requests without remembering any requests it responded to previously.[25] From a protocol design standpoint this is extremely efficient, and it fits well with the original design goal of providing access to a massive collection of interrelated documents.[26] Unfortunately, the lack of memory severely limits what can be accomplished through a stateless protocol. For example, without a mechanism for remembering previous requests, adding anything to a web-based shopping cart or even remembering who was shopping would be impossible. These are some of the problems the HTTP cookie was explicitly designed to overcome.

Cookies have two important design decisions that are defensive in nature. [27] First, cookies are stored on the client (i.e. as a part of a web browser) rather than the server.[28] This puts control of whether a server remembers a series of actions taken in the hands of the user rather than the server.[29] If a user wanted to be completely forgotten after every web request they made, they could instruct their browser to never store cookies.[30] Second, cookies are only accessible by the web domain that set them.[31] For example, if you visit www.example.com and www.wikipedia.com in that order, then www.wikipedia.com would not be able to read any of the information stored in cookies by www.example.com. The designers could have chosen to allow all web domains to read any information set in a web cookie, but they limited access as strictly as possible while still overcoming the limitations of a stateless protocol.

Cookies still have privacy problems despite their design.[32] For example, cookies are simple text files, but there are no constraints that prevent programmers from encoding data in a non-human-readable format.[33] In an ideal scenario, users may wish to examine the information stored in their cookies. This would allow them to

24. *See generally* ROY T. FIELDING ET AL., HYPERTEXT TRANSFER PROTOCOL--HTTP/1.1 (Jim Gettys ed., 1999), *available at* http://www.w3.org/Protocols/rfc2616/rfc2616.txt.

25. *Id.*

26. *Id.*

27. Aaron K. Massey & Annie I. Antón, *Behavioral Advertising Ethics*, *in* INFORMATION ASSURANCE AND SECURITY ETHICS IN COMPLEX SYSTEMS: INTERDISCIPLINARY PERSPECTIVES 162, 166 (Melissa Jane Dark ed., 2010).

28. *Id.*

29. *Id.*

30. *Id.*

31. *Id.*

32. Massey & Antón, *supra* note 27 at 165.

33. *See id.* at 166.

make more informed decisions about what information they would like to continue sharing with the websites they visit. Consider also the confusion over opt-in cookies and opt-out cookies. These terms refer to the default data collection practices employed by a website.[34] An opt-in cookie means that no information about the user is collected by default and the user must set a cookie in order for their information to be collected.[35] An opt-out cookie means that information about the user's interaction with a website would be collected by default but the user could choose to inform the server not to collect this information by setting an opt-out cookie.[36] Since no standard exists that allows a user to differentiate between the two, users seeking to limit their data collection are left with two unpalatable options when managing their cookies: (1) delete them all and remember to manually reset all opt-out cookies on their next visit to opt-out sites or (2) keep them all and remember to manually delete all opt-in cookies on their next visit to opt-in sites.

Cookies remain the primary mechanism for a website to determine that the same computer or device is returning.[37] Any policy or regulation attempting to mitigate the privacy problems presented by cookies must yield to this basic fact. Without the ability to differentiate computers or devices from one another, the Internet would be unusable. Replacing the cookie would require massive re-architecting and redevelopment of information technologies. Still, cookie-created privacy problems can be mitigated. In addition to the dilemma users face, anti-spyware and anti-virus software often delete all cookies, including opt-out cookies, by default.[38] Creating a standard to differentiate between opt-in and opt-out cookies would mitigate these situations.

These problems are perhaps obvious with 20 years of hindsight and a fair evaluation of web cookie design would consider that they were initially deployed rather early in the history of the modern Internet. Netscape Navigator introduced them in 1994 and Microsoft's Internet Explorer introduced them in 1995.[39] It is also fair to say that the cost of server hardware and the extremely rapid growth of the Internet also influenced the design choices that led to the web cookie.[40] Still, for all its problems, the HTTP cookie was clearly designed to put users in control.

---

34. Massey & Antón, *supra* note 27.

35. *Id.* at 166.

36. *Id.*

37. FTC, ONLINE BEHAVIORAL ADVERTISING: MOVING THE DISCUSSION FORWARD TO POSSIBLE SELF-REGULATORY PRINCIPLES (2007) [hereinafter FTC], https://www.ftc.gov/sites/default/files/documents/public_statements/online-behavioral-advertising-moving-discussion-forward-possible-self-regulatory-principles/p859900stmt.pdf.

38. *Id.*

39. Massey & Antón, *supra* note 27.

40. *See* Paul Ohm, *The Rise and Fall of Invasive ISP Surveillance*, 2009 U. ILL. L. REV. 1417 (2009) (detailing the rapid rise of the Internet and its effect on network technologies).

B. Transparent Tracking Technologies

Building a critical mass of users is one of the biggest problems nascent social networks must overcome. In 2012, Path — a social networking application — used Apple's API for accessing contact information to build their network.[41] Arun Thampi, a programmer from Singapore, noticed the problem while using a man-in-the-middle proxy to examine the Path application on iOS.[42] Apple's API allowed Path to query the contacts stored on an iOS device without notifying the user.[43] Using this API, Path simply uploaded the entire contacts of the user's contacts database.[44] Here's how Thampi described his discovery:

> Upon inspecting closer, I noticed that my **entire address book (including full names, emails and phone numbers) was being sent as a plist to Path.** Now I don't remember having given permission to Path to access my address book and send its contents to its servers, so I created a completely new 'Path' and repeated the experiment and I got the same result – my address book was in Path's hands.[45]

Apple responded to this privacy incident by putting a system in place that notifies the user when applications want permission to access sensitive personal information, like contacts or location data.[46] Such a system provides direct transparency to the user at the time an application first requests the data.[47] Apple also provided ongoing control to the user in the system settings for iOS, which allows the user to periodically review which applications have access to which types of data, and they have also included a similar feature in OS X, Apple's desktop operating system.[48] These transparencies allow users to understand and control how their information is being used.[49]

Path also responded quickly to the incident. Dave Morin, the CEO and Co-Founder of Path, as well as an early employee at Facebook, responded to Thampi's blog post by apologizing at length on Path's company blog:

> We believe you should have control when it comes to sharing your personal information. We also believe that actions speak louder than words. So, as a clear signal of our commitment to your privacy, **we've deleted the entire collection of user uploaded contact information from our**

41. *Path Uploads Your Entire iPhone Address Book to its Servers*, MCLOVIN'S BLOG (Feb. 8, 2012), http://mclov.in/2012/02/08/path-uploads-your-entire-address-book-to-their- servers.html [hereinafter *Path*].

42. *Id.*

43. Aaron Massey & Travis Breaux, *Interference*, *in* INTRODUCTION TO IT PRIVACY: A HANDBOOK FOR TECHNOLOGISTS (Travis Breaux ed., 2014).

44. *Id.*

45. *Path*, *supra* note 41 (emphasis in original).

46. Massey & Breaux, *supra* note 43.

47. *Id.*

48. *Id.*

49. *Id.*

> **servers.** Your trust matters to us and we want you to feel completely in control of your information on Path.[50]

Path's privacy violation is strikingly similar to the 2010 violation by another social network called Google Buzz.[51] Google invited Gmail users to try Buzz.[52] Users who accepted the invitation found that Buzz made their Gmail contacts publically available for others to see as a part of bootstrapping the network.[53] Had the invitation clearly stated that Gmail contact information would be made public, users would have been able to make a more informed decision.[54] Both Path and Google wanted users to connect conveniently with friends on their network.[55] To that end, they used contact information that was originally collected for one purpose (e.g., private email and phone calls) for another secondary purpose (e.g., bootstrapping social network relationships).[56] Users formally complained about Google Buzz to the U.S. Federal Trade Commission (FTC), and the FTC found that Google had used deceptive business practices.[57] As a part of the settlement, Google is required to comply with third party audits of their privacy practices for the next 20 years.[58]

Transparency is not a perfect solution. There is a risk that regular requests for data may turn into mindless click-throughs for users simply seeking to get something done quickly.

C. Privacy-oblivious Tracking Technologies

Although technologies are typically designed with a set of clear purposes in mind, they are often used to achieve radically different goals. E-ZPass is an electronic toll collection system designed to allow motorists to conveniently travel on toll roads, over bridges, and through tunnels without paying at a booth for every transit.[59] This convenience comes with a tradeoff. Each transit is logged for billing purposes.[60] What was once an ephemeral transaction is now recorded, searchable,

---

50. Dave Morin, *We Are Sorry*, PATH (Feb. 8, 2012), http://blog.path.com/post/17274932484/we-are-sorry (emphasis in original).

51. Fed. Trade Comm., FTC Charges Deceptive Privacy Practices in Google's Rollout of Its Buzz Social Network (Mar. 30, 2011), *available at* http://www.ftc.gov/opa/2011/03/google.shtm.

52. *Id.*

53. *Id.*

54. *Id.*

55. *See* Massey & Antón, *supra* note 27, at 164.

56. Fed. Trade Comm., *supra* note 51.

57. *Id.*

58. *Id.*

59. *See* E-ZPASS GROUP, http://www.e-zpassiag.com/about-us (last visited Apr. 24, 2015).

60. *See* E-ZPASS GROUP, http://www.e-zpassiag.com/about-e-zpass/how-does-it-work (last visited Apr. 24, 2015).

and can be repurposed for other needs. E-ZPass records serve as a good example of this and have been repurposed as evidence of cheating in divorce proceedings.[61]

Telephone records may be the most famous example of a privacy-oblivious tracking technology. *Smith v. Maryland* addressed the question of whether telephone metadata—the numbers dialed, duration of the call—required a warrant to be accessed by law enforcement.[62] They do not.[63] The *Smith* decision applied Justice Harlan's concurrence from *Katz v. United States*, which established that the Fourth Amendment applied to areas where people have a "reasonable expectation of privacy."[64] The court found that Smith did not have a reasonable expectation that the numbers he was dialing or the duration of the call, which were known to third parties, would remain private.[65] This reasoning is known as "third party doctrine."[66]

Third party doctrine is a hotly debated policy topic, and it is central to privacy concerns in a cloud-computing world where massive amounts of information—including extremely sensitive or personal information—is shared with third parties through modern technologies like Dropbox or Facebook.[67] Critics argue that the "reasonable expectation of privacy" test is too flimsy to be interpreted consistently and moves the "reasonableness" burden from the government's search to the individual's expectation.[68] Supporters argue that third party doctrine ensures that the Fourth Amendment remains technology neutral—that it will neither expand nor contract to protect more or less than it did previously based on new developments in technology.[69] Regardless of the policy debate, third party doctrine is established precedent. Privacy-oblivious technologies or repurposed technologies will continue to create an avenue for government requests for potentially sensitive information until this precedent changes or better design and engineering practices obviate or minimize the exposure of information.[70]

---

61. Chris Newmarker, *E-ZPass Records Out Cheaters in Divorce Court*, NBC NEWS, (Aug. 10, 2007, 3:30 PM), http://www.nbcnews.com/id/20216302/.

62. Smith v. Maryland, 442 U.S. 735, 737 (1979).

63. *Id.* at 745–46.

64. Katz v. United States, 389 U.S. 347, 360 (1967).

65. *Smith*, 442 U.S. at 742.

66. *See generally* Jim Harper, *Reforming Fourth Amendment Privacy Doctrine*, 57 AM. U. L. REV. 1381 (2008) (introducing Third Party Doctrine).

67. *Id.* at 1401–02.

68. *Id.* at 1383–91 (arguing that Justice Harlan's test in *Katz* "does not tether courts to solid conceptual footings" and that "the Fourth Amendment focuses on the reasonableness of the government's actions in undoing that privacy, not on the reasonableness of the individual's expectations.").

69. Orin S. Kerr, *The Case for Third Party Doctrine*, 107 MICH. L. REV. 561, 580 (2009) (arguing that "[i]f we accept that the Fourth Amendment should stay technology neutral, then we should accept that rule both when new technological practices threaten to expand Fourth Amendment protection as when they threaten to constrict it. Just as the Fourth Amendment should protect that which technology exposes, so should the Fourth Amendment permit access to that which technology hides.").

70. Massey and Breaux, *supra* note 43.

Although government requests for information are a good example of the privacy implications of repurposed technologies, they are not the only example. Internet browsers may be uniquely identifiable even without the use of technologies explicitly designed for identification, like IP addresses or cookies.[71] Information theory lends an explanation for this. Information entropy is a measure of the uncertainty of an information source.[72] For example, if we know an individual's gender, birthday, and home zip code, then we may be in a position to uniquely identify them. Latanya Sweeney famously used this information to identify Governor William Weld's record in a publicly released dataset.[73] She also determined that these three pieces of information would uniquely identify 87% of the U.S. population.[74] In another study, two computer scientists used the same principles of information theory to re-identify approximately 80% of the individuals in a supposedly anonymous dataset released by Netflix for the purpose of soliciting improvements to their movie recommendation algorithms.[75] These broad findings about the challenge of releasing data in a truly anonymous format have inspired quite a bit of debate among academics.[76]

In the case of an Internet browser, all the information the browser provides to the client reduces information entropy.[77] For example, browsers communicate their set of installed plugins and a user-agent string,[78] which is defined as a part of the HTTP specification, to allow servers to know what technologies they have available for rendering web content.[79] This information may be enough to uniquely identify a web browser. Peter Eckersley demonstrated a browser fingerprinting technique, called Panopticlick, which was able to uniquely identify 94.2% of browsers based on information transmitted by browsers for purposes other than unique iden-

---

71. *See* Peter Eckersley, *How Unique is Your Web Browser?*, ELEC. FRONTIER FOUND., *available at* https://panopticlick.eff.org/browser-uniqueness.pdf.

72. C. E. SHANNON, A MATHEMATICAL THEORY OF COMMUNICATION, 11-16 (1948).

73. Latanya Sweeney, *Uniqueness of Simple Demographics in the U.S. Population* (Lab. for Int'l Data Privacy, Working Paper LIDAP-WP4, 2000).

74. *See* Massey and Breaux, *supra* note 43. Note that these statistics are a bit outdated now. The U.S. Census has been conducted twice since the initial study, and U.S. demographics have changed much over that same time period. However, the principles of information theory remain the same; these data points greatly reduce information entropy for the U.S. population.

75. Arvind Narayanan & Vitaly Shmatikov, *Robust De-Anonymization of Large Sparse Datasets* (Feb. 5, 2008), https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf (last visited Apr. 24, 2015).

76. *See* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1703 (2010) (highlighting the challenges and implications of information theory and re-identification technologies for policy makers).

77. Eckersley, *supra* note 71.

78. *See* Rubinstein & Good, *supra* note 12.

79. Fielding et al., *supra* note 24.

tification.[80] Clearly, if privacy is not considered as a central design concern in the development of a new technology, the resulting privacy-oblivious technology may pose serious privacy concerns for future users.

D. Surreptitious Tracking Technologies

A great many technologies are designed to track individuals without their knowledge. The government designs some of these technologies for law enforcement and other surveillance purposes. Businesses design some of these technologies to further their own ends. All of them are designed with two key goals in mind: (1) to track users in some context and (2) to avoid detection by the users being tracked.[81] In the context of the brief ethics discussion at the beginning of this section, these are all technologies with ethical imperatives embedded into their design.[82]

Perhaps the perfect example of a surreptitious tracking technology is the StingRay. Originally designed by the Harris Corporation, the StingRay is a branded version of a technology known as an International Mobile Subscriber Identity catcher or IMSI-catcher.[83] The operation of a StingRay can be summarized succinctly.[84] During normal operation, cell phones regularly communicate with cellular base stations or cell sites.[85] These communications allow the cellular network to route and connect text messages or incoming calls, and they are a necessary part of the network. The StingRay imitates a cell site, allowing it to collect in real time location data, text messages, and the content of incoming calls.[86] It is small enough to be handheld, carried in a vehicle, or even mounted to a drone.[87] StingRays are used by intelligence agencies, the military, and law enforcement. They are impossi-

---

80. Eckersley, *supra* note 71. This paper also describes several approaches that may improve the results in section 3.1. Note that information theory suggests that as browsers become more complex, they will communicate more information. Since this study was completed, the HTML 5 specification has been adopted, adding a variety of new features that each compliant browser must implement. Each new feature is another opportunity to improve browser fingerprinting by reducing information entropy.

81. These technologies must track users in some context to be defined as "tracking" technologies. Similarly, they must be designed to avoid detection to be defined as "surreptitious" technologies.

82. Narayanan, *supra* note 22.

83. *See* Stephanie K. Pell & Christopher Soghoian, *A Lot More Than a Pen Register, and Less Than a Wiretap: What the StingRay Teaches Us About How Congress Should Approach the Reform of Law Enforcement Surveillance Authorities*, 16 YALE J.L. & TECH. 134, 144–148 (2013) [hereinafter A Lot More Than a Pen Register].

84. *Id.*

85. *Id.*

86. *Id.*

87. *Id.*

ble for communicating parties to detect, and they can operate without the assistance of telecommunications companies.[88]

Consider this tracking scenario, which is only possible to do with a StingRay. A suspected criminal seeks to use a burner phone for the first time. Law enforcement on the ground positively identify the suspect making a phone call while walking into a store from a public street. Using their StingRay, they are able to capture a single call coming from a mobile phone inside the store. The contents of the call confirm that the individual is the suspect they are seeking, and they are able to arrest the individual exiting the store. Even in this scenario, the StingRay must, as a part of its normal operation, intercept data from other mobile devices in the area. Law enforcement can positively identify the suspect's phone call by relying on the signal strength. These innocent third parties would never know that their data was intercepted or that they were tracked by law enforcement.

StingRays raise many important legal questions about surveillance technologies,[89] but perhaps the most interesting aspect of the StingRay is its technical operation. The StingRay and other IMSI-catchers rely on cellular network protocols that have been designed into devices and cell sites for decades. Although previously only available to governments willing to pay as much as $400,000, it is now possible for dedicated private hobbyists to build homemade IMSI-catchers for less than $1,000.[90] These technologies will only become cheaper, and the longer the cellular network infrastructure remains vulnerable, the cost to prevent their operation[91] will only become more expensive.

Surreptitious tracking technologies are not exclusively designed for use by governments; they are also designed by businesses.[92] Apple's Safari web browser blocks third party cookies by default, which makes placing the initial cookie needed to track users more difficult for third party ad networks.[93] In 2011, Google bypassed this default setting to set a cookie for their advertising network without noti-

88. *Id.*

89. *See* A Lot More Than a Pen Register, *supra* note 83, at 163–64; Stephanie K. Pell & Christopher Soghoian, *Your Secret StingRay's No Secret Anymore: The Vanishing Government Monopoly Over Cell Phone Surveillance and its Impact on National Security and Consumer Privacy*, 28 HARV. J. L. & TECH. 1, 32 (2014) [hereinafter Your Secret StingRay's].

90. Depending on the capabilities, it may be possible to build a passive identification-only device for less than $100. *See* Your Secret StingRay's, *supra* note 89, at 46–54.

91. Because of the wide variety of capabilities available to a StingRay, the technical details of how to prevent their operation are non-trivial. StingRays depend on weak encryption protocols and poor security practices prevalent throughout cellular networks. Backwards compatibility for existing devices alone would make upgrading the network a multi-decade effort. *See generally* A Lot More Than a Pen Register, *supra* note 83; Your Secret StingRay's, *supra* note 89.

92. FTC, GOOGLE WILL PAY $22.5 MILLION TO SETTLE FTC CHARGES IT MISREPRESENTED PRIVACY ASSURANCES TO USERS OF APPLE'S SAFARI INTERNET BROWSER (Aug. 9, 2012), *available at* https://www.ftc.gov/news-events/press-releases/2012/08/google-will-pay-225-million-settle-ftc-charges-it-misrepresented.

93. *Id.*

fying their users and in violation of their stated privacy practices.[94] These actions resulted in a fine of $22.5 million from the Federal Trade Commission in 2012[95] and a settlement of $17 million for a case involving 37 states and the District of Columbia.[96]

Verizon employs a different technique for tracking users without their knowledge: they simply inject a unique identifier as an extra HTTP header for all web traffic generated on Verizon devices.[97] The design of this technology allows Verizon devices to be tracked even if subscribers have opted out, cleared their cookies, or used private browsing mode.[98] A wide variety of advertising agencies relied on Verizon's unique header and were able to use it to re-construct cookies that were missing or deleted by the user, allowing them to continue tracking users who attempted to prevent themselves from being tracked.[99]

Surreptitious tracking technologies rely in part on secrecy. The companies behind these technologies use nondisclosure agreements and intellectual property protections to prevent scrutiny.[100] Prosecutors have dropped serious criminal charges to protect the details of law enforcement use of StingRay devices.[101] If these technologies are not well understood by regulators, policy makers, and consumers, then no legal framework for privacy can sufficiently address the resulting violations.

IV. AMBIGUITY IN LAWS AND REGULATIONS

The privacy frameworks outlined in Section II are insufficient not because they contain ambiguities, but because they contain ambiguities that provide too little context to be disambiguated. Laws and regulations often contain intentional

94. *Id.*

95. *Id.*

96. Claire Cain Miller, *Google to Pay $17 Million to Settle Privacy Case*, N.Y. TIMES, Nov. 18, 2013, http://www.nytimes.com/2013/11/19/technology/google-to-pay-17-million-to-settle-privacy-case.html?_r=0.

97. *See* Jonathan Mayer, *How Verizon's Advertising Header Works*, WEB POLICY (Oct. 24, 2014), http://webpolicy.org/2014/10/24/how-verizons-advertising-header-works/; *See also* Jacob Hoffman-Andrews, *Verizon Injecting Perma-Cookies to Track Mobile Customers, Bypassing Privacy Controls*, *Deeplinks Blog*, ELECTRONIC FRONTIER FOUNDATION, (Nov. 3, 2014), https://www.eff.org/deeplinks/2014/11/verizon-x-uidh.

98. *Id.*

99. Jonathan Mayer, *The Turn-Verizon Zombie Cookie*, WEB POLICY (Jan 14, 2015), http://webpolicy.org/2015/01/14/turn-verizon-zombie-cookie/.

100. FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 1–18 (2015).

101. Ellen Nakashima, *Secrecy around police surveillance equipment proves a case's undoing* (Feb. 22, 2015), WASH. POST, http://www.washingtonpost.com/world/national-security/secrecy-around-police-surveillance-equipment-proves-a-cases-undoing/2015/02/22/ce72308a-b7ac-11e4-aa05-1ce812b3fdd2_story.html.

ambiguity.[102] For example, regulators sometimes employ generalized standards for information technology rather than explicit rules detailing precise functionality.[103] Consider the simple example of encryption technologies. An explicit rule would require a particular encryption algorithm, such as 256-bit AES encryption.[104] Explicitly requiring a strong encryption standard in a regulation may appear to be a best practice, but it actually has a critical weakness: what if the encryption standard is found later to be broken? How will the regulation be updated? Instead of an explicit rule, a general standard could be adopted. For example, the regulation could require the use "reasonably secure encryption mechanisms." This phrasing is intentionally ambiguous. The engineer implementing technologies that must comply with this standard must reify the phrase into a specific encryption technique, and the regulatory system in which the technology is deployed must be able to determine whether a given encryption technique qualifies as "reasonably secure."

Generalized standards require technical expertise to be interpreted. Unfortunately, traditional approaches to requirements engineering are not sufficient for engineers seeking to disambiguate these regulations; engineers and policy makers must collaborate to successfully interpret intentionally ambiguous standards.[105] For these scenarios, engineers need new methods to improve and demonstrate legal compliance in software systems.[106] Evidence suggests that although individual engineers are able to identify portions of a legal text that are ambiguous, they vary greatly on their classifications of these ambiguities.[107] Improving regulatory compliance in requirements engineering is an active area of research,[108] and it is an opportunity for policy makers and engineers to collaborate to improve compliance.

Ambiguity that can consistently be interpreted by both policy makers and engineers must be the goal for future privacy regulations. The four categories of tracking technologies outlined in Section III are intentionally ambiguous; they de-

---

102. Paul N. Otto, *Reasonableness Meets Requirements: Regulating Security and Privacy in Software*, 59 DUKE L.J. 309, 314 (2009).

103. *Id.* (arguing that broad standards allow engineers and policy makers the flexibility needed to allow regulations the flexibility needed to successfully achieve their intended outcomes).

104. National Institute of Standards and Technology, *Announcing the Advanced Encryption Standard (AES)*, FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION 197 (November 26, 2001), http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf.

105. Aaron K. Massey et al., *Identifying and Classifying Ambiguity for Regulatory Requirements*, REQUIREMENTS ENGINEERING CONFERENCE (RE), 2014 IEEE 22ND INTERNATIONAL (2014), *available at* http://dx.doi.org/10.1109/RE.2014.6912250.

106. Otto, *supra* note 102.

107. *Id.*

108. This is evidenced by a distinct increase in publications involving legal compliance concerns at the annual IEEE International Requirements Engineering Conference and by the development of the International Workshop on Requirements Engineering and Law. *See Workshop Overview*, Eighth International Workshop on Requirements Engineering and Law (Aug. 25, 2015), http://gaius.isri.cmu.edu/relaw/2015/.

scribe generally four types of tracking technologies based on technical design aspects of those technologies.[109] Although these categories are limited to a specific context (i.e., tracking technologies), they are general enough to cover the context and remain interpretable by, and therefore useful to, both engineers and policy makers.

V. CONCLUSION

Effective communication between engineers and policy makers must become commonplace to protect privacy and improve the state of technology policy. Engineers must seek to better understand and incorporate laws and regulations into the design of their technologies; this is not only an ethical imperative,[110] but it is also a critical aspect of developing professional standards for software engineering. Policy makers and lawyers must seek to understand the technologies beyond a surface level. Far too often policy makers identify a general policy solution or principle and then fail to state it with enough specificity to be accessible to the engineers building tomorrow's technologies.

The development of new technologies continues. As everyday things become computerized and the Internet of Things becomes a reality, we will once again experience a societal shift in our understanding of technology that will undoubtedly affect laws and regulations. The collection of Big Data and its analysis by algorithms kept secret by governments and businesses will also strain our ability to craft coherent regulation. These developments will preserve our eternal September unless engineers and policy makers undertake a concerted effort to get to October together.

109. *See supra* Part III.

110. The ACM Code of Ethics requires ACM members to "know and respect existing laws" well enough to recognize when a law or rule is "immoral or inappropriate." *See Code of Ethics*, ASSOCIATION FOR COMPUTING MACHINERY, http://www.acm.org/about/code-of-ethics (last visited Apr. 24, 2015).