

IS DATA SPEECH?

Jane Bambauer*

Privacy laws rely on the unexamined assumption that the collection of data is not speech. That assumption is incorrect. Privacy scholars, recognizing an imminent clash between this long-held assumption and First Amendment protections of information, argue that data is different from the sort of speech the Constitution intended to protect. But they fail to articulate a meaningful distinction between data and other more traditional forms of expression. Meanwhile, First Amendment scholars have not paid sufficient attention to new technologies that automatically capture data. These technologies reopen challenging questions about what “speech” is.

This Article makes two overdue contributions to the First Amendment literature. First, it argues that when the scope of First Amendment coverage is ambiguous, courts should analyze the government’s motive for regulating. Second, it highlights and strengthens the strands of First Amendment theory that protect the right to create knowledge. Whenever the state regulates in order to interfere with the creation of knowledge, that regulation should draw First Amendment scrutiny.

In combination, these claims show clearly why data must receive First Amendment protection. When the collection or distribution of data troubles lawmakers, it does so because data has the potential to inform and to inspire new opinions. Data privacy laws regulate minds, not technology. Thus, for all practical purposes, and in every context relevant to privacy debates, data is speech.

* Associate Professor of Law, University of Arizona, James E. Rogers College of Law; J.D., Yale Law School; B.S., Yale College. Thanks to Derek Bambauer, Toni Massaro, Leslie Francis, Barbara Atwood, Chris Robertson, Eric Goldman, Eugene Volokh, Teneille Brown, James Grimmelman, Frank Pasquale, Woody Herzog, Peter Swire, Heidi Anderson, Felix Wu, Marc Blitz, Rebecca Tushnet, Dan Hunter, Jonathan Zittrain, David Post, Andrea Matwyshyn, Mike Madison, Berin Szoka, Geoff Manne, Joe Tomain, Kirsten Engel, Ellie Bublick, Jean Braucher, Robert Glennon, Larry Ponoroff, Ted Schneyer, David Marcus, Roy Spece, Marc Miller, Billy Sjoström, John Swain, Carol Rose, Brent White, Aurora Kaiser, Nancy Sjoström, and Endre Stavang for helpful suggestions and support.

INTRODUCTION.....	58
I. WHAT IS “DATA,” AND WHAT IS “SPEECH”?.....	64
A. <i>Data</i>	65
B. <i>Speech</i>	66
II. DATA IN FIRST AMENDMENT PRECEDENT	70
A. <i>Existing Data</i>	71
1. <i>Gleanings from privacy law</i>	72
2. <i>Gleanings from commercial speech regulations</i>	74
3. <i>Gleanings from copyright law</i>	75
B. <i>Data Creation as Nonexpressive Conduct</i>	77
C. <i>Data Creation as Expressive Conduct</i>	84
III. THE RIGHT TO CREATE KNOWLEDGE	86
A. <i>The Negative Right to Create Knowledge</i>	87
B. <i>Data and First Amendment Objectives</i>	91
1. <i>Marketplace of ideas and public good theories</i>	92
2. <i>Deliberative democracy and a check on state power</i>	97
3. <i>Self-Determination</i>	100
C. <i>The Level of Scrutiny</i>	105
IV. COPING WITH SCRUTINY	110
A. <i>Concern 1: Scrutiny Will Kill Privacy and Other Good Things</i>	110
B. <i>Concern 2: If Data Is Speech, Regulators Are Hamstrung</i>	114
C. <i>Concern 3: Scrutiny Is Subjective at Best, Political at Worst</i>	117
CONCLUSION.....	117

[I]f there is any principle of the Constitution that more imperatively calls for attachment than any other it is the principle of free thought . . .

—Justice Oliver Wendell Holmes, Jr.¹

INTRODUCTION

When does factual information become speech? The appealing, simple answers reside at the extremes.

Perhaps information is always speech. After all, information communicates, educates, and persuades. A single plain fact can do more to change minds and alter debates than a thousand opinions. But this rule goes too far. There are many times that an event will leave a mark that has the potential to retell its story. A car may careen into a barrier and leave a streak of paint. Long after the car is towed, the streak states, in a way, when and where the crash occurred, how fast the car was traveling during impact, and what color the car was. The streak of paint can be received and interpreted by a human to create

1. *United States v. Schwimmer*, 279 U.S. 644, 654-55 (1929) (Holmes, J., dissenting).

knowledge.² But if a city repainted the barrier, we would not interpret this as a decision related to speech. Likewise a crack in the sidewalk might tell the story of a frost, but a municipal ordinance requiring property owners to maintain their sidewalks would not be an act of censorship. Every cell contains DNA, the body's ultimate archive of information, and yet the proper disposal of used syringes does not, and should not, implicate First Amendment scrutiny.

Since data is expressed in alphanumeric symbols, it certainly looks a lot more like traditional speech than a crack in the sidewalk.³ However, conceptually it is sometimes not so far off. When data is the byproduct of other events and services—transactions between a home computer and a website's server, or between a cell phone and a cell tower—these records are no different from other unanticipated marks created by the bustle of life. They have no intended author, and no intended audience. They are mere footprints.

If information isn't always speech, then perhaps it is never speech. After all, the quintessential First Amendment litigant expresses some unpopular idea or opinion. Maybe the domain of the First Amendment is opinion, and other products of human subjectivity. This rule runs into unsalvageable problems even more quickly than the last. Some of the most important modern speech cases concerned the unadorned reporting of raw information. The Supreme Court protected the *New York Times's* distribution of the Pentagon Papers from prior restraint by the federal government on the basis of its First Amendment rights to do so.⁴ The *New York Times* had received a leak of the Pentagon Papers—a report created by the Department of Defense to document the U.S. involvement in the Vietnam War. Many details in the report conflicted with the information that the Johnson Administration had told the public, so the Pentagon Papers were significant not only for the details within them, but also for the inevitable inference that President Johnson had deceived Americans. Justices Black and Douglas had no trouble recognizing the speech interests in raw facts—the fact of the Pentagon Papers' existence and the facts that were reported in them.⁵ It was not necessary for the *New York Times* to editorialize in order to access the First Amendment's protection.⁶

A distinction can be made between statements of fact that are observed and written by a human and those that are collected mechanically, and it might be tempting to draw the First Amendment line between the two. But the distinc-

2. It is, as Gregory Bateson has put it, "a difference which makes a difference." GREGORY BATESON, *STEPS TO AN ECOLOGY OF MIND* 453 (1972) (emphasis omitted).

3. Although the word "data" is the plural form of "datum," in the context of this Article, I usually use the word as a mass noun (like "information" or "the media").

4. *N.Y. Times Co. v. United States*, 403 U.S. 713, 714 (1971) (per curiam).

5. *Cf. id.* at 715 (Black, J., concurring).

6. Moreover, even facts that are not connected to matters of grave public concern, such as the price of a prescription drug, may receive the treatment and protection of speech. *See, e.g., Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, Inc.*, 425 U.S. 748, 770 (1976).

tion is untenable. Suppose Congress were to pass a law mandating the destruction of mechanically captured climate science data and the discontinuation of the collection of core samples. It is implausible to think that a court would not employ some form of First Amendment scrutiny.⁷

The two absolute positions, that data always is speech or never is, simply fail to describe an understanding of the First Amendment that society could tolerate. If neither of the poles can be correct, the truth is somewhere in the messy middle.

Delineating exactly when information receives speech protection, and why, would not have been worth the energy before the era of mass computing, but the building momentum of data analytics and the competing interests of personal privacy have turned this once-arcaic question into a matter of great significance and urgency.

Privacy advocates naturally turn to lawmakers to combat the effects of increased data collection and cheap data storage. They are not alone. As smart phones and other recording devices become ubiquitous, corporations have come to the well, too, pressing legislators to create or strengthen laws that protect their interests in secrecy. Of course, these political forces have little overlap in support. Consumer groups press for privacy laws while corporations lobby for trade secret laws and various industry-specific protections (for example, so-called “ag-gag” laws for animal agriculture companies⁸ or criminal statutes outlawing card-counting devices for casinos⁹), but the motivation for these disparate efforts is the same: to preserve the information status quo.

These efforts clash with the First Amendment.

This Article contends that the freedom of speech carries an implicit right to create knowledge. When the government deliberately interferes with an individual’s effort to learn something new, that suppression of disfavored knowledge is presumptively illegitimate and must withstand judicial scrutiny.

The application of this rule to information law is straightforward. Data is not automatically speech in every context; as the opening hypotheticals illustrate, data can be generated without any expectation to be reviewed and interpreted. But asking whether all data should be treated as speech misses the

7. If Congress is careful, this could be crafted in a viewpoint-neutral way since all researchers, regardless of their beliefs about global warming, would be burdened by the regulation.

8. Mark Bittman, Op-Ed., *Who Protects the Animals?*, N.Y. TIMES OPINIONATOR (Apr. 26, 2011, 9:29 PM), <http://opinionator.blogs.nytimes.com/2011/04/26/who-protects-the-animals> (coining the term “ag-gag”); see also UTAH CODE ANN. § 76-6-112 (LexisNexis 2013) (defining the crime of “agricultural operation interference”).

9. NEV. REV. STAT. § 465.075 (2013) (criminalizing the use of a device which “[a]nalyzes the probability of the occurrence of an event relating to [a] game”). For a thorough discussion of the First Amendment implications of this and similar card-counting statutes, see Adam Kolber, *Card Counting and Freedom of Thought* 1, 5 (May 29, 2013) (unpublished manuscript) (on file with author).

point: any time the state regulates information precisely *because* it informs people, the regulation rouses the First Amendment.

The right to create knowledge reinforces American commitments to autonomy and intellectual curiosity. It at once transcends and supports “speaker” and “listener” rights by protecting observation and thought—the very things that make speaking and listening so valuable. It deserves to be celebrated. Nevertheless, I will proceed with caution because the implications of the right are quite far reaching and profound.

For one thing, as I describe in Part II, the right to create knowledge explodes the classic distinction between information and information *gathering*, which has long been used by courts to define a boundary between protected speech and unprotected conduct. Mechanically captured recordings of conversations and images are often treated as information-gathering conduct, and have avoided First Amendment scrutiny on that basis. A large and growing body of wiretap statutes and information security laws rely heavily on the distinction. But a rigorous review of the cases distinguishing information from information gathering reveals that the reasoning is flawed. The cases treat a journalist who observes an event differently depending on whether she records the incident by taking meticulous notes using pad and paper (protected) or by using the video function on her iPhone (unprotected). The line arbitrarily favors old technologies (memory, pens, even word processors) over new ones.

Recent, better-reasoned case law has already begun to correct the inconsistency.¹⁰ If the dissemination of mechanical recordings receives First Amendment protection (which it does),¹¹ then the creation of those same recordings must have First Amendment significance, too.

Many readers will feel uneasy about the breadth of a First Amendment that interferes with the modern attempts to limit the collection of personal data. The proposals in this Article can be used to challenge the constitutionality of many popular privacy statutes (for example, the Health Insurance Portability and Accountability Act of 1996¹²). They throw into doubt the Federal Trade Commission’s varied privacy initiatives and frustrate the White House’s proposed “Consumer Privacy Bill of Rights.”¹³ All-party-consent wiretap statutes would

10. See discussion *infra* Part II.C.

11. *Bartnicki v. Vopper*, 532 U.S. 514, 527 (2001) (“It is true that the delivery of a tape recording might be regarded as conduct, but given that the purpose of such a delivery is to provide the recipient with the text of recorded statements, it is like the delivery of a handbill or a pamphlet, and as such, it is the kind of ‘speech’ that the First Amendment protects.”).

12. Pub. L. No. 104-191, 110 Stat. 1936 (codified as amended at scattered sections of 18, 26, 29, and 42 U.S.C.).

13. Do-Not-Track Online Act of 2011, S. 913, 112th Cong. (2011); Do Not Track Me Online Act, H.R. 654, 112th Cong. (2011); FED. TRADE COMM’N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS (2012) [hereinafter FTC PRIVACY REPORT], available at <http://ftc.gov/>

have to be justified by important or compelling state interests, and the various “Do Not Track” bills, as well as the pending and future attempts to limit the collection and disclosure of location-tracking data, would be subjected to constitutional scrutiny, too. It would be quite natural to have reservations about a theory of constitutional law that has such sweeping effects.

Anticipating the disruption that constitutional scrutiny would cause, privacy scholars have built a consensus conveniently placing data outside the scope of the First Amendment’s protection. Neil Richards has argued that personal data is not speech because it is more commodity than expression.¹⁴ Ashutosh Bhagwat recognizes that there is some basis for treating data as speech, but because the consequences are “dramatic and troubling,” he advises jurists to evade scrutiny by characterizing data as less valuable, purely private speech.¹⁵ And Tim Wu concludes that data cannot be speech because computers do not have constitutional rights.¹⁶

These scholars write in stubborn service to the privacy and consumer regulations that seem most sensible to them.¹⁷ Certainly there are many reasons to believe that privacy regulations make good public policy. But optimal policies are not immune from First Amendment constraints.¹⁸ More importantly, these

os/2012/03/120326privacyreport.pdf; WHITE HOUSE, CONSUMER DATA PRIVACY IN A NETWORKED WORLD: A FRAMEWORK FOR PROTECTING PRIVACY AND PROMOTING INNOVATION IN THE GLOBAL DIGITAL ECONOMY (2012) [hereinafter WHITE HOUSE PRIVACY REPORT], available at <http://www.whitehouse.gov/sites/default/files/privacy-final.pdf>; Jamin Melvin, “Do Not Track” Internet Spat Risks Legislative Crackdown, REUTERS (July 23, 2012, 4:08 PM EDT), <http://www.reuters.com/article/2012/07/23/us-internet-tracking-idUSBRE86M17R20120723>. The FTC’s involvement in industry-wide talks about a voluntary “Do Not Track” option may be an example of what Derek Bambauer has called “soft censorship.” See Derek E. Bambauer, *Orwell’s Armchair*, 79 U. CHI. L. REV. 863, 867 (2012).

14. Neil M. Richards, *Reconciling Data Privacy and the First Amendment*, 52 UCLA L. REV. 1149, 1169, 1173 (2005); see also Shubha Ghosh, *Informing and Reforming the Marketplace of Ideas: The Public-Private Model for Data Production and the First Amendment* (Univ. of Wis. Law Sch. Legal Studies Research Paper Series, Paper No. 1189, 2012), available at <http://ssrn.com/abstract=2000102>.

15. Ashutosh Bhagwat, *Sorrell v. IMS Health: Details, Detailing, and the Death of Privacy*, 36 VT. L. REV. 855, 856, 875-76 (2012).

16. Tim Wu, Op-Ed., *Free Speech for Computers?*, N.Y. TIMES (June 19, 2012), <http://www.nytimes.com/2012/06/20/opinion/free-speech-for-computers.html>. Of course, by this logic, books cannot be speech either because they, too, are not persons.

17. See, e.g., Agatha M. Cole, *Internet Advertising After Sorrell v. IMS Health: A Discussion on Data Privacy & the First Amendment*, 30 CARDOZO ARTS & ENT. L.J. 283, 310-12 (2012) (finding that *Sorrell* does not constrain “Do Not Track” legislation because the proposed law does not restrict “expressive uses of data”).

18. Justice Holmes, the chief architect of modern First Amendment law, cautioned the courts against using policy considerations as their gauge. A passage from his heralded dissent in *Abrams v. United States* succinctly captures the dangers of this approach: “Persecution for the expression of opinions seems to me perfectly *logical*.” 250 U.S. 616, 630 (1919) (Holmes, J., dissenting) (emphasis added).

scholars do not provide any coherent vision of the First Amendment. They do not successfully reconcile their arguments with longstanding social commitments to the free flow of information.

Meanwhile, First Amendment scholars have largely ignored how new technologies challenge the definitional questions of speech.¹⁹ Seth Kreimer has called attention to the unsound distinction between mechanically captured information and protected expression in the context of camera and video recordings,²⁰ but scholars have yet to propose a satisfying analytical framework for the variety of data collection practices that will give rise to First Amendment questions.

This Article offers such a framework in Part III by combining two important insights, each of which has a strong legacy in case law and legal scholarship. First, the Article highlights and strengthens the strands of First Amendment theory that protect the creation of knowledge. Expanded knowledge is an end goal of American speech rights, and accurate information, along with other, more subjective expressions, provides the fuel. Second, if a court is uncertain whether a regulation targets “speech,” it should analyze the regulation purposively rather than simply considering the regulation in the abstract. A motive analysis can quickly and concretely identify regulations that interfere with free expression.²¹ Together, these principles suggest that state action will trigger the First Amendment any time it purposefully interferes with the creation of knowledge.

Thus, for all practical purposes, and in every context relevant to the current debates in information law, data is speech. Privacy regulations are rarely incidental burdens to knowledge. Instead, they are deliberately designed to disrupt knowledge creation.²² Ag-gag laws and trade secret protections, too, operate on disfavored forms of increased knowledge.²³ These disruptions in knowledge

19. Fred Cate and Robert Litan began to explore the First Amendment application to personal information, but concluded their work with a series of open questions. Fred H. Cate & Robert Litan, *Constitutional Issues in Information Privacy*, 9 MICH. TELECOMM. & TECH. L. REV. 35, 57-58 (2002). Eugene Volokh warned that the rush to regulate information in the name of privacy will conflict with the First Amendment, but he starts with the presumption that data privacy laws burden speech. Eugene Volokh, *Freedom of Speech and Information Privacy: The Troubling Implications of a Right to Stop People from Speaking About You*, 52 STAN. L. REV. 1049 (2000). Volokh also limits his analysis to restrictions on the *communication* of information, and does not comment on the process of gathering data in the first place.

20. Seth F. Kreimer, *Pervasive Image Capture and the First Amendment: Memory, Discourse, and the Right to Record*, 159 U. PA. L. REV. 335 (2011).

21. Intent-based analyses are not new to First Amendment case law, though they are overlooked in the literature. I discuss cases applying motive-based analyses below in Part III.

22. See generally Daniel J. Solove, *The Virtues of Knowing Less: Justifying Privacy Protections Against Disclosure*, 53 DUKE L.J. 967 (2003).

23. Trademark protection, based on theories of dilution, can also disrupt the spread of accurate information. See Rebecca Tushnet, *Trademark Law as Commercial Speech Regulation*, 58 S.C. L. REV. 737, 738-39 (2007).

creation might sometimes, or even frequently, be justified in order to protect compelling societal interests—for instance, living without constant surveillance.²⁴ Although the First Amendment creates a barrier to the enforcement of new and existing information laws, that barrier is not insurmountable. It simply requires, as it should, a lively inquiry into whether the harms caused by the collection of information are probable enough, and serious enough, to outweigh the right to learn things.²⁵

The Article proceeds in four Parts. Part I defines “data” and “speech” so that the parameters of this investigation are clear at the outset. It also explores whether data is expressive enough to warrant First Amendment protection. Part II evaluates the sometimes contradictory case law that has analyzed whether the First Amendment applies to regulations of data and information. Part III presents the culmination of the Article’s investigation: data should receive speech protection any time it is regulated *as* information. This conclusion fits most of the cogent precedents, and it is the most defensible on normative grounds. Part IV lays out the implications of the speech rule proposed by this Article, and responds to some obvious objections. In particular, this Part responds to fears that the application of the First Amendment will lead to a parade of horrors when existing privacy laws are challenged.

This will not be a particularly smooth journey. In 1956, John Biggs, Jr., Chief Judge of the Third Circuit, noted that the state of the law relating to privacy and speech is “still that of a haystack in a hurricane.”²⁶ The state of affairs has only deteriorated with the advent of mass computing.

I. WHAT IS “DATA,” AND WHAT IS “SPEECH”?

This Part addresses two important definitional questions. Since the terms “data” and “speech” invite multiple meanings, defining them with precision here will focus and constrain the rest of the discussion.

24. Louis Brandeis, famous for his protection of speech, was also a champion for the right to be let alone—the right to have a sufficient amount of breathing space and respite from social judgment. This right was, to Brandeis, a necessary component of liberty. Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 193 (1890). I too have endorsed the right to seclusion, and believe it can and should pass constitutional scrutiny in some contexts. See Jane Yakowitz Bambauer, *The New Intrusion*, 88 NOTRE DAME L. REV. 205 (2012). However, as Burt Neuborne pointed out over twenty years ago, when the government uses information regulations to socially engineer a desired outcome, it runs the risk of “selective manipulation of information as a tool of social control.” Burt Neuborne, *The First Amendment and Government Regulation of Capital Markets*, 55 BROOK. L. REV. 5, 35 (1989).

25. Marc Blitz, motivated by the values of projects like Google Earth, proposes a similar test to protect the right to create maps. Marc Jonathan Blitz, *The Right to Map (and Avoid Being Mapped): Reconceiving First Amendment Protection for Information-Gathering in the Age of Google Earth*, 14 COLUM. SCI. & TECH. L. REV. 115, 121-24, 202 (2012).

26. *Ettore v. Philco Television Broad. Corp.*, 229 F.2d 481, 485 (3d Cir. 1956).

A. *Data*

When this Article uses the term “data,” it means a fixed record of a fact. I am about to delve more deeply into this concept, so if this simple definition is satisfactory, consider moving on to the next Subpart defining “speech.”

Semantic explanations of the word “data” require working backwards from the end goal. Data is a subset of information, and information is special because humans can transform it into new factual knowledge—a new understanding about something that has occurred. We do this by receiving and interpreting the patterns in the information. Information will not always provoke a new understanding about the state of the world from a human (just as a teacher cannot always lead a student to understanding), but it might, and the potential is key.

This Article will use “information” to mean any objective representation of something that has occurred—any representation of a fact.²⁷ It need not be fixed. A man yelling “timber” has momentarily created information that a tree is falling, though it will soon dissolve into memory and the air. It also need not be man-made, as a fossil can represent facts as clearly as any drawing or description of the original specimen.

This is an expansive definition, but it is constrained, importantly, by factual representation. Utterances sometimes convey facts (“Timber!”), but not always. Thus, when a person yells “Republicans suck” or “beauty is truth, truth beauty,” he is not representing a fact and, thus, is not uttering information about Republicans or beauty. These are opinions, ideas, hypotheses, and artistic expressions.²⁸ Information reports only something that has occurred.

The scope of this Article is narrower still. I will investigate the constitutional treatment of *data*—information that somebody has deliberately caused to be captured and recorded into a fixed, man-made format. Though I will spend much of my energy thinking about electronic data that has recorded some type of event—the time and details of a file transfer, or the time and location of a cell tower ping—data does not have to be electronic. Hand-collected tallies, cash register tapes, and photographic portraits are data. Even parts of John Keats’s *Ode on a Grecian Urn* are data to the extent they describe an actual urn.

27. This raises many worthwhile questions about how to distinguish between objectivity and subjectivity, and more fundamentally, whether objectivity is possible. This Article does not engage these questions. The definition used here is in line with systems theory’s use of the term, but it is not how the term is used in information theory. Computer scientists use a definition of “information” that is both broader and narrower than this Article’s. In information theory, information is a sequence of signals that can be interpreted as a message of any sort (not exclusively factual). JOHN R. PIERCE, AN INTRODUCTION TO INFORMATION THEORY: SYMBOLS, SIGNALS, AND NOISE 8 (2d rev. ed. 1980).

28. A recording of the person who yells it, however, would be a factual representation of what was said, by whom, and how.

This Article confines its investigation to data for two reasons. First, man-made data has more similarities to traditional speaker-listener arrangements than other types of information (though there are some limits to the similarities), so it might have a more promising claim to the First Amendment's protections. Second, unlike other forms of information, the quantity of data has exploded in the digital age.²⁹ The sudden change in the status quo makes data a likely candidate for policy debates and regulation. In other words, this is the information that governments worry about.

This Article will also refer to "records" (as in "business records" or "medical records"). A record is data created with the specific purpose of preserving information. Thus, it can be distinguished from "functional data," such as server data, which is produced in order to carry out a function or service.³⁰ Functional data is data too; it came to be because browser programs or other software were deliberately coded to create the data. However, it is conceptually distinct from a record because it was not necessarily created for the purpose of preserving information. Machine-to-machine communications were originally produced to carry out the steps necessary for web browsing. Though the data turned out to be very valuable for its information content, it was not necessarily created with that purpose in mind. So, records and functional data are distinct subsets of data. Some scholars, myself included, have homed in on the distinction between functional data and other types of records as a useful boundary for law.³¹ But as this Article explains, First Amendment doctrine should not capitalize on this distinction.

Finally, for the purposes of this Article, data will be presumed to be accurate. Data that turns out to be inaccurate, either through deliberate attempts at deceit or unintentional error, presents some interesting First Amendment and legal liability questions, but the nature of those questions is very different. In the great conflict between privacy regulation and information flow, accuracy is the source of both harm and value.

B. *Speech*

Equally important is the definition of "speech." The goal of this Article is to understand whether regulations on the collection or transfer of data implicate the First Amendment, thus requiring the government to justify and narrowly tailor the regulations. It is not particularly important whether First Amendment

29. See *The Data Deluge*, ECONOMIST (Feb. 25, 2010), <http://www.economist.com/node/15579717>.

30. Server data is the series of messages sent between an end user's computer or device and a website's servers. These messages instruct the server to send the website's code and web objects back to the end user's machine, so that the machine can build the website for the end user to view. Marc Cohen, *How the Web Works—in One Easy Lesson*, MARC'S SPACE (Mar. 13, 2011), <http://mkcohen.com/how-the-web-works-in-one-easy-lesson>.

31. Bambauer, *supra* note 24, at 239-40.

interests are raised under the banner of “speaker rights,” “listener rights,” or under some other concept, such as the semi-developed “right to receive information.”³² Ultimately the Article seeks to understand whether legal restraints on the creation or dissemination of data draw First Amendment scrutiny on any basis.

Because some forms of data are already very familiar forms of expression—maps, almanacs, photographs, and the like—it is quite natural to refer to them as “speech.” So this Article will use the term “speech” as convenient shorthand for anything protected by the First Amendment (under the Free Speech Clause) on any basis. For our purposes, if data is constitutionally protected at all, it is “speech.”

Another caveat: this Article will assume that anything found within the scope of speech will receive some accompanying judicial scrutiny so long as it does not fall within one of the exceptional categories of “unprotected” speech—such as fighting words, obscenity, fraud, and so forth.³³ But it would be careless to proceed with this assumption without acknowledging that it is a subject of scholarly debate.

Though courts resolving speech issues routinely assert that the First Amendment applies to speech except for the exceptional categories of “unprotected” speech, Frederick Schauer has questioned the plausibility of this pro forma rhetoric. Schauer notes that First Amendment concerns are only raised when courts notice certain familiar issues—political speech or insults, for example. Meanwhile, the great majority of laws regulating utterances and communications do not draw any constitutional scrutiny at all.³⁴ Schauer points to antitrust and sexual harassment laws, among others, to illustrate his point.³⁵ Neil Richards echoes Schauer’s reasoning to argue that consumer data privacy laws do not trigger scrutiny even if they do regulate information—a form of speech.³⁶

32. Marc Jonathan Blitz, *Constitutional Safeguards for Silent Experiments in Living: Libraries, the Right to Read, and a First Amendment Theory for an Unaccompanied Right to Receive Information*, 74 UMKC L. REV. 799 (2006) (promoting an independent First Amendment right to receive information); Christina E. Wells, *Reinvigorating Autonomy: Freedom and Responsibility in the Supreme Court’s First Amendment Jurisprudence*, 32 HARV. C.R.-C.L. L. REV. 159, 172 (1997) (identifying distinct speaker-centric and listener-centric lines of cases).

33. See Erwin Chemerinsky, *Content Neutrality as a Central Problem of Freedom of Speech: Problems in the Supreme Court’s Application*, 74 S. CAL. L. REV. 49, 61-62 (2000).

34. Frederick Schauer, *The Boundaries of the First Amendment: A Preliminary Exploration of Constitutional Salience*, 117 HARV. L. REV. 1765, 1768 (2004).

35. These seeming exceptions can be understood, though, as statutory law forbidding certain conduct—for example, conditioning employment on gender or colluding on the price of a good. Though this conduct is inevitably carried out through communications, the communications merely evidence the decision or action that Congress seeks to forbid.

36. Richards, *supra* note 14, at 1173-74. Robert Post, too, believes the scope of speech protection does not go much further than the protection of speech of “public concern.”

Granted, the Supreme Court's choice to distinguish not only between speech and conduct (which, while difficult, is necessary), but also between protected and unprotected speech, breeds some uncertainty about whether classifying something as "speech" automatically means that it receives judicial protection.³⁷ Since courts are inconsistent in defining the boundary between protected and unprotected speech, Schauer and Richards are justified in casting doubt on whether First Amendment scrutiny applies by default.

However, many of the examples Schauer cites as evidence that First Amendment protection is exceptional, even for traditional forms of expression like utterances, fail to support his theory.³⁸ For example, Schauer makes much of the fact that antitrust claims may be brought against firms based on the speech used to reach price-fixing agreements,³⁹ but these laws target and prohibit the business agreement to collude on pricing. Enforcement uses the actual words uttered by the firms' directors or agents only as evidence of an agree-

ROBERT C. POST, *DEMOCRACY, EXPERTISE, AND ACADEMIC FREEDOM: A FIRST AMENDMENT JURISPRUDENCE FOR THE MODERN STATE* 11 (2012). However, to make sense of cases that unambiguously protect speech that is not political, or of "public concern" (which itself eludes definition), Post is forced to identify other First Amendment doctrines, such as a First Amendment interest in "competence." *Id.* at 25, 35, 85. These epicycles could be avoided if Post abandoned his commitment to the centrality of the "public concern" distinction.

37. Defamatory speech is an apt example. In some cases, the Court insists that there is no constitutional value in false, defamatory statements, and that such statements receive protection only incidentally to protect other speech. *See, e.g., Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 340-41 (1974); *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 268 (1964) (citing and accepting earlier precedent). At other times, the Supreme Court suggests that defamatory speech does receive direct protection, albeit relaxed protection in light of the low value of the speech. *N.Y. Times*, 376 U.S. at 273 ("If neither factual error nor defamatory content suffices to remove the constitutional shield from criticism of official conduct, the combination of the two elements is no less inadequate."). In *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749 (1985), the members of a divided Court endorsed both interpretations in their fractured opinions. *Compare id.* at 769-70 (White, J., concurring in the judgment) (suggesting that defamatory speech receives only derivative protection), *with id.* at 775-76 (Brennan, J., dissenting) (suggesting that defamatory speech receives direct protection).

38. Robert Post makes a similar argument that speech protection does not usually apply to regulations of speech. He supports the claim using the example of an American Dental Association rule that prohibits dentists from telling their patients their opinions about the dangers of dental amalgams and from persuading their patients to have their amalgam fillings removed. POST, *supra* note 36, at 12. Post does not reflect on the fact that the rule is promulgated by a licensing board whose functional purpose is to regulate the practice of dentistry. A dentist's ability to practice dentistry is tied to certain acts or omissions that are required to comply with licensing privileges. *See, e.g., Planned Parenthood of Se. Pa. v. Casey*, 505 U.S. 833, 884 (1992) (upholding speech restrictions on abortion providers on the basis of reasonable licensing requirements). The other way to look at this, of course, is that Post labors to make sense of a speech restriction that doesn't deserve to be saved. This may be an inappropriate and unconstitutional condition on a government licensing benefit.

39. Schauer, *supra* note 34, at 1781, 1805-06.

ment—a contract—to do something that violates public policy (in this case, to sell goods at noncompetitive prices).⁴⁰

Contrary to Schauer’s thesis, intentional regulations of speech as speech have, as a default, triggered First Amendment scrutiny. Recently, the Supreme Court has been on a mission to apply the First Amendment broadly, and to interpret its exceptions narrowly. Cases from the last three years have scrutinized the regulation of violent videogames,⁴¹ animal torture videos,⁴² behavioral marketing,⁴³ and even intentional lies.⁴⁴ Chief Justice Roberts has promised that “[e]ven [w]holly neutral futilities” will receive the same First Amendment protection as Keats’s poems.⁴⁵

One could argue that the categories of “unprotected” speech might better be understood as speech for which a strict scrutiny calculus that has already, tacitly, been performed. Elena Kagan’s account of the First Amendment describes the unprotected speech classification as an exception that applies to “unusually trustworthy kinds of content-based restrictions.”⁴⁶ By “trustworthy,” Kagan means that the legislature’s purported motives and their true motives are the same, and are indeed compelling.⁴⁷ Child pornography is “unprotected” speech, but if it were protected speech, regulations would probably survive strict scrutiny because of the compelling state interests involved. The same could be said for fraud.

Finally, this Article addresses only the issue of First Amendment coverage. It does not attempt to determine what level of scrutiny ought to apply to the regulation of data.⁴⁸ The first-order analysis of whether the First Amendment

40. The Sherman Antitrust Act defines criminal behavior in terms of contract. *See* 15 U.S.C. § 1 (2012). The enforcement of hostile work environment sexual harassment claims does implicate, as Schauer says, the regulation of speech, *see* Schauer, *supra* note 34, at 1805, but it is also not clear that these claims would fail to draw First Amendment scrutiny. Eugene Volokh has nicely cataloged and critiqued the scenarios in which speech is regulated as part of a course of action or as evidence of a particular arrangement. *See* Eugene Volokh, *Speech as Conduct: Generally Applicable Laws, Illegal Courses of Conduct, “Situation-Altering Utterances,” and the Uncharted Zones*, 90 CORNELL L. REV. 1277 (2005).

41. *Brown v. Entm’t Merchs. Ass’n*, 131 S. Ct. 2729 (2011).

42. *United States v. Stevens*, 559 U.S. 460 (2010).

43. *Sorrell v. IMS Health Inc.*, 131 S. Ct. 2653 (2011).

44. *United States v. Alvarez*, 132 S. Ct. 2537 (2012).

45. *Stevens*, 559 U.S. at 479 (second alteration in original) (quoting *Cohen v. California*, 403 U.S. 15, 25 (1971)) (internal quotation mark omitted).

46. Elena Kagan, *Private Speech, Public Purpose: The Role of Governmental Motive in First Amendment Doctrine*, 63 U. CHI. L. REV. 413, 414 (1996).

47. *Id.* at 415, 481.

48. However, I do comment on this briefly at the end of Part III. I do not comment, except for here, on the many ways that expression can be affected by the government’s *failure* to pass laws promoting and encouraging speech. Because the First Amendment provides a negative right to the governed, not a positive one, the state’s omission will not be a violation of the First Amendment even though the effects could be dramatic. To appreciate

applies *at all* is complex enough to merit an article of its own, and so I direct my focus there.⁴⁹

No account of the First Amendment treatment of data, whether normative or descriptive, can be complete without some reflection on the qualities of data—qualities that make it at once similar to and very different from other more familiar forms of expression. So before moving on to the discussion about what the First Amendment does (Part II) and what it should do (Part III), this Part will conclude with some preliminary observations about the nature of data.

Data communicates. It tells a narrative just as effectively as prose, imagery, and music to those with the training to interpret it. Its style is dry, but this does not interfere with its ability to light up the mind. A database can be interpreted directly by a person with the help of a codebook, and it can also be translated into other more familiar forms of expression like maps, charts, graphs, and descriptive sentences. Lest there be any doubt about data's intimate connection to other forms of expression, one may recall that the very first form of writing was data: the accounting records of traders in ancient Mesopotamia.⁵⁰ Data provided the building blocks of the rest of written language.

Modern data collection practices raise uniquely difficult constitutional line-drawing problems. When data is created to inform a broad public audience, or even to inform a specific audience, its conceptual similarities to pamphlets and other types of traditional speech are obvious. But when it is created without the intent to inform anybody of anything in particular, the parallels to traditional speech begin to break down.

To understand whether this type of data can or should receive the protection of the First Amendment, I will next explore how real conflicts have been resolved by the courts.

II. DATA IN FIRST AMENDMENT PRECEDENT

The descriptive analysis of the First Amendment case law as it relates to data runs into a contradiction. When data already exists and is transmitted from one person or entity to another, the case law strongly supports the conclusion that raw facts are protected by the First Amendment. These cases are described in Subpart A.

One might assume that the creation of data tends to receive protection as well, since the collection of data is a necessary precursor to having and sharing it. Indeed, some cases suggest this is so, but at present the authorities are in

the role that speech-facilitating statutes can play, see Jack M. Balkin, *The Future of Free Expression in a Digital Age*, 36 PEPP. L. REV. 427 (2009).

49. Frederick Schauer and Robert Post would describe this as an inquiry about “coverage,” as distinguished from “protection.” POST, *supra* note 36, at 1; Schauer, *supra* note 34, at 1769.

50. MATT RIDLEY, *THE RATIONAL OPTIMIST: HOW PROSPERITY EVOLVES* 160 (2010).

conflict. A long strand of privacy case law confidently concludes that data creation is conduct insufficiently expressive to warrant First Amendment protection. These cases are described in Subpart B.

Subpart C discusses some recent cases challenging and rejecting the assumption that mechanical information gathering is nonexpressive conduct.

To prepare for the normative discussion in Part III, Subpart C will also give some attention to the logical strengths and weaknesses of each of the diverging trends in the case law. The cases finding that mechanical information gathering does not receive speech protections reach their holdings after making assumptions about technologies that have not aged well. In contrast, the cases recognizing First Amendment protection of data gathering are more consistent with free speech values and are better equipped to stand the test of time.

A. Existing Data

A few years ago, privacy advocates and data aggregators watched anxiously as the Supreme Court decided *Sorrell v. IMS Health Inc.*⁵¹ Justice Kennedy's opinion very nearly resolved whether data is protected speech. Big Data, held by Big Pharma, was at the center of the case. The State of Vermont passed legislation prohibiting pharmaceutical companies from receiving and using prescription data to customize their advertising to doctors. IMS Health, a data aggregator, brought a First Amendment challenge.⁵² Five Justices joined in Justice Kennedy's majority opinion finding the law unconstitutional. Part of this opinion suggested that the restriction on transfers of data between willing givers and receivers was automatically a restriction of speech. Data consists of facts, after all, and facts are "the beginning point for much of the speech that is most essential to advance human knowledge and to conduct human affairs."⁵³ According to Justice Kennedy, there was a "strong argument" that First Amendment scrutiny applied on that basis alone.⁵⁴ But in the end, Justice Kennedy opted to resolve the case on narrower grounds.⁵⁵

Why couldn't Justice Kennedy pull the trigger and declare that data merits speech protection?⁵⁶ He no doubt had misgivings about the broad and unanticipated consequences that such a declaration might bring about.

51. 131 S. Ct. 2653 (2011).

52. *Id.* at 2660-62.

53. *Id.* at 2667.

54. *Id.*

55. *Id.* at 2663-64 (relying on viewpoint discrimination).

56. This is not the first time Justice Kennedy exhibited some misgivings about granting databases complete constitutional protection. A decade earlier, Justice Kennedy wanted to grant certiorari in a case that would have allowed the Court to consider whether the FCRA violated the First Amendment. Justice Kennedy hoped the Court would consider whether accumulations of consumer data constituted speech. *Trans Union LLC v. Fed. Trade Comm'n*, 536 U.S. 915, 916 (2002) (Kennedy, J., dissenting from denial of certiorari).

Justice Kennedy's reservations notwithstanding, lower courts have shown they are quite prepared to apply the First Amendment to already-existing raw information that one individual or entity wishes to share with another. Prior to *IMS Health*, the cases addressing this issue had insisted that raw information is speech without hesitation. Occasionally they had done much more, recommending that free speech rights protect uninhibited analysis and access to information, along with the freedom to disseminate it.

This Subpart catalogs the cases in three areas of law where the constitutional protection of data sharing has naturally had reason to surface: privacy, advertising, and copyright. The results across each of these disciplines consistently protect the free communication of data.

1. *Gleanings from privacy law*

The Court wrote on a fairly clean slate when it decided *IMS Health*. Very few cases have raised First Amendment challenges to data privacy statutes. But in the challenges that have been brought, courts have concluded unequivocally that the communication of raw data is speech.

In *Trans Union Corp. v. Federal Trade Commission*, a credit reporting agency challenged the constitutionality of the Fair Credit Reporting Act (FCRA), which forbids companies from sharing consumer credit reports except for specified purposes.⁵⁷ The D.C. Circuit never doubted that the consumer reports were speech. Relying on earlier Supreme Court precedent, the court found that consumer reports received protection but warranted more lenient intermediate scrutiny as commercial speech.⁵⁸ Thus, at least in the view of the D.C. Circuit, existing business records were entitled to First Amendment protection of some form.

The Tenth Circuit case of *U.S. West, Inc. v. FCC* goes further still.⁵⁹ *U.S. West* also addressed a First Amendment challenge to a privacy regulation—section 222 of the Telecommunications Act of 1996,⁶⁰ which restricts telecommunications providers from disclosing or using customer data except for limited purposes. The telecommunications data at issue in *U.S. West* included any “information that relates to the quantity, technical configuration, type, destination, and amount of use of a telecommunications service subscribed to by any customer.”⁶¹ This definition is expansive. It includes not only the sorts of

57. 245 F.3d 809 (D.C. Cir. 2001), *cert. denied*, 536 U.S. 915.

58. *Id.* at 818 (citing *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749 (1985)). In *Trans Union*, the “credit reports” were actually marketing lists that were created in part based on information that would be available only to creditors and credit reporting agencies. This type of document falls within the FCRA’s definition of “consumer report.” *Id.* at 812-13.

59. 182 F.3d 1224 (10th Cir. 1999).

60. 47 U.S.C. § 222 (2011).

61. *See id.* at 1228 n.1 (quoting 47 U.S.C. § 222(f)(1)(A)-(B)).

information that the telecommunications provider records in its bills and other business records (though billing information is certainly covered⁶²), but also data describing when, where, and to whom a customer places a call. This is information generated in the process, and for the purpose, of providing telecommunications service. It is, in other words, *functional data*.⁶³

The court found that section 222 infringed on the telecommunications service provider's First Amendment rights.⁶⁴ But it did not do so by concluding that the customer data itself was speech. Instead, the court located the First Amendment interest in the communications that U.S. West intended to make to its customers with the *aid* of the data. The First Amendment attached to the advertising U.S. West wished to tailor to their customers' usage habits. So technically, the court never identified the data as a source of First Amendment rights. Instead, the court focused its analysis on the indirect burden imposed on U.S. West's traditional commercial speech.⁶⁵

But if the ultimate question is whether data receives First Amendment protection, the distinction is inconsequential. If section 222 infringed on U.S. West's right to advertise to its customers, it did so only indirectly through its effect on data. The fact that the court treated the indirect burden as a First Amendment problem reveals that data has a sufficiently close relationship to free expression to merit derivative protection.

Indirect burdens on speech provoke First Amendment scrutiny only when the intended target of the regulation is so closely linked to speech, so indispensable to First Amendment rights, that it is an impediment to speech itself. A tax on ink, for example, would draw First Amendment scrutiny.⁶⁶ A tax on metal wouldn't, even though metal is used to make printing presses. And speed limits wouldn't, even though newspaper delivery could be speedier if the law did not apply.⁶⁷ The FCC regulation did not restrict U.S. West from making any particular communication to its customers. Instead, it deprived U.S. West of a resource the company would have liked to use to tailor the messaging. By choosing to analyze the privacy regulations under the First Amendment, the *U.S. West* court tacitly endorsed an understanding of the First Amendment that treats

62. *See id.*

63. *See supra* notes 30-31 and accompanying text.

64. *U.S. West, Inc.*, 182 F.3d at 1239-40.

65. *See id.* at 1232.

66. *See Minneapolis Star & Tribune Co. v. Minn. Comm'r of Revenue*, 460 U.S. 575, 582-83 (1983); *see also Fla. Bar v. Went For It, Inc.*, 515 U.S. 618 (1995) (finding that a restriction on the distribution of attorney solicitations via mail is a burden on speech, but that it withstands scrutiny). *But see Young v. Am. Mini Theatres, Inc.*, 427 U.S. 50 (1976) (finding that zoning laws requiring that adult motion picture theaters maintain a certain distance between themselves only incidentally burdened speech).

67. *See Frederick Schauer, Cuban Cigars, Cuban Books, and the Problem of Incidental Restrictions on Communications*, 26 WM. & MARY L. REV. 779 (1985).

accurate information (to which the potential speaker otherwise has access) as an element, or fundamental tool, of speech.

Alternatively, the *U.S. West* court may have believed that a privacy regulation like section 222 could never be an *incidental* burden to speech because the purposes underlying the regulation are aimed at achieving some desired effect on future communications.⁶⁸ Either way, the court recognized a constitutionally significant connection between readily accessible facts and speech.

2. *Gleanings from commercial speech regulations*

Prescription drug prices and beer ingredients are truly dull facts, but their communication from a business to its customers is protected speech.⁶⁹

The Supreme Court's opinion in *Virginia State Board of Pharmacy* has the most to say about the constitutional protection of information in commercial settings. The Court found that pharmacies had a constitutional right to advertise the prices of prescription drugs.⁷⁰ This holding readily supports a conclusion that data is speech when it is distributed. Even if a person's decision to transmit information is driven by economic self-interest, "[t]hat hardly disqualifies him from protection under the First Amendment."⁷¹

The *Virginia State Board of Pharmacy* Court also recognized a "strong interest in the free flow of commercial information"—a constitutional interest that attaches to the *recipient* of information rather than the transmitter.⁷² This right to receive information flows naturally from a conception of "speech" that aims to protect freedom of thought as an ingredient of autonomy. To illustrate this, the Court praised an Illinois case protecting a manufacturer's right to advertise the origin of a product so that its consumers could access the information, and use it to conform their purchases to their antipathy for the outsourcing of American manufacturing jobs.⁷³ The Court also pointed to an

68. Restrictions on data might, thus, automatically fail the "unrelated[ness]" rule set out in *Buckley v. Valeo*, 424 U.S. 1, 18 (1976).

69. See *Rubin v. Coors Brewing Co.*, 514 U.S. 476, 478 (1995) (holding that a federal law prohibiting beer labels from displaying alcohol content violated the First Amendment); *Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, Inc.*, 425 U.S. 748, 756-57 (1976) (holding that a state law prohibiting pharmacies from advertising prescription drug prices violated the First Amendment).

70. *Va. State Bd. of Pharmacy*, 425 U.S. at 756-57. Note that the Court applied intermediate scrutiny since the advertisements consisted solely of commercial speech. See *id.* at 770-71.

71. *Id.* at 762. The reporting of research findings is also given the protection of speech. See *Bd. of Trs. of Leland Stanford Junior Univ. v. Sullivan*, 773 F. Supp. 472, 479 (D.D.C. 1991).

72. *Va. State Bd. of Pharmacy*, 425 U.S. at 764.

73. *Id.* (citing *Chi. Joint Bd. v. Chi. Tribune Co.*, 435 F.2d 470 (7th Cir. 1970)).

earlier Supreme Court case finding First Amendment protection for statements identifying where legal abortions were performed.⁷⁴

Furthermore, the Court not only recognized the public's interest in accessing information, but it also found that the right to receive information exists in the abstract, even if the particular communications that happen to be made using the information are not especially illuminating. Because information *can* be edifying, the First Amendment will protect that potential regardless of how the information is ultimately used. The Court used a hypothetical pharmacist to illustrate the expansive protection: "Our pharmacist, for example, could cast himself as a commentator on store-to-store disparities in drug prices, giving his own and those of a competitor as proof. We see little point in requiring him to do so, and little difference if he does not."⁷⁵

Virginia State Board of Pharmacy was not the first time that the Court recognized a First Amendment right to receive information,⁷⁶ but this case marks its strongest pronouncement. Moreover, the "right to receive" has been expanded by some courts to a right to access information, which protects individuals who pluck the information directly from the world, rather than relying on the communications of another speaker. This expansion covers the right to create original data, as opposed to receiving data already in existence. These cases are explored in Subpart C below.

3. *Gleanings from copyright law*

Facts cannot be copyrighted.⁷⁷ This is so first and foremost because the Intellectual Property Clause of the U.S. Constitution allows Congress to create copyrights only in fixed works that are *original*. A fact is not created by anybody. "The first person to find and report a particular fact has not created the fact; he or she has merely discovered its existence."⁷⁸

It is conceivable that a raw fact could be outside the scope of the Intellectual Property Clause and also unprotected by the First Amendment, but it would be odd. If this were the case, the government would be powerless to grant *one* person exclusive control of a fact, yet it would have the power to deny *everybody* access to and use of it.

Instead, the treatment of facts under copyright law promotes an interpretation of the First Amendment that protects data. In *Feist*, when the Supreme

74. *Id.* at 759-60 (citing *Bigelow v. Virginia*, 421 U.S. 809 (1975)).

75. *Id.* at 764-65.

76. *Martin v. City of Struthers*, 319 U.S. 141, 143 (1943) (finding that the First Amendment protects the right to receive literature).

77. See Michael C. McFarland, *Intellectual Property, Information, and the Common Good*, B.C. INTELL. PROP. & TECH. F. (June 5, 1999), http://www.bc.edu/bc_org/avp/law/st_org/iptf/resources/index.html.

78. *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 347 (1991).

Court determined that compilations of facts cannot be copyrighted, one of the reasons was that they “are part of the public domain available to every person.”⁷⁹ Similarly, in *Harper & Row, Publishers, Inc. v. Nation Enterprises*, the Court articulated its strongest statement about the interaction between the First Amendment and copyright protection when it explained that the “idea/expression dichotomy strike[s] a definitional balance between the First Amendment and the Copyright Act by permitting free communication of facts while still protecting an author’s expression.”⁸⁰ It follows that the unobstructed communication of facts is a First Amendment priority.⁸¹

More recently, a lawsuit brought under the Digital Millennium Copyright Act (DMCA) happened to present the Second Circuit with the issue of whether computer code is speech. The DMCA prohibits the distribution of technologies that can be used to circumvent the encryption used to limit access to copyrighted files.⁸² In *Universal City Studios, Inc. v. Corley*, a journalist challenged the constitutionality of the law in a civil enforcement action.⁸³ The journalist had published source code on a hacking enthusiast’s website. Beginning its analysis, the court said “[c]ommunication does not lose constitutional protection as ‘speech’ simply because it is expressed in the language of computer code.”⁸⁴ However, because source code simultaneously has qualities of expressive description *and* executable computer commands, the court reasoned that computer code’s status as speech depends on the manner and purpose for which it is distributed. If the code is shared to inform readers about *how* to decrypt copyright-

79. *Id.* at 348 (quoting *Miller v. Universal City Studios, Inc.*, 650 F.2d 1365, 1369 (5th Cir. 1981)) (internal quotation mark omitted).

80. 471 U.S. 539, 556 (1985) (alteration in original) (quoting *Harper & Row, Publishers, Inc. v. Nation Enters.*, 723 F.2d 195, 203 (2d Cir. 1983)) (internal quotation marks omitted). More recently, though less forcefully, the Court noted the First Amendment limitations on the government’s attempt to restrict the communication of facts in *Eldred v. Ashcroft*, 537 U.S. 186, 219-20 (2003). Scholars, too, have argued that the First Amendment is at least partly responsible for defining the boundaries of copyright. See Yochai Benkler, *Free as the Air to Common Use: First Amendment Constraints on Enclosure of the Public Domain*, 74 N.Y.U. L. REV. 354, 440 (1999); Alan E. Garfield, *The Case for First Amendment Limits on Copyright Law*, 35 HOFSTRA L. REV. 1169, 1173 (2007).

81. Other case outcomes are consistent with this understanding. The court in *Barclays Capital Inc. v. Theflyonthewall.com, Inc.*, 650 F.3d 876 (2d Cir. 2011), for example, found that the Copyright Act preempted the plaintiff’s claim for hot news appropriation in part to advance First Amendment values that ensure no one person will have control over the knowledge of newsworthy events. But note that it protected the company’s right to “break” the news it had gathered and was silent as to the First Amendment protection of the gathering itself. *See id.* at 902-03, 907.

82. 17 U.S.C. §§ 1201(a)(2), (b)(1) (2012).

83. 273 F.3d 429, 435-36 (2d Cir. 2001).

84. *Id.* at 445. This may satisfy residual doubts that raw data and web logs are expressive. The court went on to state: “Mathematical formulae and musical scores are written in ‘code,’ *i.e.*, symbolic notations not comprehensible to the uninitiated, and yet both are covered by the First Amendment.” *Id.*

ed files, then the code will be treated as protected information distribution. If it is shared with the purpose, and the instruction, to be installed and run on a computer, then the distribution can be treated as conduct.⁸⁵ The former scenario is akin to distributing the description for making a bomb, while the latter is akin to distributing an actual bomb.⁸⁶

This outcome seems bizarre at first blush. The difference between sharing code for informational purposes and for conduct purposes is a matter of mental state alone. As a practical matter, this standard will be very difficult to apply. But, as a theoretical matter, it is eminently sensible. When a government regulates bombs, it constrains what a person can *have* and *do*. When a government regulates the dissemination of instructions for making a bomb, it constrains what a person can *know*.

The protection of potentially useful knowledge thematically connects most of the cases described above, in which courts grappled with the First Amendment's application to existing data. Though none of them define speech to include a right to uninhibited curiosity and knowledge creation, neither do they worry unduly about whether transmission of raw information is expressive enough to draw parallels to other, more traditional forms of speech. The holdings have begun to create a mosaic illustrating that the right to speech requires, and assumes, a right to learn new things. Thus, the cases addressing the dissemination of data align well with the thinker-centered First Amendment promoted in this Article.

However, the next set of cases, addressing the creation of new data, is in obvious tension with the right to create knowledge.

B. *Data Creation as Nonexpressive Conduct*

Mr. Dietemann was a quack. He occasionally invited people into his home office and for a modest fee, or sometimes for no fee at all, he would use gadgets, clay, and minerals to treat their medical ailments.⁸⁷

In 1963, *Life Magazine* published an article titled *Crackdown on Quackery*, which detailed the persistence of witch doctors in the age of modern medicine. As part of the story, and in collaboration with the District Attorney's Office of Los Angeles County, two *Life Magazine* reporters arranged to have an

85. *Id.* at 448-49 (contrasting the facts before it with *Commodity Futures Trading Commission v. Vartuli*, 228 F.3d 94, 111 (2d Cir. 2000), which found that code combined with instructions and encouragement to run the program was regulable nonspeech). Criminal prohibitions of tax shelters draw a similar distinction between political advocacy and encouragement to commit unlawful behavior. See *United States v. Raymond*, 228 F.3d 804 (7th Cir. 2000), *overruled on other grounds by Hill v. Tangherlini*, 724 F.3d 965 (7th Cir. 2013).

86. However, the difference between these two is razor thin—as thin as a difference in intent.

87. *Dietemann v. Time, Inc.*, 449 F.2d 245, 245 (9th Cir. 1971).

appointment with Mr. Dietemann under the pretense of having legitimate interest in his services. The reporters carried a hidden radio transmitter and a hidden camera, with which they surreptitiously recorded and photographed Mr. Dietemann's examination. The published article depicted Mr. Dietemann's practice, sparing no comical, humiliating detail. It included a photograph of Mr. Dietemann grabbing one of the reporters' breasts and waving a wand.⁸⁸

Mr. Dietemann sued Time, Inc. (the parent company of *Life Magazine*) for invasion of privacy.⁸⁹ Despite the journalists' deceit in inducing Mr. Dietemann to invite them into his office, the Ninth Circuit decided that the vivid written account of everything that occurred that night received full speech protection.⁹⁰ If Mr. Dietemann's claim had been based on the reporters' use of memory and words to reconstruct the events, he would not have been able to overcome the magazine's First Amendment defense.⁹¹

However, the court found that the First Amendment did not extend to the reporters' use of mechanical recording devices. It reasoned that, while the magazine had a First Amendment right in the information it observed (and, thus, a right to report it in traditional news media), the First Amendment does not immunize information-gathering *conduct*, which can be regulated by the state like any other conduct. "The First Amendment has never been construed to accord newsmen immunity from torts or crimes committed during the course of newsgathering. The First Amendment is not a license to trespass, to steal, or to intrude by electronic means . . ."⁹² In the end, the acts of mechanically recording and photographing Mr. Dietemann resulted in liability for Time, Inc. (though the publication of the images did not).⁹³

Distinguishing speech from conduct is the First Amendment's Sisyphean task. There is no set of properties that can reliably identify which human activities are expressive enough to be speech and which are not.⁹⁴ Surely information-gathering practices can fall squarely on the conduct side of the speech/conduct divide. If the reporters had never been invited into Mr. Dietemann's

88. *Id.* at 245-46.

89. *Id.* at 245.

90. *Id.* at 249 ("He invited two of defendant's employees to the den. One who invites another to his home or office takes a risk that the visitor may not be what he seems, and that the visitor may repeat all he hears and observes when he leaves.").

91. *Id.*

92. *Id.* (emphasis added).

93. *Id.* at 250.

94. As Post has noted:

[I]t is not possible constitutionally to distinguish speech from action on the ground that the former communicates ideas or uses language. The implication of this conclusion is quite significant, for it suggests that speech cannot be distinguished from action because of some common property that "speech" possesses but that "action" does not.

POST, *supra* note 36, at 3-4; see also Geoffrey R. Stone, *Flag Burning and the Constitution*, 75 IOWA L. REV. 111, 114-15 (1989) (highlighting this difficulty in the context of flag burning).

house, physically overtaking Mr. Dietemann and barging inside his house could not be immunized from the laws of trespass and battery, even if the team reasonably believed that newsworthy information was likely to emerge from the disruption. Likewise, holding up a postman in order to gather interesting information would be the same bad conduct as holding up a postman for any other reason. But the *Dietemann* case has an unresolved tension in its facts. The reporters *were* invited into Mr. Dietemann's home and unquestionably retained their First Amendment right to report everything they saw and heard in as much detail as they could.⁹⁵ What makes information collected using "electronic means" so special?

The harm the *Dietemann* court sought to avoid had nothing to do with conduct. The electronic devices did not harm Mr. Dietemann or his property at the time of recording. The harm had everything to do with content—specifically, the accuracy, credibility, and rich detail in mechanically recorded information that far exceeded the capabilities of human memory. The court decided that Mr. Dietemann did not "take the risk that what is heard and seen will be transmitted . . . in full living color and hi-fi to the public at large or to any segment of it that the visitor may select."⁹⁶ In other words, information gathering is protected by the First Amendment only up to a point: only when it is constrained by the natural limitations of the human newsgatherer (perhaps aided by pens and other older technologies).⁹⁷ The rule emerging from *Dietemann*, which has been sanctioned and followed by the Supreme Court, is that the disclosure of mechanical information is speech, while the recording of that mechanical information is not speech.⁹⁸

Given that the cases summarized in the last Subpart demonstrate that already-existing data transmitted between two people is speech, the *Dietemann* rule is anomalous: data is speech, but the creation of it is not. An equivalent rule for other forms of speech would immediately draw suspicion: Selling books is protected speech; could printing them possibly be regulable conduct?

95. This unrestrained right of the journalists to report what they had observed may have its roots in *Sidis v. F-R Pub. Corp.*, 113 F.2d 806 (2d Cir. 1940), a curious case that embedded a right to report information surreptitiously collected by undercover journalists into the American common law without explicitly referencing the First Amendment. See Samantha Barbas, *The Sidis Case and the Origins of Modern Privacy Law*, 36 COLUM. J.L. & ARTS 21 (2012).

96. *Dietemann*, 449 F.2d at 249.

97. Recording and communications technologies, starting with spoken language itself, have become critical to the human experience because they enhance what Tyler Burge calls "substantive content memory." Tyler Burge, *Memory and Persons*, 112 PHIL. REV. 289, 289 (2003).

98. *Bartnicki v. Vopper*, 532 U.S. 514, 526-31 (2001) (finding that a radio station's broadcast of newsworthy recordings that it knew were recorded via an illegal wiretap received First Amendment protection, while suggesting that the person who made the illegal recordings could be prosecuted without the same scrutiny).

Music is speech, but could it be that recording and producing an album is conduct?⁹⁹

On the other hand, perhaps the *Dietemann* court simply recognized that data has a unique place in the First Amendment sphere because of its potential to convert everything that has ever happened into protected speech, with little distortion. Mechanical recordings may have seemed out of range for the rights traditionally associated with speech. Moreover, the interests of Mr. Dietemann that the court sought to protect are quite strong: we all would hope that our homes, at the very least, can shelter us from public observation.¹⁰⁰

Putting aside for a moment the wisdom of the *Dietemann* rule differentiating information from information gathering, lawmakers have certainly depended on it. Many states have all-party-consent wiretap statutes, which criminalize the recording of conversations unless each and every party to the conversation has given consent.¹⁰¹ Retailers in California may not ask for or record a consumer's zip code.¹⁰² Privacy reports produced by the White House and by the Federal Trade Commission urge Congress to pass new laws prohibiting the unnecessary creation of personal data and requiring the deletion of data after a period of time.¹⁰³ "Do Not Track" legislation would be expected to put constraints on the collection or retention of certain types of data, too.¹⁰⁴ All of these were developed with little concern for First Amendment constraints because the activity targeted by the regulations can be characterized as information *gathering*.¹⁰⁵

99. Except through the application of neutral regulations, such as traffic and labor laws, the answer of course is no. A law that burdened production in order to thwart the production is another matter.

100. In Part IV, I discuss the considerable public interests in preserving seclusion and suggest that laws tailored to these interests should be able to withstand First Amendment scrutiny. Thus, the outcome of *Dietemann* does not seem nearly as troubling as its reasoning, which would remove all prohibitions of recordings from First Amendment coverage.

101. See CAL. PENAL CODE § 631-32 (West 2013); CONN. GEN. STAT. § 52-570d(a) (2013); FLA. STAT. § 934.03 (2013); 720 ILL. COMP. STAT. 5/14-2 (2013); MD. CODE ANN., CTS. & JUD. PROC. § 10-402 (LexisNexis 2013); MASS. GEN. LAWS ch. 272, § 99 (2012); MICH. COMP. LAWS § 750.539c (2013); MONT. CODE ANN. § 45-8-213 (2013); N.H. REV. STAT. ANN. § 570-A:2 (2013); 18 PA. CONS. STAT. § 5704 (2012); WASH. REV. CODE § 9.73.030 (2013).

102. *Pineda v. Williams-Sonoma Stores, Inc.*, 246 P.3d 612, 616 (Cal. 2011).

103. These concepts are referred to as data "minimization" and "destruction." See FTC PRIVACY REPORT, *supra* note 13, at 29; WHITE HOUSE PRIVACY REPORT, *supra* note 13, at 52.

104. See *supra* note 13.

105. The division between information and information gathering is Neil Richards's sole basis for his conclusion that privacy laws do not implicate the First Amendment. Richards notes that the distinction *can* cause problems at the extreme, but he does not regard these problems to be unavoidable. "One can imagine science fiction-style hypotheticals that would bring information collection rules within this doctrine—for example, a law forbidding the keeping of records or outlawing cameras." See Richards, *supra* note 14, at 1189.

First Amendment rules on the creation of data can be explored through the courts' treatment of documentary photographs, since they are a familiar subset of data. Digital cameras and Adobe Photoshop have taught us to understand photographs as data—as nothing more than a grid of color pixels.¹⁰⁶ Just like other forms of data, photographs are sometimes created for one purpose even though their long-term value turns out to be completely unrelated. Other times, photographs are taken without any particular audience or purpose in mind except to satisfy the curiosity of the photographer. Still other images wind up having tremendous interest from a large audience even though they were taken without any human looking through the lens, such as the 2012 Miami “zombie attack” caught by security cameras.¹⁰⁷ Much can be learned by looking at the precarious legal treatment of documentary photographs. As goes photography, so goes data.

Consistent with *Dietemann*, some courts have justified treating photography as nonexpressive conduct because it can be done without an audience in mind. Security cameras create images without the involvement of anybody resembling an author. Arguably, the First Amendment should require, at the very least, somebody who can be identified as a speaker, and somebody else who could be an intended audience. As one court put it: “To achieve First Amendment protection, a plaintiff must show that he possessed: (1) a message to be communicated; and (2) an audience to receive that message, regardless of the medium in which the message is to be expressed.”¹⁰⁸

Courts have used tests like this one to find that photographers have no First Amendment right to take photographs or videos in public places. In *Porat v. Lincoln Towers Community Ass'n*, a man who was ticketed for photographing a public courtyard lost his claim attempting to vindicate his First Amendment rights.¹⁰⁹ He lost because he had described himself as a “photo hobbyist,” a

106. In fact, photographs have always been grids of data. Black and white film is simply a grid of light-sensitive silver grains that record their exposure to light. See Charles Woodworth, *How Photographic Film Works*, HOWSTUFFWORKS, <http://electronics.howstuffworks.com/film7.htm> (last visited Dec. 18, 2013).

107. See Brad Lendon, *Miami “Zombie” Attack Video: Footage Shows Entire Attack by Rudy Eugene on Ronald Poppo, Police Say*, WPTV.COM (May 31, 2012), http://www.wptv.com/dpp/news/news_archives/miami-zombie-attack-video-footage-shows-entire-attack-by-rudy-eugene-on-ronald-poppo-police-say.

108. *Porat v. Lincoln Towers Cmty. Ass'n*, No. 04 Civ. 3199(LAP), 2005 WL 646093, at *4 (S.D.N.Y. Mar. 21, 2005) (citing *Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp.*, 515 U.S. 557, 568 (1995)), *aff'd*, 464 F.3d 274 (2d Cir. 2006); see also *Texas v. Johnson*, 491 U.S. 397, 404 (1989). However, both of these cases assessed whether actions that would otherwise be conduct were sufficiently expressive to receive First Amendment protection. Regulation of a photograph, or a database, is not concerned with the conduct required to create the record; it is the record itself at issue. Attempts to ban photography or the collection of data have nothing to do with the acts that were necessary to collect the information and have everything to do with the content of the information itself.

109. 2005 WL 646093, at *4-5 (citing *Hurley*, 515 U.S. at 568).

phrase that unexpectedly doomed his case, because, according to the district court, “[h]e effectively disclaim[ed] any communicative property of his photography as well as any intended audience by describing himself as a ‘photo hobbyist,’ and alleging that the photographs were only intended for ‘aesthetic and recreational’ purposes.”¹¹⁰

In contrast, in *Pomykacz v. Borough of West Wildwood*, a woman who followed around the town’s mayor photographing him, and whose conduct may have satisfied the actus reus elements of criminal harassment, persuaded the court that she had a First Amendment interest in her photographs because they were part of her political activism and could be used to corroborate her theories that the mayor engaged in nepotism (if, indeed, these theories were borne out).¹¹¹

These cases certainly craft a clear enough rule, and they just as certainly train future photographers about what to say during a deposition (e.g., claim to be an activist, never a hobbyist); but should the First Amendment shield only speakers who have audiences from government interference?

Superficially, these cases look like they are mandated by the Supreme Court’s edict that speech must contain an “intent to convey a particularized message” likely to be “understood by those who viewed it.”¹¹² But the Court established this requirement in the context of flag desecration, an area that forced the Court to draw a line between nonspeech action and symbolic acts consciously designed to send a message.¹¹³ The rule is a mismatch for the pro-

110. *Id.* at *5 (citation omitted); see also *Carson v. Cnty. of Stanislaus*, No. 1:10-cv-02133-OWW-SMS, 2011 WL 1532533, *1-3 (E.D. Cal. Apr. 20, 2011) (dismissing the First Amendment claim because an inference about a particularized message could not be drawn from a lawyer’s attempt to take a photograph of a district attorney investigator); *Larsen v. Fort Wayne Police Dep’t*, 825 F. Supp. 2d 965, 979-80 (N.D. Ind. 2010) (finding that a father’s attempt to video record his daughter’s choir concert does not communicate an idea for First Amendment purposes).

111. 438 F. Supp. 2d 504, 506-07, 512-13 (D.N.J. 2006). Courts will also find a First Amendment interest if the photographer has no particular audience in mind, but hopes very much to capture something interesting which might later prove to be valuable to an audience. Such was the case with Beau Lambert, who took his video camera to downtown Des Moines in the hopes of filming something newsworthy. After he got his wish and filmed a fatal brawl in the streets, he successfully used the First Amendment to demand the return of his confiscated videotape. *Lambert v. Polk Cnty.*, 723 F. Supp. 128, 130-31, 133, 135 (S.D. Iowa 1989).

112. *Spence v. Washington*, 418 U.S. 405, 410-11 (1974) (per curiam); see also *Rumsfeld v. Forum for Academic & Institutional Rights*, 547 U.S. 47, 65-66 (2006) (holding that if an observer cannot understand the message of some particular conduct without additional explanation, that is “strong evidence” that the conduct is not “inherently expressive”); CHARLES C. HAYNES ET AL., *FIRST AMENDMENT CTR., THE FIRST AMENDMENT IN SCHOOLS* 77 (2003).

113. See *Spence*, 418 U.S. at 405-06; see also *Johnson*, 491 U.S. 397 (finding an expressive interest in flag burning); *Smith v. Goguen*, 415 U.S. 566 (1974) (avoiding a decision on the First Amendment interest in sewing a flag to the seat of one’s trousers).

cess of taking a photograph. Few photographers would locate speech interests in the *act* of taking pictures. The photographs themselves, and not the acts involved in creating them, are the source of the photographer's First Amendment interest.

Taken to its logical end, a free speech rule requiring an audience would fail to protect diaries. This result denies the value of private thought and intellectual growth that scholars and jurists have traditionally endorsed.¹¹⁴

In any event, even if protected speech requires a speaker and an audience, photography bans ought to attract scrutiny by their nature. The very purpose of a photography ban is to prevent a wider audience from seeing the scene. With the exception of bans on *flash* photography, which are often designed to prevent the deleterious effects of light on sensitive objects, photo bans are designed to cut down on communicative potential.¹¹⁵

In time, the rule that mechanically capturing information is nonexpressive conduct will prove to be unworkable. All-party-consent wiretap statutes have already begun to fall apart under the weight of increased judicial scrutiny. These cases and others like it are discussed next.

114. C. Edwin Baker specifically addresses the First Amendment protections of a diary (as well as solving a problem or singing out loud alone). See C. Edwin Baker, *Scope of the First Amendment Freedom of Speech*, 25 UCLA L. REV. 964, 993 (1978). Martin Redish also insists the First Amendment must cover the keeping of a diary. See Martin H. Redish, *Freedom of Thought as Freedom of Expression: Hate Crime Sentencing Enhancement and First Amendment Theory*, CRIM. JUST. ETHICS, Summer/Fall 1992, at 29, 30-31. Naturally, Seana Shiffrin, whose theory of a thinker-based First Amendment forms the backbone for my proposal here, also agrees that "diaries and other forms of discourse meant primarily for self-consumption" should indisputably be within the scope of speech protections. See Seana Valentine Shiffrin, *A Thinker-Based Approach to Freedom of Speech*, 27 CONST. COMMENT. 283, 285 (2011). To save the speech value of a diary, the rule at the very least must be modified so that the intended audience can be the author himself. With this modification, the photo hobbyist should have access to First Amendment protections, too. Moreover, the limit can be avoided through tricks. A photographer can meet the threshold by alleging that he intends to display the photograph somewhere. Surely a sophisticated corporation could ensure it meets the same intent rule if the issue were to arise in the context of its data. To be safe, it could even build in routine protections by hiring people to look at the data it generates, guaranteeing that the audience requirement is met. But it would be odd if First Amendment analysis of data could be radically changed just by moving data across a human eyeball. Surely the First Amendment deserves a more thoughtful limiting principle than this.

115. The same can be said for the regulation of data more generally. By the time litigants ask courts for First Amendment protection of their databases, the data has already proven to be interesting to some audience. The prescription data in *IMS Health* was interesting to pharmaceutical companies that make office visits to the prescribing doctors. A borrower's loan repayment history data is interesting to other banks considering whether to extend credit. Anticipation of an audience is the very thing that inspires governments to regulate data flow. Thus, the distinction between information and information gathering is dubious in this context.

C. *Data Creation as Expressive Conduct*

Over the last fifteen years, several state police departments, exercising exceedingly poor judgment, have chosen to arrest citizens who recorded the conduct of on-duty police officers.¹¹⁶ Some (though not all) of the courts dismissed the charges based on a First Amendment right to record. However, with one exception, the right was crafted narrowly, as a right to record public officials performing their public duties.¹¹⁷ The Department of Justice has written an opinion letter recognizing the right on narrower terms still—protecting a person’s right to record “police activity.”¹¹⁸

But last year, in *American Civil Liberties Union of Illinois v. Alvarez*, the Seventh Circuit enjoined the enforcement of Illinois’s all-party-consent wiretap statute out of recognition to a broader right to record¹¹⁹: a right to record anything.¹²⁰ The majority (over a thoughtful dissent by Judge Richard Posner) arrived at its position because a well-functioning First Amendment must unite the dissemination of speech with the creation of speech:

The act of *making* an audio or audiovisual recording is necessarily included within the First Amendment’s guarantee of speech and press rights as a corollary of the right to disseminate the resulting recording. . . .

. . . .

. . . . [T]here is no fixed First Amendment line between the act of creating speech and the speech itself

This observation holds true when the expressive medium is mechanical rather than manual. . . .

. . . .

Audio and audiovisual recording are communication technologies, and as such, they enable speech. Criminalizing all nonconsensual audio recording

116. *See, e.g., Kelly v. Borough of Carlisle*, 622 F.3d 248, 251 (3d Cir. 2010); *Commonwealth v. Hyde*, 750 N.E.2d 963, 964-65 (Mass. 2001); *see also Colten v. Kentucky*, 407 U.S. 104, 109 (1972) (finding that a person does not have a “constitutional right to observe the issuance of a traffic ticket,” even when he is not violating any trespass or traffic law).

117. *See, e.g., Glik v. Cunniffe*, 655 F.3d 78, 79 (1st Cir. 2011); *Gilles v. Davis*, 427 F.3d 197, 212 n.14 (3d Cir. 2005); *Smith v. City of Cumming*, 212 F.3d 1332, 1333 (11th Cir. 2000); *Robinson v. Fetterman*, 378 F. Supp. 2d 534, 541-42 (E.D. Pa. 2005) (finding a right to record in part because “[t]he activities of the police, like those of other public officials, are subject to public scrutiny”); *see also United States v. Rosen*, 445 F. Supp. 2d 602, 633 (E.D. Va.) (finding that “[t]he collection and discussion of information about the conduct of government” is a core value of the First Amendment), *amended by* No. 1:05cr225, 2006 WL 5049154 (E.D. Va. 2006), *aff’d*, 557 F.3d 192 (4th Cir. 2009).

118. Letter from Jonathan M. Smith, Chief of Special Litig. Section of Civil Rights Div., U.S. Dep’t of Justice, to Mark H. Grimes, Balt. Police Dep’t 2-3 (May 14, 2012), available at http://www.justice.gov/crt/about/spl/documents/Sharp_ltr_5-14-12.pdf.

119. 679 F.3d 583, 586-87 (7th Cir.), *cert. denied*, 133 S. Ct. 651 (2012).

120. The court reserved for another day the decision whether surreptitious recordings and recordings in nonpublic places receive different First Amendment treatment. *Id.* at 606.

necessarily limits the information that might later be published or broadcast—whether to the general public or to a single family member or friend—and thus burdens First Amendment rights.¹²¹

Other courts have inched toward the same position by advancing a right to access information. In *S.H.A.R.K. v. Metro Parks Serving Summit County*, an animal rights group brought a civil rights claim against a county for disposing of its video footage of a deer culling in a state park.¹²² The activists had set up video cameras in trees to record during hours that the park was closed to visitors. The Sixth Circuit found that the animal rights activists had a First Amendment interest in recording the deer culling, and more generally in “access to information.”¹²³ However, the right was differentiated from, and given significantly less scrutiny than, the right to expression, and the animal rights activists ultimately lost their case.¹²⁴

A right to access information (or, more precisely, a right to be free from government restraint on access to information) is at odds with *Dietemann* and other cases that presume the First Amendment imposes absolutely no constraint on the tort of intrusion upon seclusion. But the *S.H.A.R.K.* approach seems necessary. If access to knowledge were not a constitutionally protected right, the intrusion tort could be boundless. At the extreme, the government could prohibit a person from recording anything at all without conflicting with the First Amendment.

This cannot be right. Though the Supreme Court has not directly addressed the question whether mechanical recordings are protected speech, it has recognized that information gathering is a necessity for speech. In *Branzburg v. Hayes*, although the Court found that neutrally applicable laws apply to the press, it recognized that “without some protection for seeking out the news, freedom of the press could be eviscerated.”¹²⁵ And in *Time, Inc. v. Hill*, the Court found that the First Amendment offered a complete defense to a family’s lawsuit against Time, Inc. for its imaginative depiction in *Life Magazine* of a home invasion the family had lived through.¹²⁶ “Exposure of the self to others in varying degrees is a concomitant of life in a civilized community. The risk of this exposure is an essential incident of life in a society which places primary

121. *Id.* at 595-97.

122. 499 F.3d 553, 557-58 (6th Cir. 2007).

123. *Id.* at 559.

124. *Id.* at 559-60, 563 (quoting *D’Amario v. Providence Civic Ctr. Auth.*, 639 F. Supp. 1538, 1543 (D.R.I. 1986)); see also *Shulman v. Grp. W Prods., Inc.*, 955 P.2d 469, 474-75, 497 (Cal. 1998) (protecting the right to broadcast conversations surreptitiously recorded in a medivac helicopter, but offering no protection for the right to record the conversations in the first place). No reasons were given for adopting a lower standard for the right to access to information.

125. 408 U.S. 665, 681-82 (1972).

126. 385 U.S. 374, 376-78, 387-88 (1967).

value on freedom of speech and of press.”¹²⁷ The “risk of this exposure” that the Court describes is the flipside of a right for individuals to observe each other. If there are constitutional limits on the restrictions for information gathering, it must mean that, despite the rhetoric to the contrary, the First Amendment protects the right to gather information in some fashion.

This is not to say that the state is powerless to create laws that intentionally obstruct access to information. As I show in Part IV, it can with proper justification. The outcome of *Dietemann* may survive scrutiny if the state can show it has a strong interest in protecting its constituents’ seclusion by outlawing surreptitious recordings in one’s own home. Mr. Dietemann is a sympathetic plaintiff because the humiliating photographs were taken without his knowledge, and on his own turf.¹²⁸ Mr. Dietemann has a liberty interest in keeping surreptitious recording equipment out of his home, and such an interest could compete in court with the reporters’ interests in recording information that they lawfully observed. Given the long-held expectations that the home provides sanctum,¹²⁹ Mr. Dietemann may have won his case despite the First Amendment arguments.

Outside of our homes, on the other hand, enforcement of an expectation that people will not record what we say or do is in conflict with *their* liberty interest in knowledge creation.

To this day, the right to access information is underdeveloped.¹³⁰ Its relationship to full speech rights is awkward. Courts recognize that the right to free speech is hollow without access to information, but the constitutional protection of information has yet to achieve coherence.

III. THE RIGHT TO CREATE KNOWLEDGE

This Part makes the normative case for the First Amendment protection of data. First, I define a “right to create knowledge,” which is a latent prerequisite for free expression.¹³¹ Speech does very little for a government’s constituents if it is not supported by commitments to free thought and information flow.

127. *Id.* at 388.

128. *See* *Desnick v. Am. Broad. Cos.*, 44 F.3d 1345, 1352-53 (7th Cir. 1995) (Posner, C.J.) (describing the facts of *Dietemann*).

129. *See* Warren & Brandeis, *supra* note 24, at 220.

130. Adding to the confusion is the fact that the same phrase is sometimes used to mean a positive right—the affirmative obligation for the government to provide public access to books and information. For a discussion of both the positive and negative qualities, see Kay Mathiesen, *The Human Right to a Public Library*, J. INFO. ETHICS, Spring 2013, at 60; and Kay Mathiesen, *The Human Right to Internet Access: A Philosophical Defense*, INT’L REV. INFO. ETHICS, Dec. 2012, at 9.

131. I use this phrasing rather than the right of “access to information” because it avoids ambiguity. “Access to information” has been used to describe many proposed rights, includ-

I outline a framework for analyzing whether the right to create knowledge has been impeded. A motive analysis will frequently suffice. When a law or regulation has the very purpose of limiting knowledge, the restriction must undergo First Amendment scrutiny. Data privacy laws have the unabashed goal of limiting, and shaping, what the government's constituents can know. There are often good reasons to do this. There may even be compelling reasons to do this. But, to be confident, direct regulations of data should draw scrutiny.

Next, this Part analyzes whether the protection of data is consistent with other competing scholarly theories about the First Amendment's purpose. The right to create knowledge in general, and the protection of data specifically, fit very comfortably within all of them. This Part closes with some thoughts about the level of scrutiny that should apply to regulations of data.

A. *The Negative Right to Create Knowledge*

The negative right to create knowledge ensures that the state will not interfere unduly with its constituents' learning. This is not an entirely new concept for courts or for scholars, but the varied instantiations of a right to knowledge have been developed in the abstract and through unusual fact patterns.

Courts have already had occasion to interpret "speech" expansively so that it encompasses the right to receive or access information.¹³² The impetus for doing so is plain: free speech will have little value if the government has substantial influence over the ideas and facts that speakers are permitted to consider. At a higher level of generality, the First Amendment safeguards the freedom of thought. The Supreme Court recognized this right in *Stanley v. Georgia*, when it disallowed enforcement of an obscenity ban, deciding that Georgia "cannot constitutionally premise legislation on the desirability of controlling a person's private thoughts."¹³³ Freedom of thought is an old and uniquely American liberty. Samuel Adams, celebrating the signing of the Declaration of Independence, commented: "[F]reedom of thought and the right of private judgment, in matters of conscience . . . direct their course to this happy country as their last asylum."¹³⁴

In 2002, the Supreme Court reaffirmed this sentiment, stating that "[t]he right to think is the beginning of freedom, and speech must be protected from the government because speech is the beginning of thought."¹³⁵ But doesn't

ing positive rights of access to government information and to libraries. See sources cited *supra* note 130.

132. See discussion *supra* Part II.A.2.

133. 394 U.S. 557, 559, 566 (1969).

134. MATTHEW C. PRICE, *THE WILSONIAN PERSUASION IN AMERICAN FOREIGN POLICY 1* (2007) (internal quotation mark omitted).

135. *Ashcroft v. Free Speech Coal.*, 535 U.S. 234, 253 (2002). This passage is preceded by the following: "First Amendment freedoms are most in danger when the government seeks to control thought or to justify its laws for that impermissible end." *Id.*; see also Palko

this have the order reversed? Thought is almost always a precursor to utterances, art, and other forms of expression. And, although a subset of thoughts is inspired by the speech of others, many thoughts are not. They are either the product of original ideas or first-time observations of the world.

While courts have been slow to flesh out the right to free thought, constitutional law scholars have laid a good deal of groundwork to provide a theoretical justification for deriving a freedom of thought from the freedom of speech.¹³⁶ Seana Shiffrin has done a particularly nice job articulating a thinker-centered theory for the First Amendment,¹³⁷ so I borrow heavily from her work to operationalize the right to create knowledge.

For Shiffrin, the First Amendment is called into service when a statute, regulation, court decision, or lawmaking activity (1) on its face exhibits a design to “ban or attempt to ban the free development and operation of a person’s mind or those activities or materials necessary for its free development and operation”; (2) has the effect of interfering too greatly with the free development and operation of a person’s mind; or (3) has a rationale which, even if not overtly designed to conflict with the free development of a person’s mind, is nevertheless unacceptably inconsistent with that right.¹³⁸

The right to create knowledge ought to follow the same framework, but with a narrower focus on the uninhibited acquisition of knowledge. The “free development and operation of a person’s mind” has nearly infinite range. Public schools, by selecting what to teach and what not to teach, might bar other options the student would otherwise have to *develop* her mind. But a school’s curriculum does not bar the student’s opportunity to acquire specific pieces of knowledge that the student can acquire elsewhere. In short, the right to create knowledge promises freedom from intentional or excessive government restraints on learning something new.

v. Connecticut, 302 U.S. 319, 326-27 (1937) (“[F]reedom of thought, and speech . . . is the matrix, the indispensable condition, of nearly every other form of freedom.”), *overruled on other grounds by* Benton v. Maryland, 395 U.S. 784 (1969).

136. See, e.g., ANTHONY LEWIS, FREEDOM FOR THE THOUGHT THAT WE HATE: A BIOGRAPHY OF THE FIRST AMENDMENT (2007); Marc Jonathan Blitz, *Freedom of Thought for the Extended Mind: Cognitive Enhancement and the Constitution*, 2010 WIS. L. REV. 1049; Dana Remus Irwin, *Freedom of Thought: The First Amendment and the Scientific Method*, 2005 WIS. L. REV. 1479 (endorsing strict scrutiny any time a regulation intrudes on free thought); Roy G. Spece, Jr. & Jennifer Weinzierl, *First Amendment Protection of Experimentation: A Critical Review and Tentative Synthesis/Reconstruction of the Literature*, 8 S. CAL. INTERDISC. L.J. 185 (1998); see also Blitz, *supra* note 25, at 182; Kolber, *supra* note 9, at 7-8 (organizing the scholarship into two categories of argument: (1) arguments that the right to free thought is *intrinsic* to the right to speech and it deserves protection for its own sake, and (2) arguments that free thought receives *instrumental* protection because it is such an important tool for speakers, much like ink).

137. See Shiffrin, *supra* note 114.

138. *Id.* at 287.

This framework avoids the need to define the nature of “speech.” Instead, the framework focuses on the nature of the state action. It asks what purpose a regulation seeks to serve and how the regulation operates in practice. The previously vexing scope question is much more manageable when the analysis centers on the regulation rather than the object of the regulation.¹³⁹

A purpose-driven test lines up with Elena Kagan’s insight that the Court’s recognition of speech interests tracks an often-implicit search for the inappropriate motives of lawmakers.¹⁴⁰ However, under the framework, a regulation promulgated with pure motives—motives unrelated to the restriction of knowledge creation—can nevertheless trigger scrutiny if the practical effect of the law causes an unreasonable hindrance to free inquiry.

Although First Amendment scholars tend to overlook purposive analyses of speech, they are not unprecedented.¹⁴¹ In *United States v. O’Brien*, the famous case about draft card burning, the Supreme Court set out the constitutional test for the regulation of an act that it had already found to have enough semantic content to be a form of speech.¹⁴² A regulation could survive scrutiny only “[1] if it furthers an important or substantial governmental interest; [(2)] if the governmental interest is unrelated to the suppression of free expression;

139. Quests to find the meaning of “speech” in the abstract, without reference to any particular state action, have led to a good deal of frustration among scholars. *See, e.g.*, Stuart Minor Benjamin, *Transmitting, Editing, and Communicating: Determining What “The Freedom of Speech” Encompasses*, 60 DUKE L.J. 1673, 1675-77 (2011). Many of the seeming contradictions in precedent can be explained through a motive analysis, even if the court failed to apply one at the time. Moreover, this test ensures that the right to knowledge creation is not “unrestrained.” *See* *Zemel v. Rusk*, 381 U.S. 1, 17 (1965) (explaining that the “right to speak and publish does not carry with it the unrestrained right to gather information”).

140. *See* Kagan, *supra* note 46, at 414. Leslie Kendrick has similarly made sense of the content-based discrimination doctrine by showing that the Court typically looks for an invidious purpose to discriminate against certain types of viewpoints or subject matters. Leslie Kendrick, *Content Discrimination Revisited*, 98 VA. L. REV. 231, 248 & n.61 (2012); *see also* Geoffrey R. Stone, *Free Speech in the Twenty-First Century: Ten Lessons from the Twentieth Century*, 36 PEPP. L. REV. 273, 277-82 (2009) (suggesting that courts are on the lookout for illegitimate motivations and explaining the influence of Equal Protection Clause jurisprudence on the development of this First Amendment doctrine).

141. *See* *Bd. of Educ. v. Pico*, 457 U.S. 853, 871 (1982) (“[W]hether petitioners’ removal of books from their school libraries denied respondents their First Amendment rights depends upon the motivation behind petitioners’ actions. If petitioners *intended* by their removal decision to deny respondents access to ideas with which petitioners disagreed, and if this intent was the decisive factor in petitioners’ decision, then petitioners have exercised their discretion in violation of the Constitution.” (footnote omitted)); *see also* *Hill v. Colorado*, 530 U.S. 703, 718 n.25 (2000) (discussing the purpose of the challenged statute); *Turner Broad. Sys., Inc. v. FCC*, 512 U.S. 622, 645 (1994) (noting that the Court has recognized that a facially neutral regulation “may be content based if its manifest purpose is to regulate speech because of the message it conveys”).

142. 391 U.S. 367, 369, 377 (1968).

and [(3)] if the incidental restriction on alleged First Amendment freedoms is no greater than is essential to the furtherance of that interest.”¹⁴³

John Hart Ely helpfully explained a few years later that the rule is “incomplete”; a regulation that does not satisfy element (2), because it intentionally interferes with speech, may still survive constitutional scrutiny, but must undergo a different and more exacting scrutiny.¹⁴⁴ So long as free thought is included under the umbrella of free expression, Shiffrin’s scheme tracks the *O’Brien* rule quite nicely, perhaps even conservatively.

The framework also supplies manageable limiting principles to the First Amendment’s protection of knowledge. Not every regulation affecting knowledge will draw scrutiny; plenty of laws are inimical to a person’s ability to learn new things without having a direct purpose to obstruct knowledge.¹⁴⁵ For example, the prohibition of theft will prevent a thief from learning the content of a book that he may plan to steal, but the purpose of the law is to enforce tangible property rights, not to inhibit the knowledge of thieves.

The thinker-based approach to speech has had a sleepy arrival to First Amendment discourse. Our government, quite fortunately, is not often in the business of overtly manipulating thought. However, the right to knowledge creation has obvious application to the modern data privacy debate. At the beginning of this exploration, the automated generation of data appeared to be a close case for speech. Though the humming of servers and the shuttering of cameras may share superficial similarities with conduct, a motive analysis reveals very clearly on which side of the speech line data privacy laws must fall. Data privacy laws are purposefully designed to interfere with somebody’s (or some company’s) knowledge.¹⁴⁶

A law prohibiting the creation, maintenance, or distribution of digital information attempts to achieve its social goals by limiting the accumulation of knowledge. Data privacy laws strive to give individuals the power to decide who does and does not get to learn about them.¹⁴⁷ As principled as these restrictions may be, they must draw scrutiny if the First Amendment is to protect

143. *Id.* at 377.

144. John Hart Ely, *Flag Desecration: A Case Study in the Roles of Categorization and Balancing in First Amendment Analysis*, 88 HARV. L. REV. 1482, 1483-84 (1975).

145. For this reason, Dana Remus Irwin’s proposed protections for the freedom of thought, which demand strict scrutiny any time a regulation intrudes into free thought, are unworkable and potentially boundless. See Irwin, *supra* note 136, at 1507.

146. Daniel Solove identifies the interference of thought and knowledge as the purpose of privacy laws; privacy laws curtail “irrational judgments” and “guard against rational judgments that society may desire to curtail.” See Solove, *supra* note 22, at 1034-36, 1041.

147. The very first principle guiding the White House’s proposed “Consumer Privacy Bill of Rights” is “[i]ndividual [c]ontrol: [c]onsumers have a right to exercise control over what personal data companies collect from them and how they use it.” See WHITE HOUSE PRIVACY REPORT, *supra* note 13, at 1.

the creation of new knowledge. Any other outcome spirals quickly into special pleading.

The same can be said for trade secret laws, antihacking statutes, and other information laws established to protect information security and maintain economic incentives.¹⁴⁸ As with data privacy laws, the government may be acting out of the best of intentions, inhibiting certain thoughts that tend to produce bad effects on society. In these cases, the government may regulate thought as a way station, or convenient midway point, to achieve more tangible and legitimate goals. Nevertheless, if freedom of thought has First Amendment protection, the government must justify its programs within the judicial scrutiny that applies to other well-meaning regulations of speech.

There are, of course, other theories of the First Amendment that do not center on thought. Since no single theory can lay exclusive claim to the First Amendment, the next Subpart analyzes how well the protection of data corresponds with alternative visions for the First Amendment.

B. *Data and First Amendment Objectives*

The objectives of the First Amendment are large and contain multitudes.¹⁴⁹ At times courts have emphasized the centrality of political discourse and have given utterances of political dissent the distinctive title of “core” political speech.¹⁵⁰ At other times, the courts have rebuked the suggestion that political speech receives greater protection than other speech.¹⁵¹ The objectives put forward by scholars are similarly fractured. The theories break into five nonexclusive categories¹⁵²: (1) the marketplace of ideas; (2) the preservation of a public

148. Pamela Samuelson has explored potential clashes between trade secrets law and the First Amendment. See Pamela Samuelson, *Principles for Resolving Conflicts Between Trade Secrets and the First Amendment*, 58 HASTINGS L.J. 777 (2007).

149. Apologies for butchering a beautiful Walt Whitman poem. WALT WHITMAN, *Song of Myself* (1891), in *SONG OF MYSELF AND OTHER POEMS BY WALT WHITMAN* 71, 131 (Counterpoint 2010). For another example of a speech analysis that looks to First Amendment goals in order to understand its scope, see Andrew Tutt, Note, *Software Speech*, 65 STAN. L. REV. ONLINE 73 (2012).

150. See *Meyer v. Grant*, 486 U.S. 414, 422 (1988).

151. See *Cohen v. California*, 403 U.S. 15, 25 (1971) (“[W]holly neutral futilities . . . come under the protection of free speech as fully as do Keats’ poems or Donne’s sermons . . .” (second alteration in original) (quoting *Winters v. New York*, 333 U.S. 507, 528 (1948) (Frankfurter, J., dissenting)) (internal quotation marks omitted)).

152. For academic works advancing these different theories, see STEVEN H. SHIFFRIN, *THE FIRST AMENDMENT, DEMOCRACY, AND ROMANCE* 46 (1990); David A. Anderson, *The Origins of the Press Clause*, 30 UCLA L. REV. 455 (1983); Baker, *supra* note 114; Daniel A. Farber, Commentary, *Free Speech Without Romance: Public Choice and the First Amendment*, 105 HARV. L. REV. 554 (1991); Schauer, *supra* note 34; James Weinstein, *Participatory Democracy as the Central Value of American Free Speech Doctrine*, 97 VA. L. REV. 491 (2011).

resource; (3) deliberative democracy (and, particularly, countermajoritarian expression); (4) a check on state power; and (5) self-determination. The expressive nature of data is quite consistent with the first four objectives, and the fifth poses some interesting questions.

1. *Marketplace of ideas and public good theories*

Proponents of the marketplace of ideas see speech as an open exchange that allows bad ideas to be bested by sounder ones.¹⁵³ The public good theory is closely related. It posits that the benefits of speech are indirect (society's gradual accumulation of knowledge and rejection of bad ideas) while its harms are direct (insult, or reputational harms), so speech needs extra protection from shortsighted regulations.¹⁵⁴ The First Amendment supports both of these goals when it limits the state from creating laws that might distort the free competition for the minds of Americans. Wrongheaded as these theories may be (the marketplace has received much just criticism),¹⁵⁵ they inarguably have some claim to the First Amendment's original intent.¹⁵⁶

The marketplace of ideas and public good theories dovetail with the right of knowledge creation described above. Indeed, the marketplace of ideas is understood to promote cognitive processes.¹⁵⁷ Recorded facts are part of the marketplace. The marketplace is expected to include "the widest possible dissemination of information from diverse and antagonistic sources."¹⁵⁸ Moreover, the marketplace of ideas is consistent with the protection of data *collection*, and not just its distribution. A person's opportunity to receive information and ideas

153. See JOHN STUART MILL, ON LIBERTY 55-56 (Curran V. Shields ed., Bobbs-Merrill Co. 1956) (1859).

154. Farber, *supra* note 152, at 560.

155. See Derek E. Bambauer, *Shopping Badly: Cognitive Biases, Communications, and the Fallacy of the Marketplace of Ideas*, 77 U. COLO. L. REV. 649 (2006). Frederick Schauer makes the sharp rebuke that the very rhetoric around the marketplace of ideas may itself be a counterexample to the notion that bad ideas eventually fade when confronted with corrective speech. Frederick Schauer, *Facts and the First Amendment*, 57 UCLA L. REV. 897, 910-11 (2010).

156. "If there be any among us who would wish to dissolve this union, or to change its republican form, let them stand undisturbed, as monuments of the safety with which error of opinion may be tolerated where reason is left free to combat it." Thomas Jefferson, First Inaugural Address (Mar. 4, 1801), in 8 THE WRITINGS OF THOMAS JEFFERSON 1, 3 (Paul Leicester Ford ed., N.Y., G.P. Putnam's Sons 1897); see MILL, *supra* note 153, at 55-56; see also DARIEN AUBURN MCWHIRTER, FREEDOM OF SPEECH, PRESS, AND ASSEMBLY 1-2 (2002) (discussing Mill's influence during the earliest battles of First Amendment interpretation).

157. POST, *supra* note 36, at 6.

158. *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 791 (1985) (Brennan, J., dissenting) (quoting *Associated Press v. United States*, 326 U.S. 1, 20 (1945)) (internal quotation mark omitted). Reference to facts always makes an appearance in discussions of the "idea" half of the idea-expression dichotomy of copyright, too. See, e.g., *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 556 (1985).

should be protected whether he receives information from another person (a traditional “speaker”) or through his direct observations of the world. A preference for knowledge received *indirectly* from other people would defy the logic of a robust marketplace.

If a chief constitutional goal is to vet ideas against one another, data is precisely the sort of grist that can corroborate or contradict various theories. Data collects for human consumption what cannot be observed directly. In the words of Robert Post, “If we wish to know whether cigarettes are carcinogenic or whether high tariffs produce market inefficiencies or whether plutonium-239 has a half-life of about 24,000 years, we cannot intelligently speak for ourselves.”¹⁵⁹

Personal data has the power to dramatically shift scientific consensus and public opinion. Consider the ulcer. The medical and popular consensus once was that ulcers were caused by stress. In fact, common law courts that required proof of physical manifestations of stress before awarding damages for negligent infliction of emotional distress accepted ulcers as such proof.¹⁶⁰ But Barry Marshall’s quest to understand a type of bacteria that can survive in stomach acid led him to the discovery that all the patients that he studied who had ulcers also had the bacteria. And when the ulcer patients were treated with antibiotics, they were cured 900% more often than those receiving conventional treatment (the conventional treatment being to “relax”).¹⁶¹ It took another dozen years for the medical community to accept Marshall’s findings, but considering they ran against a deeply rooted truism, the change in treatment and understanding was swift.¹⁶²

Swifter still was the rejection of the long-accepted theory that humans killed off all of their contemporary hominid populations when *Homo sapiens* left Africa to go dominate the rest of the world.¹⁶³ DNA sequencing of both a Neanderthal and a sample of modern Americans allowed geneticists David

159. POST, *supra* note 36, at xi-xii. It should be noted that Post focuses here on an entirely different question from the one that I address. Post is setting up a discussion about the inherent tension between speech and expertise, between a desire to increase speech while also developing the skills to discriminate between claims and reject those that do not survive accepted methodologies.

160. *See, e.g.,* Richardson v. J.C. Penney Co., 649 P.2d 565, 566-67 (Okla. Civ. App. 1982) (allowing the plaintiff to recover for negligent infliction of emotional distress because he had a “bleeding ulcer”).

161. Kathryn Schulz, *Stress Doesn’t Cause Ulcers! Or, How to Win a Nobel Prize in One Easy Lesson: Barry Marshall on Being . . . Right*, SLATE (Sept. 9, 2010, 6:35 AM), http://www.slate.com/blogs/thewrongstuff/2010/09/09/stress_doesn_t_cause_ulcers_or_how_to_win_a_nobel_prize_in_one_easy_lesson_barry_marshall_on_being_right.html.

162. *Id.*

163. *See* Elizabeth Kolbert, *Sleeping with the Enemy: What Happened Between the Neanderthals and Us?*, NEW YORKER (Aug. 15, 2011), http://www.newyorker.com/reporting/2011/08/15/110815fa_fact_kolbert.

Reich and Svante Pääbo to discover that *Homo sapiens* bred with Neanderthals.¹⁶⁴ Some of us are their distant descendants.

Examples of data changing minds are abundant, and they are not limited to health and genetic data. A study by Benjamin Edelman using Internet consumer data found that web users in the most conservative states had the highest per capita subscriptions to pornographic websites.¹⁶⁵ Utah leads the pack, even after controlling for age, income, and marital status.¹⁶⁶ The insight drawn from the study was that scarcity of supply in the brick-and-mortar world did not reduce the demand for pornography.¹⁶⁷

Julie Cohen faults intellectuals (like me) who assume that more information will lead to more truth.¹⁶⁸ In Cohen's opinion, advocates of the data revolution inadvertently promote a corporate Big Data ideology that promotes stereotyping and categorization and may even produce self-fulfilling prophecies.¹⁶⁹ Cohen's description of a monolithic Big Data culture is badly misinformed. Any data analyst worth his keep has to be humble and open-minded, and will abandon the stereotypes that tend to be perpetuated and overvalued in society.¹⁷⁰ The data analyst will do so not in the hope that his predictive algorithm has any possibility of being perfectly accurate, but in the hope that it will

164. *Id.*

165. See Benjamin Edelman, *Red Light States: Who Buys Online Adult Entertainment?*, J. ECON. PERSP., Winter 2009, at 209, 216, 219.

166. *Id.* at 217; see also Dawn House, *Utah Is No. 1—for Online Pornography Consumption*, SALT LAKE TRIB. (Mar. 2, 2009, 7:39 PM), http://www.sltrib.com/ci_11821265 (noting Utah also has the second-highest ranking for Google searches of the words “hot sex” and “naughty”).

167. Luis Urbina, *Fifteen Questions with Benjamin G. Edelman '02*, HARVARD CRIMSON (Mar. 10, 2009), <http://www.thecrimson.com/article/2009/3/10/fifteen-questions-with-benjamin-g-edelman>.

168. See Julie E. Cohen, *What Privacy Is for*, 126 HARV. L. REV. 1904, 1919-22 (2013).

169. As she notes:

Big Data is the ultimate expression of a mode of rationality that equates information with truth and more information with more truth, and that denies the possibility that information processing designed simply to identify “patterns” might be systematically infused with a particular ideology. . . . But the denial of ideology is itself an ideological position.

Id. at 1924. One might think that Cohen is making an epistemological relativist argument of the “lies, damn lies, and statistics” sort, suggesting that no data methodology can have any greater claim to truth than any other. But she acknowledges that data analysis is able to produce increased “predictive rationality.” Instead, Cohen argues that equating accurate description with “truth” is a value judgment, and one that should be questioned. See *id.* at 1925.

170. To the extent data analytics has an ideology at all, it is the “Money Ball” ideology: a willingness and readiness to reject the prevailing stereotypes and conventional wisdom in the face of countervailing evidence. See MICHAEL LEWIS, *MONEY BALL: THE ART OF WINNING AN UNFAIR GAME* (2003).

be a little less inaccurate.¹⁷¹ To have a chance at bucking the conventional stereotypes and improving knowledge, he needs data.

A person's access to accurate data is no guarantee that he will learn anything new. The experiments of Daniel Kahneman, Dan Kahan, and many others have shown that confirmation biases and other heuristics lead people to cherry-pick among bits of data to confirm their preexisting beliefs.¹⁷² Worse still, human confirmation biases are so deeply ingrained that we have cognitive biases to keep us from recognizing our own cognitive biases, and more biases to prevent us from recognizing those biases, and so forth.¹⁷³ Thus, people on opposite sides of the debate on whether global warming is caused by humans become *more* confident, and more entrenched in their positions, the more accurate information they have.¹⁷⁴

But biases do have limits. Accurate information increases the cost of maintaining a false belief. Eventually, accurate information reaches a tipping point and is able to overwhelm and correct the false beliefs of all but the most stubborn. Thus, after Harold Camping's first prediction of the end of the world

171. As Nate Silver has put it, increasing data will allow a skilled forecaster to tune his predictive model to be “*less subjective, less irrational, and less wrong.*” NATE SILVER, *THE SIGNAL AND THE NOISE: WHY SO MANY PREDICTIONS FAIL—BUT SOME DON’T* 259 (2012).

172. DANIEL KAHNEMAN, *THINKING, FAST AND SLOW* 80-81 (2011); Dan M. Kahan et al., “*They Saw a Protest*”: *Cognitive Illiberalism and the Speech-Conduct Distinction*, 64 *STAN. L. REV.* 851, 883-84 (2012); Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 *REV. GEN. PSYCHOL.* 175, 175 (1998).

173. See DAN ARIELY, *THE (HONEST) TRUTH ABOUT DISHONESTY: HOW WE LIE TO EVERYONE—ESPECIALLY OURSELVES* 25-27 (2012) (describing experiments showing that people will cheat less when they must confront data showing that they have cheated—even if they have no risk of being caught); Robin Hanson, *Enhancing Our Truth Orientation*, in *HUMAN ENHANCEMENT* 359, 360-61 (Julian Savulescu & Nick Bostrom eds., 2009) (citing various works in the field); Thomas Goetz, *Harnessing the Power of Feedback Loops*, *WIRED* (June 19, 2011, 9:45 AM), http://www.wired.com/magazine/2011/06/ff_feedbackloop (describing an experiment with driver speed that confirms that documentation of one's behavior and confrontation with the data affects behavior by preventing willful self-blindness).

174. Dan M. Kahan et al., *The Polarizing Impact of Science Literacy and Numeracy on Perceived Climate Change Risks*, 2 *NATURE CLIMATE CHANGE* 732 (2012), <http://www.nature.com/nclimate/journal/v2/n10/full/nclimate1547.html>; see also PHILIP E. TETLOCK, *EXPERT POLITICAL JUDGMENT: HOW GOOD IS IT? HOW CAN WE KNOW?* 122-23 (2005) (describing an experiment showing that experts are unable to imagine possibilities if a theory rival to their own turns out to be correct); Louis Menand, *Everybody's an Expert: Putting Predictions to the Test*, *NEW YORKER* (Dec. 5, 2005), available at http://www.newyorker.com/archive/2005/12/05/051205crbo_books1 (describing Tetlock's conclusion that “[h]uman beings who spend their lives studying the state of the world, in other words, are poorer forecasters than dart-throwing monkeys”). For an example of a cat outperforming experts in the purchase of stocks, see Neetzan Zimmerman, *Ordinary Housecat Beats Professional Wealth Managers in Year-Long Stock-Picking Challenge*, *GAWKER* (Jan. 14, 2013, 10:25 AM), <http://gawker.com/5975736/ordinary-housecat-beats-professional-wealth-managers-in-year+long-stock+picking-challenge>.

proved to be incorrect on September 6, 1994, and after his second prediction also proved wrong on May 21, 2011, there were only twenty-five followers left to fret over his third prediction of the end of the world—October 21, 2011.¹⁷⁵ Even cult members are a little bit Bayesian.

Data can also illustrate the public good theory of the First Amendment because it tends to provide greater value to society at large than it does to any individual described by the data, who may naturally prefer to control and prevent its spread. Even the privacy rules embedded in the Health Insurance Portability and Accountability Act of 1996 (HIPAA) have costs that are often hidden from view, since data cannot easily be moved around for research purposes.

Adverse drug reactions provide an example.¹⁷⁶ Vioxx, the drug once prescribed for arthritis, was sold for over five years before its manufacturer, Merck, withdrew it from the market in 2004.¹⁷⁷ Though small-scale studies found a correlation between Vioxx and increased risk of heart attack, the FDA did not have convincing evidence until it analyzed data about 1.4 million HMO members.¹⁷⁸ By the time Vioxx was pulled, it had caused between 88,000 and 139,000 unnecessary heart attacks, and around 28,000 avoidable deaths.¹⁷⁹ The Vioxx debacle is a haunting illustration of the importance of large-scale data research. The FDA explored possible “what if” scenarios during a 2007 hearing. If researchers had been granted access to seven million patient records, one researcher estimated that the relationship between Vioxx and heart attacks would have been clear in under three years.¹⁸⁰ With access to 100 million records, it would have been discovered in just three months.¹⁸¹ HIPAA allows for the nonconsensual sharing of medical records for research purposes under certain strict conditions,¹⁸² but the friction caused by these regulations has real consequences.

175. See *Judgment Day Doesn't Dawn*, NEWS24 (Oct. 21, 2011, 11:06 AM), <http://www.news24.com/World/News/Judgment-Day-doesnt-dawn-20111021>.

176. I am indebted to Barbara Evans, who turned my attention to this really cool, illustrative example. She has described the Vioxx study in her own work. See Barbara J. Evans, *Seven Pillars of a New Evidentiary Paradigm: The Food, Drug, and Cosmetic Act Enters the Genomic Era*, 85 NOTRE DAME L. REV. 419, 455-56 (2010).

177. Barnaby J. Feder, *Merck's Actions on Vioxx Face New Scrutiny*, N.Y. TIMES (Feb. 15, 2005), <http://www.nytimes.com/2005/02/15/business/15merck.html>.

178. Rita Rubin, *How Did Vioxx Debacle Happen?*, USA TODAY (Oct. 12, 2004), http://usatoday30.usatoday.com/news/health/2004-10-12-vioxx-cover_x.htm.

179. FDA, *Merck, and Vioxx: Putting Patient Safety First?: Hearing Before the S. Comm. on Fin.*, 108th Cong. 13-14 (2004) (statement of David J. Graham, Associate Director for Science, Office of Drug Safety, FDA).

180. Evans, *supra* note 176, at 456 & nn.250-51.

181. *Id.*

182. Data can be shared with the broader research community only if the data is de-identified in compliance with 45 C.F.R. § 164.514(a)-(b). This usually hampers researchers from linking records between health providers. Alternatively, a health provider may go through the steps of creating a “limited data set” that removes all direct identifiers (again

2. *Deliberative democracy and a check on state power*

The third and fourth categories view the First Amendment as an institutional check on the tyranny of the majority and on government actors, respectively. Under these theories, free speech is compromised when the government regulates speech that empowers political dissenters, or that for some other reason it does not care for.¹⁸³

On the surface, data seems to have little in common with archetypal political speech—protests, op-eds, and the like. Privacy scholars often argue that personal information is wholly different from the types of “core political speech” that most deserve constitutional protection.¹⁸⁴ Likewise, some First Amendment scholars agree that the right to free speech is meant to protect dissident opinions most ardently, with its protections dropping off as expression becomes less political.¹⁸⁵

This protester-centric vision for the First Amendment is occasionally echoed by courts,¹⁸⁶ but it is a cramped and unsatisfying version of free speech. It would be unrecognizable to Louis Brandeis. While Justice Brandeis promoted a theory of the First Amendment that ensured countermajoritarian messages could be heard and received by others, he was not concerned exclusively, or even primarily, with political speech. Justice Brandeis envisioned speech rights protecting the individual development and learning we all do privately in order to better contribute publicly.¹⁸⁷ “[T]he final end of the State was to make men

preventing any linkages with other health provider records), and that also is accompanied by a data release agreement in compliance with 45 C.F.R. § 164.514(e).

183. See ALEXANDER MEIKLEJOHN, *POLITICAL FREEDOM: THE CONSTITUTIONAL POWERS OF THE PEOPLE* 26-27 (1960). The deliberative democracy theories best explain the emphasis on content neutrality as a trigger for heightened First Amendment scrutiny. This idea, and its shortcomings, are explored in Steven J. Heyman, *Spheres of Autonomy: Reforming the Content Neutrality Doctrine in First Amendment Jurisprudence*, 10 WM. & MARY BILL RTS. J. 647, 650-53 (2002).

184. See Bhagwat, *supra* note 15; Richards, *supra* note 14; Solove, *supra* note 22, at 984 (concluding that not all forms of speech have high value, but also embracing a balancing approach to speech and privacy).

185. See, e.g., Robert H. Bork, *Neutral Principles and Some First Amendment Problems*, 47 IND. L.J. 1, 20 (1971) (arguing that only overt political speech should receive First Amendment protection, and not all of it); Weinstein, *supra* note 152, at 492-93.

186. See *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 759 (1985); *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 343 (1974). But the Supreme Court has also interpreted “public concern” expansively at times. For example, in *Florida Star v. B.J.F.*, the name of a rape victim (and not just the facts surrounding the rape) was held to have a sufficient nexus to public concern. 491 U.S. 524 (1989).

187. Louis Brandeis was influenced by the philosophy of John Dewey, who believed learning was ultimately a social activity, with significant interactive effects:

His focus on free speech as part of the mutual and potentially reinforcing relationship between the individual and society closely resembled Dewey’s postwar analysis of free speech, as Dewey himself recognized. For Brandeis, as for Dewey, by protecting free speech the state liberates individuals, who in turn contribute to society.

free to develop their faculties It is the function of speech to free men from the bondage of irrational fears.”¹⁸⁸ Both Justices Holmes and Brandeis were heavily influenced by the ideas of John Dewey, who believed that very private and internal acts of learning formed the base of democratic liberalism.¹⁸⁹ So, to Justice Brandeis at least, deliberative-democracy goals were closely linked to free thought and self-determination.¹⁹⁰

But even if the First Amendment were constrained to political speech, raw information has a claim to its protection. Data can lead to the discovery or evidence of government scandals. For example, most logs of IP addresses will have no relation to political debate, but the IP addresses used to access a shared e-mail account are highly political when they happen to belong to David Petraeus and his biographer, Paula Broadwell.¹⁹¹ And collectively, data can unearth scandals and phenomena that could not be discovered without it (as when Justin Wolfers used sports-wagering data to discover that college basketball players accept bribes to play worse¹⁹²). Any type of data, if there’s enough of it, can be used to mine for insights.

Consider *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, the Supreme Court opinion that first recognized a distinction between core speech of “public concern” and less crucial speech of “purely private concern.”¹⁹³ In *Dun & Bradstreet*, a credit reporting firm had sent a false credit report to a potential creditor of the plaintiff. The firm refused to adequately disclose the error to the

DAVID M. RABBAN, *FREE SPEECH IN ITS FORGOTTEN YEARS* 343 (1997). See generally JOHN DEWEY, *DEMOCRACY AND EDUCATION: AN INTRODUCTION TO THE PHILOSOPHY OF EDUCATION* (The Free Press 1966) (1916); JOHN DEWEY, *MY PEDAGOGIC CREED* (N.Y., E.L. Kellogg & Co. 1897).

188. *Whitney v. California*, 274 U.S. 357, 375-76 (1927) (Brandeis, J., concurring), *overruled in part by* *Brandenburg v. Ohio*, 395 U.S. 444 (1969). Justice Holmes joined Justice Brandeis’s concurrence in *Whitney*. *Id.* at 380.

189. See David Kennedy, *John Dewey, in THE CANON OF AMERICAN LEGAL THOUGHT* 113, 113-20 (David Kennedy & William W. Fisher III eds., 2006). More recently, Robert Post has recognized that political discourse must rely on access to accurate facts: “To preserve the self-government of the people, we must preserve their access to knowledge.” POST, *supra* note 36, at 95.

190. “For all his emphasis on the contributions individual free speech makes to democratic governance, Brandeis also believed that a democratic state has obligations to promote individuality, including its expression through free speech, as an end in itself.” RABBAN, *supra* note 187, at 370.

191. See Kim Zetter, *Email Location Data Led FBI to Uncover Top Spy’s Affair*, WIRE (Nov. 12, 2012, 2:17 PM), <http://www.wired.com/threatlevel/2012/11/gmail-location-data-petraeus>.

192. Wolfers discovered a statistical anomaly—teams that were heavily favored to win a game too frequently won the game by *just a little bit less* than the spread. See Justin Wolfers, *Point Shaving: Corruption in NCAA Basketball*, 96 AM. ECON. REV. (PAPERS & PROC.) 279 (2006).

193. 472 U.S. 749, 757-59 (1985) (“We have never considered whether the *Gertz* balance obtains when the defamatory statements involve no issue of public concern.”).

recipients of the erroneous credit report even though it had substantial evidence that the credit report was wrong.¹⁹⁴ In upholding an award for punitive damages for the disclosure of the false report, the Supreme Court distinguished credit reports from speech of public concern, such as the public political indictments that led to suits in *New York Times Co. v. Sullivan* and *Gertz*. Although the Court had once said punitive damages could not be assigned to speech—even false speech—it changed its mind in *Dun & Bradstreet* because the credit report concerned purely private speech that would not bear on public discourse.¹⁹⁵

The opinion could have been construed narrowly, to apply only to *false* speech of private concern, but it has not been understood that way. The Supreme Court has cited with approval the distinction between speech of “public or private concern” without regard to whether the speech is false or defamatory.¹⁹⁶ But today, as we grapple with the aftermath of a home mortgage crisis, credit reports are vital to the debate. Whether in the aggregate or as stand-alone examples, the details of credit reporting are highly relevant to understanding the intricate dance between lenders and borrowers that led to the collapse of an unsustainably leveraged housing market.

If the credit report example seems unique, consider climate science data. This data, too, has dipped in and out of the “public concern” category enough times to raise questions about the viability of the distinction. Throughout most of the twentieth century, temperature records across the country were presumably matters of public concern when they were first recorded. After all, the public has an obvious interest in the current temperature. But as time goes on, the public’s concern for these records wanes. Except for extreme temperature readings, the public does not have much interest in the weather reports from five years ago. But today, some of the most important climate science data is historical temperature and ecological information—information that was collected for a different purpose, before we were aware of global warming. And so, the same data has bounced back into public concern because it relates to a live political debate.

Robert Post, one of the strongest supporters of the “public concern” criterion, has alternatively defined its scope as “all efforts deemed normatively necessary for influencing public opinion,”¹⁹⁷ and as all “processes of public opinion formation.”¹⁹⁸ Post has to define speech of “public concern” capaciously so that it has the chance to cover cases in which the expression has a small audi-

194. *Id.* at 751-52.

195. *Id.* at 759-60.

196. *See, e.g.*, *Snyder v. Phelps*, 131 S. Ct. 1207, 1215 (2011).

197. POST, *supra* note 36, at 18.

198. *Id.* at 28.

ence, or an unfocused message—flag-burning cases, for example.¹⁹⁹ Its broad sweep also mitigates (though does not entirely avoid) putting discretion into the hands of the government to decide what is “normatively necessary,” or what is “proper” (a test that Dan Solove accepts²⁰⁰) for the public to know. But under Post’s formulation, it is difficult to imagine speech that would *not* be a matter of public concern. Even a narrow conversation between two people, or a private record kept by one person for that matter, has the potential to be amplified over time and eventually fertilize public debate. Thus these activities are arguably both “normatively necessary for influencing public opinion” and “processes of public opinion formation.”

Justice Brennan’s dissent in *Dun & Bradstreet* recognized the importance that all information—even seemingly mundane facts—has on self-government. “[T]he choices we make when we step into the voting booth may well be the products of what we have learned from the myriad of daily economic and social phenomenon that surround us.”²⁰¹ These economic and social phenomena are frequently revealed through data. Data on wages, electricity use, or home prices can be unrelated to public discourse at one moment only to come to the service of urgent public debate at the next. Just like talking, e-mailing, and other traditional forms of speech, data feeds the slow, often-messy process of public opinion formation. In Hannah Arendt’s words: “[F]actual truth informs political thought”²⁰²

3. *Self-Determination*

The autonomy or self-determination theories of the First Amendment have the most obvious overlap with the approach I’ve taken in this Article because the free development of the mind would, presumably, incorporate a freedom from a government’s constraints on the information we are able to access—whether that information is raw data or a book. But data’s relationship to self-determination is complex.

On the one hand, the ability to collect and mull over data is very helpful for independent thought. In the future, as we adjust to the information revolution, we may even find data to be essential to our thought processes. Future generations may marvel at the way we live today, wondering how a person can main-

199. *Id.* at 15. For the Court’s current approach to flag burning, see *United States v. Eichman*, 496 U.S. 310 (1990).

200. Solove, *supra* note 22, at 1000, 1016, 1026. Solove endorses the definition of “public concern” articulated by the Second Restatement of Torts: information that the “the public has a proper interest in learning about.” *Id.* at 1000 (quoting RESTATEMENT (SECOND) OF TORTS § 652D (1977)) (internal quotation mark omitted).

201. *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 788 (1985) (Brennan, J., dissenting).

202. HANNAH ARENDT, *BETWEEN PAST AND FUTURE: EIGHT EXERCISES IN POLITICAL THOUGHT* 238 (1968).

tain a decent quality of life without raw data about anything and everything, ready to be probed and analyzed so that his decisions can be optimized. Surely we ourselves struggle to imagine the quality of life for the generations of humans who had to live illiterately before the invention of Gutenberg's press.

On the other hand, privacy laws are developed to protect self-determination, too. Our freedom to develop as fully autonomous individuals will be stunted if our private information is on constant display to everybody else. Stripped of all privacy, people will naturally and rationally engage in the sort of self-restraint and self-censorship that serve neither themselves nor society at large. This inherent tension makes data privacy regulation difficult to calibrate. No doubt there are important, competing interests in privacy and access to information. At times, data access will pit liberty against liberty.

Privacy advocates and scholars have seized on the liberty-preserving aspects of privacy, but they've done so using generic platitudes that tend to overstate privacy's relationship to autonomy.²⁰³ Consider Julie Cohen's recent attempt to fortify the value of privacy against the pressures of encroaching technology.²⁰⁴ Privacy, according to Cohen, is a necessary precondition to liberty because it "shelters dynamic, emergent subjectivity from the efforts of commercial and government actors to render individuals and communities fixed, transparent, and predictable."²⁰⁵ Cohen goes on to say that surveillance by corporations in time creates a citizenry who lacks the capacity for democratic self-government. How we would reach these dire results is less than clear, but the idea seems to be that as corporations predict and model how we are using personal data, our lives will begin to imitate the very pigeonholes into which we've been put.²⁰⁶

Granted, humans are not as autonomous and immune to social pressures as they are assumed to be in classical accounts of liberal democracies. But every form of speech can be used to manipulate people. The consequences are part of

203. For example, Marc Rotenberg, executive director of the Electronic Privacy Information Center, asks: "Does the Obama administration really want to be on the opposite side of the European effort to upgrade and modernize its privacy law which is at its core about the protection of a *fundamental freedom*?" Natasha Singer, *Data Protection Laws, an Ocean Apart*, N.Y. TIMES (Feb. 2, 2013), <http://www.nytimes.com/2013/02/03/technology/consumer-data-protection-laws-an-ocean-apart.html> (emphasis added).

204. See Cohen, *supra* note 168, at 1905.

205. *Id.*

206. Certain modern data practices are, according to Cohen, "designed to produce a particular way of knowing and a mode of governance designed to produce a particular kind of subject. Its purpose is to produce tractable, predictable citizen-consumers whose preferred modes of self-determination play out along predictable and profit-generating trajectories." *Id.* at 1917. In fact, data analysts make conscious efforts to avoid "overfitting" the data and producing models that do not track actual behaviors and preferences. See, e.g., *Don't Overfit!*, KAGGLE, <http://www.kaggle.com/c/overfitting> (last visited Dec. 18, 2013). These methods anticipate that models based on past data will tend to exaggerate the predictability of people and correct the models by reintroducing uncertainty.

the elaborate free speech experiment created by the Constitution. The Bill of Rights casts a bet that the harms from assuming people have more free will than they really do are not as bad as the harms that come from assuming people have *less* than they really do.

Moreover, a person who is categorized in one instance will be the categorizer in the next and will rightly expect the liberty to judge and form his own opinions. To be sure, a person's data is often used to make determinations and decisions about him, which might, at least temporarily, constrain him as compared to a world without data. A creditor will size him up when setting his interest rate, and a prospective partner on an online dating service will use generalizations and stereotypes in order to assess whether he is worthy of a first date. But he, too, is free to gather data and make judgments about the best creditor or best prospective mate. As Frederick Schauer has argued, generalizing based on categories—based on the reductions that Cohen finds so demeaning—is not only unavoidable, but can be positively admirable in a world of constrained resources.²⁰⁷ Without it, we must choose either arbitrariness or unconscious, unaccountable generalizations. There are no other options.

Scholars and public intellectuals tap into their imaginations to predict how a person's autonomy can be limited when companies have drastically increased access to her personal data, but they are much less creative when it comes time to predict how a person's life can be improved.²⁰⁸

Consider salt. Most of us have long internalized the advice that we should reduce our consumption of salt to improve our health and avoid hypertension, but the connection between salt and heart disease (as well as stroke) has no basis in fact. The studies finding an association are laughably flawed. The first, which launched a century-long war on salt, was based on six French patients.²⁰⁹ More recent research shows no relationship or the opposite relationship; for example, one study found that subjects who had low levels of sodium in their urine had *increased* risk of heart disease.²¹⁰ In 2013, the Centers for Disease Control and Prevention (CDC) commissioned the Institute of Medicine (IOM) to assess the state of the research and determine whether reducing sodi-

207. See FREDERICK SCHAUER, *PROFILES, PROBABILITIES, AND STEREOTYPES* (2003) (laying out the case for generalizations and stereotypes for many (but not all) contexts).

208. See, e.g., VIKTOR MAYER-SCHÖNBERGER, *DELETE: THE VIRTUE OF FORGETTING IN THE DIGITAL AGE* (2009) (describing disturbing shifts in power as a result of increased information); ELI PARISER, *THE FILTER BUBBLE: WHAT THE INTERNET IS HIDING FROM YOU* (2011) (describing how data analytics will reduce our exposure to competing ideas); Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 *UCLA L. REV.* 1701, 1750-51 (2010) (describing how each of us will be beset by our own personal "databases of ruin").

209. Melinda Wenner Moyer, *It's Time to End the War on Salt*, *SCI. AM.* (July 8, 2011), <http://www.scientificamerican.com/article.cfm?id=its-time-to-end-the-war-on-salt> (describing a meta-analysis in the *American Journal of Hypertension*).

210. *Id.*

um intake to less than 2300 milligrams per day (the current CDC recommendation) has health benefits.²¹¹ The IOM produced a report finding no such evidence.²¹² Astonishingly, the CDC has ignored, even resented, the IOM's conclusions and has chosen to continue to recommend the reduction of salt consumption.²¹³ Perhaps a reversal in course would be too embarrassing for the CDC and the medical profession at large, or maybe it is too tempting to conclude that something that tastes so good must be bad for us. More likely, the CDC doesn't feel confident that the data available today can be used to draw any conclusions. Consumer data collected on food purchases and activity levels for a large population could quickly disentangle the effects of salt from the effects of eating junk foods (which tend to be loaded with salt as well as sugars and fats). But since we do not routinely pool consumer data and open it to research, many more generations of people may continue to avoid a flavoring that they like—indeed, a flavoring that can be used to make healthy foods more enjoyable—out of a false belief in the dangers of salt.²¹⁴

A recent article by Jean Twenge in the *Atlantic* unearths a similar problem plaguing the conventional wisdom on fertility.²¹⁵ Remember that *Time* article in 2002 that informed us about research finding that female fertility drops off dramatically at age thirty-five?²¹⁶ Those findings were based on French birth records from 1670 to 1830.²¹⁷ Just as more recent research debunks the theory that high salt intake leads to negative health outcomes, so too does better and newer research run counter to the theory that fertility shifts at age thirty-five, but many open questions loom. Imagine how much more we would understand about fertility if the buying patterns of women in consumer databases (like the Target database that received so much attention last year²¹⁸) were mined to re-

211. Mike Mitka, *IOM Report: Evidence Fails to Support Guidelines for Dietary Salt Reduction*, 309 JAMA 2535, 2535 (2013).

212. *Id.*

213. *See id.*

214. Likewise, if consumer data had been linked to medical records, we would have learned long ago that taking antioxidant supplements greatly increases the chances of dying from cancer. *See* Paul A. Offit, Op-Ed., *Don't Take Your Vitamins*, N.Y. TIMES (June 8, 2013), <http://www.nytimes.com/2013/06/09/opinion/sunday/dont-take-your-vitamins.html>.

215. *See* Jean Twenge, *How Long Can You Wait to Have a Baby?*, ATLANTIC (June 19, 2013, 10:05 PM EDT), <http://www.theatlantic.com/magazine/archive/2013/07/how-long-can-you-wait-to-have-a-baby/309374>.

216. *See* Nancy Gibbs, *Making Time for a Baby*, TIME, Apr. 15, 2002, at 48.

217. Twenge, *supra* note 215.

218. *See* Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES MAG. (Feb. 16, 2012), <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>; *see also* Neil M. Richards, *The Dangers of Surveillance*, 126 HARV. L. REV. 1934, 1939-40 (2013) (describing practices like Target's as "Big Data surveillance" and encouraging greater restrictions).

veal when they were trying to become pregnant and when they succeeded.²¹⁹ Analysis of this sort no doubt has a great “ick” factor. But what we stand to gain is significant. If age turns out to have less influence over fertility than previously thought, the options for men and women change quite a bit. New couples would not be pressured into having children by the ticking of a biological clock, and women in not-so-new relationships might make different choices about whether to stay with their mate or take a few more years to find a better match. These decisions arguably have greater effect on the autonomy and quality of life of a data subject than the foreboding feeling of being watched.

Once the mind is opened to the potential upshot of massive data collection, the enhancements to self-determination seem boundless. Consumer behavior is already being used to predict whether a person is likely to adhere to her doctor’s recommended course of treatment.²²⁰ In the future, apps might be developed to collect and analyze biometric data so that patients can be advised about their health and medical treatment by the minute. And automated data collection is already allowing employers to hire based on actual performance measures rather than the noisier signals of formal education and personal interviews (both of which can introduce bias based on race, class, and other factors).²²¹ In time, employers will be able to use metrics to design recruitment, pay, and promotion systems that are increasingly based on merit.²²² In 2011 and 2012, over 100 U.S. children were injured or killed by products that had already been recalled.²²³ With more permissive data laws, notice of recalled

219. The personal data collected from the iPhone app Glow can be mined to learn more about fertility, too. Glow offers advice to couples who are trying to conceive based on the wisdom of medical consultants as well as, the hope is, crowdsourced wisdom from other users of the app. Harry McCracken, *Glow: An iPhone App That Aims to Get You Pregnant*, TIME (Aug. 8, 2013), <http://techland.time.com/2013/08/08/glow-an-iphone-app-that-aims-to-get-you-pregnant>. In the future, the developers hope to use passive collections of data that do not rely on users’ interactions with the app. *Id.*

220. See Tara Parker-Pope, *Keeping Score on How You Take Your Medicine*, N.Y. TIMES WELL BLOG (June 20, 2011, 5:23 PM), <http://well.blogs.nytimes.com/2011/06/20/keeping-score-on-how-you-take-your-medicine>.

221. See Matt Richtel, *How Big Data Is Playing Recruiter for Specialized Workers*, N.Y. TIMES (Apr. 27, 2013), <http://www.nytimes.com/2013/04/28/technology/how-big-data-is-playing-recruiter-for-specialized-workers.html?pagewanted=all>.

222. We may also learn that some industries are better or worse than we think in terms of rewarding merit. Vivienne Ming, the chief scientist of Gild (a headhunting firm that uses data to find talent in the rough), believes that Silicon Valley firms are less open minded than their reputation suggests. *Id.* Meanwhile, some researchers are finding that law firms, with their famously stuffy reputation for erecting glass ceilings for women on the partnership track, in fact treat men and women equally when their work-family commitments are the same. See Kenneth G. Dau-Schmidt et al., *Men and Women of the Bar: The Impact of Gender on Legal Careers*, 16 MICH. J. GENDER & L. 49 (2009).

223. See SARAH JOHNSTON, KIDS IN DANGER, TWO YEAR CHECK-UP: SAFERPRODUCTS.GOV 7 (2013), available at http://www.kidsindanger.org/docs/research/SaferProducts_Report.pdf.

products and food could be delivered directly to the purchasers by e-mail or text message.

Clearly, data can be used to increase the autonomy of its subjects. In time, we might come to realize that it would be irresponsible for a company *not* to collect and analyze the trails of data that consumers leave everywhere they go. Ariel Porat and Lior Strahilevitz argue that Big Data should be used to customize informed consent in medical procedures and consumer contracts so that they better fit the likely interests and preferences of the patient and consumer.²²⁴ These would be individuality-respecting and autonomy-enhancing improvements over how we live now.

Even so, there are limits to the autonomy-enhancing value of data. If we had no control at all over who could observe us, when, or why, our ability to act authentically would be constrained. Some tensions between liberty-preserving privacy and liberty-preserving knowledge gathering are unavoidable. But these tensions should be resolved within First Amendment scrutiny, not in deciding the First Amendment's coverage.²²⁵ After all, the creation and access to impersonal data (chemistry research or astronomy data, for example) promote self-determination through knowledge without any of the competing concerns for sufficient privacy protection.

C. *The Level of Scrutiny*

This Article cannot determine what level of scrutiny should apply to regulations of data flow because the answer will depend on context. Data disseminated in an advertisement, for example, will receive the lesser protections afforded to commercial speech under the *Central Hudson* test just like any other advertising speech.²²⁶ Also, like other forms of speech, data could presumably be categorized as a matter of “public” or “purely private” concern, which would affect the level of scrutiny. The divisions between commercial and non-commercial speech, and between public and private concern, have drawn criticism from First Amendment scholars because of problems with their consistency and administrability.²²⁷ The doctrines will be as challenging to apply to data as they are to other types of expression.

224. Ariel Porat & Lior Jacob Strahilevitz, *Personalizing Default Rules and Disclosure with Big Data*, 112 MICH. L. REV. (forthcoming 2014), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2217064.

225. In Part IV, I discuss compelling interests in seclusion and confidentiality, which should be capable of overcoming First Amendment scrutiny (assuming the law is narrowly tailored).

226. *Cnt. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n*, 447 U.S. 557, 564 (1980) (permitting restrictions of commercial speech if the state can assert a substantial interest to be achieved by the restriction, and the regulation is tailored to that interest).

227. See, e.g., Ashutosh Bhagwat, *The Test That Ate Everything: Intermediate Scrutiny in First Amendment Jurisprudence*, 2007 U. ILL. L. REV. 783, 785; Volokh, *supra* note 19, at

What can be said with confidence is that data should not be relegated in all cases to a lower form of protection.²²⁸

It may be tempting to exploit divisions between commercial and non-commercial speech and insist that data collected, maintained, or transmitted by a corporation is *always* commercial. The commercial speech designation may even seem to correspond nicely with a First Amendment that privileges free thought, since it is the human mind that deserves sanctity, and not the corporate one.²²⁹ Seana Shiffrin makes this point, concluding that “protection for commercial and non-press, business corporate speech is a less central matter, one that reasonably may involve weaker protections.”²³⁰

Proponents of privacy regulations have made the same human-corporate distinction, since the corporate actors generating the largest collections of data—Big Data businesses—are also believed to be the most likely to abuse their informational advantage and harm consumers. But caution is warranted. Shiffrin sensibly recognizes that corporations who trade in communications—that is, the press—cannot be regulated under lesser constitutional scrutiny. A corporation that sells books and newspapers has significant First Amendment interests in avoiding regulation of its product. Whether these interests are direct or derivative of the authors’ and readers’ rights is an interesting but purely academic matter. The press gets protection. The natural follow-up question, one that Shiffrin’s work did not have reason to explore in any detail, is what it means to be a “press” business corporation.

A corporation that generates and subsequently uses or sells data, even if the revenue stream is ancillary to its primary product or service, has a cognizable argument that it is in the business of communications, and is therefore analogous to a traditional press corporation. LexisNexis is an example of a corporation that should at least challenge the assumption that “the press” is a static concept; Lexis is in the business of aggregating traditional forms of speech (newspaper articles, laws, and court opinions) and creating original content for

1095-96; Eugene Volokh, *The Trouble with “Public Discourse” as a Limitation on Free Speech Rights*, 97 VA. L. REV. 567, 567-68 (2011).

228. Ashutosh Bhagwat has argued that regulating disclosures of data should invariably draw intermediate scrutiny for regulations of purely private speech since “[t]he disclosure of large amounts of data, especially personal data, generally has no real connection to self-governance.” Bhagwat, *supra* note 15, at 876. This argument is addressed under the deliberative democracy theories discussed above in Part III.B.2.

229. It is worth reflecting on the ways in which human memory and decisionmaking have become more similar to those of a corporation. Technologies have allowed us to outsource the memorization of information to our cell phones and the Internet and to make algorithmic decisions that require the processing power of a computer. See Erez Reuveni, *Copyright, Neuroscience, and Creativity*, 64 ALA. L. REV. 735, 766-68 (2013). Cass Sunstein describes human knowledge as a dispersed network of information and inferences, with various types of errors and inefficiencies. See CASS R. SUNSTEIN, *INFOTOPIA: HOW MANY MINDS PRODUCE KNOWLEDGE* (2006).

230. Shiffrin, *supra* note 114, at 286.

a fee. This is entirely standard “press” sort of business. But it also maintains one of the largest databases of personal information culled from private and public records, and it sells components of this database to clients and customers.²³¹ Lexis’s trade in data is not meaningfully different from its trade in other types of information and cannot be differentiated from its “press” functions without creating limitations on the meaning of “the press” that have not previously been recognized.²³² Other times, some of the most socially valuable and useful information is generated in the course of other nonspeech endeavors. The results of pharmaceutical clinical trials,²³³ the outcomes of standardized tests,²³⁴ or the customer lists of Swiss banks,²³⁵ for example, may have much more in common with political speech than they do with advertisements depending on how they are used.

Debates over corporate speech divide First Amendment scholars into two camps. Those who regard the First Amendment as protecting individual liberties through negative rights are, generally speaking, nonplussed by the extension of First Amendment rights to corporations. On the other hand, those who subscribe to what Kathleen Sullivan calls the “equality-based” model of the First Amendment do not see the need to protect corporate speakers who, at least as a group, have the wealth and power necessary to ensure their message will be heard.²³⁶ The Supreme Court not so long ago voiced its concern about the “corrosive and distorting effects of immense aggregations of wealth” as a justification for regulations of corporate speech.²³⁷

The equality camp is no doubt dismayed by the Supreme Court’s expansion over the years of corporate speech rights (recognized most emphatically in *Citizens United*²³⁸). The perils of corporate speech are seen as redoubled when corporations also have unprecedented amounts of information about their consumers and the public at large. Corporate collections of personal data are be-

231. *LexisNexis Public Records*, LEXISNEXIS, <http://www.lexisnexis.com/en-us/products/public-records.page> (last visited Dec. 18, 2013).

232. See Eugene Volokh, *Freedom for the Press as an Industry, or for the Press as a Technology? From the Framing to Today*, 160 U. PA. L. REV. 459 (2012) (exploring the original meaning of the Press Clause and its relevance to today’s news environment).

233. See Ben Goldacre, Op-Ed., *Health Care’s Trick Coin*, N.Y. TIMES (Feb. 1, 2013), <http://www.nytimes.com/2013/02/02/opinion/health-cares-trick-coin.html>.

234. See Anemona Hartocollis, *Racial Gap in Test Scores Found Across New York*, N.Y. TIMES (Mar. 28, 2002), <http://www.nytimes.com/2002/03/28/nyregion/racial-gap-in-test-scores-found-across-new-york.html?pagewanted=all&src=pm>.

235. See Liz Alderman, *Greek Editor Is Arrested After Publishing a List of Swiss Bank Accounts*, N.Y. TIMES (Oct. 28, 2012), <http://www.nytimes.com/2012/10/29/world/europe/greek-editor-arrested-after-publishing-list-of-swiss-bank-accounts.html>.

236. Kathleen M. Sullivan, *Two Concepts of Freedom of Speech*, 124 HARV. L. REV. 143, 144-45 (2010).

237. *Austin v. Mich. State Chamber of Commerce*, 494 U.S. 652, 659-60 (1990), *overruled by Citizens United v. FEC*, 558 U.S. 310 (2010).

238. 558 U.S. 310.

lieved to so wildly skew the existing imbalance of power between corporations and their consumers that they create, in the words of Mike Madison, “a whole new ballgame.”²³⁹ Frank Pasquale believes that Google has grown so powerful that it is a “de facto lawmaker” on the Internet.²⁴⁰ And Ashutosh Bhagwat concludes that privacy regulations are necessary to protect citizens against the (unspecified) “specific and tangible harms” of corporate data holders.²⁴¹

The intuitions of Madison, Pasquale, and others ally with the old adage that “knowledge is power.” But there is a limit to this logic. Information is useful, but it is not, on its own, coercive.²⁴² Moreover, corporations do not have the dominance that academics often ascribe to them. Half of the corporations that existed in 1980 are no longer around—a happy sign of our economic and scientific vitality.²⁴³ Among corporations, those in the communications and high-tech sectors may stumble quickest of all. A few memorable anecdotes show that the illuminati tend to overreact to corporate power. Thirty years ago, AT&T was fingered as the corporation that had too much access to personal communications data as well as too great a share of the market in communications technologies. Its steady decline since then has had much more to do with

239. Mike Madison, *Big Data Democracy*, MADISONIAN.NET (Nov. 14, 2012), <http://madisonian.net/2012/11/14/big-data-democracy>.

240. Frank Pasquale, *Beyond Innovation and Competition: The Need for Qualified Transparency in Internet Intermediaries*, 104 NW. U. L. REV. 105, 112 (2010).

241. Bhagwat, *supra* note 15, at 877-78. Bhagwat argues that, because privacy laws are designed to protect citizens rather than to manage the reputation of the government, these laws should be treated categorically differently than other speech restrictions and entitled to less scrutiny. *Id.* He makes no attempt to explain how his proposal can be squared with the plethora of First Amendment cases pairing the torts of public disclosure and intentional infliction of emotional distress. *See, e.g.*, *Snyder v. Phelps*, 131 S. Ct. 1207 (2011) (finding a First Amendment limitation on the tort of intentional infliction of emotional distress); *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46 (1988) (same); Daniel J. Solove & Neil M. Richards, *Rethinking Free Speech and Civil Liability*, 109 COLUM. L. REV. 1650, 1656-1660 (2009) (describing First Amendment limitations on the torts of defamation and public disclosure). Owen Fiss’s theory that the government can regulate private speech to ensure that it does not “impoverish public debate” would also be relevant to scholars endorsing exceptions for privacy from corporate data collection. *See* OWEN M. FISS, *THE IRONY OF FREE SPEECH* 42 (1996). The trouble is that Fiss’s theory of the First Amendment invites courts to legitimize restrictions on disfavored speech by framing the restriction in the amorphous “impoverish[ment]” language. Nelson Tebbe argues that Fiss’s approach has special relevance to government speech. Thus, impoverishment may not be a sufficiently strong check against a self-serving or overreaching government. The theory might be most useful and appropriate, however, in the context of the government’s own speech. *See* Nelson Tebbe, *Government Nonendorsement*, 98 MINN. L. REV. 648 (2013).

242. Shiffrin justifies heightened protections for speech precisely because speech is simultaneously more exact and precise in its message, and more respecting of others’ autonomy, than “a wordless punch in the nose” could ever be. Shiffrin, *supra* note 114, at 305.

243. *See* RIDLEY, *supra* note 50, at 111.

rapidly advancing technologies than antitrust lawsuits.²⁴⁴ Twenty years ago Microsoft looked like the greatest corporate threat to democracy, with the exact same results.²⁴⁵ Ten years ago we began to worry about Google, and now Facebook is a contender. Half of today's largest corporations did not exist thirty years ago,²⁴⁶ and if they do not harness the power of data to make themselves more useful and more efficient to their consumers, they will cease to exist in the next thirty years. As a descriptive matter, information is less a tool for corporate domination and more a tool for mere survival.

Thus, one should view with a healthy amount of skepticism Microsoft's, Apple's, and Google's purportedly altruistic interests in "Do Not Track" and other regulations that keep the information environment static.²⁴⁷ As innovative as each of these companies has been, they are not likely to be tomorrow's innovators. Maintaining the status quo through privacy law can help slow the inevitable rise of some new, unknown company that can make even better use of consumer data.

The concerns I describe here do not necessarily justify the use of heightened scrutiny in all cases; the contours of First Amendment scrutiny are complex and deserve careful consideration. Instead, these concerns are meant to raise doubts about the proposition that data automatically should receive less constitutional protection based on a generalized sense that it is "commercial," or of "purely private concern," or produced by corporations.

Whatever the level of scrutiny, a First Amendment right to create knowledge will lead to some consequences that are difficult to accept: the leveling of popular consumer privacy laws, long-established trade secrets protections, or all-party-consent wiretap statutes. The next Part addresses the most obvious objections and concerns.

244. See ROBERT W. CRANDALL & CHARLES L. JACKSON, TECH. POLICY INST., ANTITRUST IN HIGH-TECH INDUSTRIES 20 (2010), available at http://techpolicyinstitute.org/files/crandalljackson%20antitrust_in_high_tech3.pdf; Editorial, *Technology Threatens Monopolies*, N.Y. TIMES (Mar. 1, 1997), <http://www.nytimes.com/1997/03/01/opinion/technology-threatens-monopolies.html>.

245. A cynic might wonder if Microsoft is leading the charge for increased privacy (e.g., by enabling the "Do Not Track" option in Internet Explorer 10 by default) because it knows it cannot compete with its competitors on the basis of data analysis and innovation. See Natasha Singer, *When the Privacy Button Is Already Pressed*, N.Y. TIMES (Sept. 15, 2012), <http://www.nytimes.com/2012/09/16/technology/in-microsofts-new-browser-the-privacy-light-is-already-on.html?pagewanted=all>.

246. RIDLEY, *supra* note 50, at 111.

247. See David Meyer, *Google's Chrome Finally Embraces Do Not Track, but with a Warning*, ZDNET (Nov. 7, 2012, 1:03 PST), <http://www.zdnet.com/googles-chrome-finally-embraces-do-not-track-but-with-a-warning-7000007022> (noting that Internet Explorer and Safari, developed by Microsoft and Apple, respectively, have already incorporated a "Do Not Track" option).

IV. COPING WITH SCRUTINY

Many insightful scholars have contemplated the prospect of a First Amendment application to data and have concluded that the results are so absurd, and so debilitating to human progress, that the premise simply cannot be correct. At the very least, they have said, it cannot be just.²⁴⁸

This Part examines the potentially troubling implications of a right to knowledge creation. The first, and most obvious, is that the consequences of First Amendment scrutiny are just too restrictive to bear. The consequences to existing and proposed privacy legislation will wreak havoc on consumers and vulnerable populations.²⁴⁹ To the contrary, the most alarming privacy problems can be addressed through laws that are narrowly tailored to compelling interests in seclusion and confidentiality and thus survive scrutiny at any level. Regulators are also free to craft restrictions on conduct that treat people differentially based on personal data without interfering with the actor's access to information. While it is true that the First Amendment will prohibit sweeping data privacy laws, these consequences are not so perverse as to merit a deliberate unraveling of First Amendment commitments.

A. *Concern 1: Scrutiny Will Kill Privacy and Other Good Things*

Constitutional scrutiny threatens to thwart a lot of law both produced through the democratic process and representative of the better judgment of the majority. This raises an entirely legitimate concern: can we handle the consequences of nearly free data flow?

If courts abandon the distinction between information collection and information disclosure, which I think they must, they will be faced with the daunting task of drawing a principled line around information-gathering practices that excessively interfere with individual liberty. Though property rules can do some of the work by imposing liability for trespasses, modern technology allows conversations to be intercepted, driving routes to be traced, and DNA to be sequenced without breaking any generally applicable property rules. Courts will have to balance society's competing interests in seclusion and information production. This will be a challenging task, but a worthwhile one. It is the cost of a fundamental right to perceive the world and create new knowledge without unjustified interference. The government does not have the authority to curate the information we receive.

Data is not the first form of speech to cause problems, but the problems data causes occur on a vastly different scale. Because modern gadgets produce rich data trails, we are surrounded by recordings of facts that were too minute

248. *See supra* notes 14-17.

249. These are the primary concerns of Neil Richards, Shubha Ghosh, and Ashutosh Bhagwat. *See supra* notes 14-15.

to be noticed, let alone written down, before the computerized era. But today, without interfering with any generally applicable rules, a person can intercept a cell phone or Internet communication of another. Or he can hack into the other's bank website or medical records portal and receive information that way. Or he can swab the other's disposed coffee cup and analyze his genotype. With the exception of the coffee cup swabbing, all of these activities run afoul of federal criminal statutes.²⁵⁰ But these statutes are not generally applicable laws; they are designed for the singular purpose of interfering with the actor's access to new knowledge. If these types of data protections have the potential to draw constitutional scrutiny, courts will have to either justify the regulation on the basis of an important public interest or allow the information collection to proceed. The thought that courts may choose the latter is, undoubtedly, a dreary prospect.

These fears are far-fetched. The judiciary certainly *could* mangle scrutiny and annihilate personal privacy just as it could have found that time, place, and manner restrictions of speech are unconstitutional. But both path dependency and common sense suggest that they will not. As Justice Frankfurter said: "Free speech is not so absolute or irrational a conception as to imply paralysis of the means for effective protection of all the freedoms secured by the Bill of Rights."²⁵¹ To ensure liberty, courts have crafted some rules that interrupt our otherwise transparent society. Some of these rules have already survived scrutiny, and others will too.²⁵²

The right to seclusion serves as a good example. Seclusion provides respite from observation and judgment. Seclusion serves a variety of social goals: It allows us to engage in productive secrets. Seclusion is where a person can practice and fail in peace. In the words of Ralph Waldo Emerson: "Solitude, the safeguard of mediocrity, is to genius the stern friend . . ."²⁵³ Since some amount of assured privacy is required in order for thoughts and ideas to breed, Paul Schwartz has suggested that seclusion has some claim to serving a First Amendment purpose.²⁵⁴

Seclusion has already proved to be a sufficiently compelling interest to survive First Amendment scrutiny. Laws protecting seclusion have been found to "serve the undisputedly substantial public interest in allowing each person to

250. The first example violates Title III, which prohibits wiretapping outside of several narrowly defined circumstances. *See* 18 U.S.C. § 2511 (2012). The second example violates the Computer Fraud and Abuse Act. *See* 18 U.S.C. § 1030 (2012).

251. *Bridges v. California*, 314 U.S. 252, 282 (1941) (Frankfurter, J., dissenting).

252. Elsewhere I have championed the significant societal interests in confidentiality and the public disclosure tort, provided that these routes to recourse are narrowly tailored to target truly harmful information flows. *See* Bambauer, *supra* note 24, at 262-68.

253. RALPH WALDO EMERSON, *THE CONDUCT OF LIFE* 76 (H.G. Callaway ed., Univ. Press of Am. 2006) (1860).

254. Paul M. Schwartz, *Free Speech vs. Information Privacy: Eugene Volokh's First Amendment Jurisprudence*, 52 *STAN. L. REV.* 1559, 1572 (2000).

maintain an area of physical and sensory privacy in which to live.”²⁵⁵ However, the need and expectation of seclusion cannot go far beyond the home, private conversations, and other narrow circumstances. Otherwise the diminishing returns of seclusion will have increasingly severe effects on the liberty of others.²⁵⁶

Seclusion is not the only legitimate basis for limiting the creation or dissemination of data. The public’s interest in the confidentiality of certain special relationships, such as between a doctor and a patient, can be a compelling reason to limit the dissemination of information because, in the case of health care relationships, public health improves if patients can be honest and forthright with their doctors.²⁵⁷ And the public’s interest in the enforcement of anti-discrimination laws can justify a narrowly drawn restriction on the sorts of information that an employer can ask a job candidate to provide.

The First Amendment is unlikely to be interpreted so inflexibly as to kill off privacy. The much more likely scenario is that privacy will help save the First Amendment from developing a reputation of impossible, impassable standards of strict scrutiny. In 1972, Gerald Gunther made an observation that would become a mantra for legal scholars. He noticed that the outcomes of constitutional cases suggested that strict scrutiny is “‘strict’ in theory and fatal in fact.”²⁵⁸ Gunther’s hypothesis may better reflect perceptions than reality. (Adam Winkler’s analysis of 222 speech cases resolved under strict scrutiny found that 22% of the regulations survived review.)²⁵⁹ Nevertheless, a finding by the Supreme Court that there is a core of basic privacy interests capable of overcoming scrutiny would put First Amendment doctrine in a sustainable position. Until then, as long as scholars, advocates, and jurists believe that an important law could never survive a speech analysis, they will naturally turn their efforts to arguing that the law doesn’t technically regulate speech at all.

That said, a right to create knowledge will have casualties. Some existing privacy laws should not be able to withstand constitutional scrutiny. For example, President Obama’s proposal for a “Consumer Privacy Bill of Rights” would give consumers exclusive control over personal data that describes them, thereby extinguishing all competing interests of potential observers.²⁶⁰ The

255. *Shulman v. Group W Prods., Inc.*, 955 P.2d 469, 495 (Cal. 1998).

256. Heidi Anderson has begun to define the core interests of seclusion. See Heidi Reamer Anderson, *Plotting Privacy as Intimacy*, 46 IND. L. REV. 311 (2013).

257. Bambauer, *supra* note 24, at 262-65.

258. Gerald Gunther, *The Supreme Court, 1971 Term—Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection*, 86 HARV. L. REV. 1, 8 (1972).

259. Adam Winkler, *Fatal in Theory and Strict in Fact: An Empirical Analysis of Strict Scrutiny in the Federal Courts*, 59 VAND. L. REV. 793, 845 (2006).

260. See Press Release, White House, Fact Sheet: Plan to Protect Privacy in the Internet Age by Adopting a Consumer Privacy Bill of Rights (Feb. 23, 2012), available at

breadth of the proposal belies any tie to an important governmental interest except for a general interest in privacy.

Other American privacy laws are inspired by an important public interest in confidentiality or seclusion in specific contexts, but may fail the tailoring analysis.

Consider the Fair Credit Reporting Act (FCRA), a federal law that penalizes anybody who accesses another's credit report for any reason outside a preset list of authorized uses (e.g., to determine creditworthiness).²⁶¹ The D.C. Circuit already decided that the FCRA survives intermediate scrutiny in *Trans-Union Corp. v. Federal Trade Commission*.²⁶² The court appropriately applied intermediate scrutiny in the case because the credit-reporting company claiming a First Amendment right had compiled marketing lists with the intent that the lists would be used for solicitations.²⁶³ Applied to these facts, the FCRA appears to be appropriately tailored to the public's interest in financial confidentiality. But the facts of another case demonstrate that the FCRA might sweep broader than the consumer interests in financial confidentiality can justify.

In *Phillips v. Grendahl*,²⁶⁴ a mother worried that her daughter's fiancé was not the man he said he was. The mother had doubts that the fiancé had practiced law in Washington, D.C., as he said he had, and she was also suspicious of his inconsistent stories about ex-wives and girlfriends. With the help of a private investigator, the mother obtained information confirming that the fiancé was indeed lying to her daughter. Among other things, she learned that he had fathered several children in different states and that child support delinquency claims had been filed against him. In the course of the investigation, the mother had received a "Finder's Report," which both she and the furnisher believed were outside the scope of the FCRA.²⁶⁵ The Finder's Report did not include any information about the fiancé's debts. However, because the report stated the existence of some of the fiancé's bank accounts, this information transformed the Finder's Report into a "consumer report" for purposes of the FCRA, and the mother was civilly liable under the Act.²⁶⁶

The mother did not attempt to assert a First Amendment defense to challenge the enforcement of the FCRA in her case, but her facts illustrate problems with the privacy law. Even if the mother's suspicions had not been validated by

<http://www.whitehouse.gov/the-press-office/2012/02/23/fact-sheet-plan-protect-privacy-internet-age-adopting-consumer-privacy-b>.

261. 15 U.S.C. §§ 1681b, 1681n (2013).

262. 267 F.3d 1138, 1142-43 (D.C. Cir. 2001).

263. *Id.* at 1141-42.

264. 312 F.3d 357 (8th Cir. 2002), *abrogated by* *Safeco Ins. Co. of Am. v. Burr*, 551 U.S. 47 (2007).

265. *Id.* at 360-61.

266. *See id.* at 365-67.

the information she sought, the privacy law interferes with her natural, even laudable, instincts to test a hunch against actual facts. By defining “credit reports” so broadly, the statute ensnared and punished a person for following her curiosity. Borrowing from Kurt Vonnegut: “I love her for that, because it was so human.”²⁶⁷

HIPAA has similar problems. The duty of doctor-patient confidentiality was originally developed within the common law tort system. The duty was removed when a doctor’s disclosure of a patient’s health information served the public interest, or the interests of other individuals (e.g., sexual partners). The common law system had the flexibility to develop the rules on a case-by-case basis, so that the public interest override could evolve without risk of undermining confidentiality.²⁶⁸ HIPAA regulations, in contrast, attempt to anticipate and account for every public policy override, and set an otherwise inflexible rule of nondisclosure.²⁶⁹ As a consequence, HIPAA’s privacy provisions have had perverse effects on access to critical research data, quality of care, and overall public health.²⁷⁰

Going forward, if I am correct about a First Amendment right to knowledge creation, courts will need to scrutinize whether a privacy law is actually tailored to specific, weighty interests in seclusion or confidentiality. A well-tailored regulation will create limitations on particular disclosures and misuses of information, rather than creating global bans on data collection and distribution. Legislatures will need to tailor use restrictions to avoid the pernicious *effects* of data flow, rather than to attempt to deplete the flows themselves.

B. *Concern 2: If Data Is Speech, Regulators Are Hamstrung*

This Article has shown that government cannot limit the collection or dissemination of data in order to achieve certain preferred ends without a compelling interest to do so. In an information economy, this is a real and significant

267. KURT VONNEGUT, *SLAUGHTERHOUSE-FIVE* 19 (Dell Publ’g Co. 1985) (1969) (“And Lot’s wife, of course, was told not to look back where all those people and their homes had been. But she *did* look back, and I love her for that, because it was so human. So she was turned into a pillar of salt.”).

268. *See, e.g.*, *Tarasoff v. Regents of the Univ. of Cal.*, 551 P.2d 334, 347 (Cal. 1976) (holding that the “public policy favoring protection of the confidential character of patient-psychotherapist communications must yield to the extent to which disclosure is essential to avert dangers to others”).

269. 45 C.F.R. § 64.512 (2013) (enumerating particular exceptions to the general disclosure prohibition rather than setting out a flexible standard).

270. *See* Michael S. Wolf & Charles L. Bennett, *Local Perspective of the Impact of the HIPAA Privacy Rule on Research*, 106 *CANCER* 474 (2006) (finding that HIPAA led to a tripling of clinical study recruitment costs). Barbara Evans has described the disastrous consequences that privacy rules can cause by complicating the aggregation and sharing of medical data for research purposes. Evans, *supra* note 176, at 431-39.

restraint on state power. However, the government can achieve its preferred ends through appropriate regulation of conduct. When a business uses information to implement a decision or course of action, the implementation is usually treated as conduct. Conduct regulations will often achieve the objectives more directly than the regulation of the information that preceded the conduct.

For example, the Fair Housing Act constitutionally prevents a landlord from implementing a decision to decline a prospective tenant on account of the tenant's race.²⁷¹ This federal law restricts the uses to which a landlord may be tempted to put a piece of information—the race of a tenant. Although the Fair Housing Act does not, and practically could not, prevent a landlord from taking notice of a prospective tenant's race during a face-to-face meeting, the landlord's documentation of the tenant's race, or his inquiry over the phone about the tenant's race, though forms of speech, will constitute strong evidence that the landlord has engaged in racially disparate rental policies.²⁷²

Thus, if increased access to personal information winds up having detrimental effects on society, Congress and other lawmakers can target *misuses* of information without conflicting with a person's right to have and understand the information.²⁷³ If the government is worried about a specific use of data, it has the power to limit that use.

Let's examine a couple concrete examples. Consumer advocates worry about price discrimination on the basis of the information collected by web-tracking mechanisms.²⁷⁴ Nothing in this Article prevents these groups from lobbying for a statute prohibiting price discrimination on the basis of web usage. Likewise, it bothers Lori Andrews a good deal that creditors base their lending decisions in part on the purchase histories of borrowers.²⁷⁵ Again, consumer advocates can lobby for banking regulations that disallow disparate lending on this basis.

When these privacy concerns are reframed as use restrictions, the concerns are revealed to be misguided. Business practices can be justly criticized when

271. Fair Housing Act of 1968 § 804, 42 U.S.C. § 3604 (2011).

272. Other examples of restrictions on the use of information include laws prohibiting insider trading and the Affordable Care Act's prohibition against health insurers setting prices or coverage based on preexisting medical conditions. Patient Protection and Affordable Care Act § 1101, 42 U.S.C. § 18001 (2011).

273. Broad restrictions on use that outlaw all uses of information, or all but a narrow set of uses, strive to contain and control knowledge and should be treated as intentional regulations of thought, especially since one use is presumably dissemination of the information. This is not to say such a law could not survive scrutiny; confidentiality laws are a form of blanket-use restriction since a doctor is expected to use the information he receives only for the limited purposes of patient treatment or public health.

274. See Declan McCullagh, *Should Amazon.com Be Able to Charge You More than Someone Else?*, CNET NEWS (May 4, 2007, 9:41 AM PDT), http://news.cnet.com/8301-10784_3-9715959-7.html.

275. Lori Andrews, Op-Ed., *Facebook Is Using You*, N.Y. TIMES (Feb. 4, 2012), <http://www.nytimes.com/2012/02/05/opinion/sunday/facebook-is-using-you.html>.

they exploit financially vulnerable populations, but the most popular business uses of personal data tend to benefit traditionally disadvantaged consumers. Web-tracking information, if it is ever used for pricing, is used to charge certain affluent web surfers (like Mac users) more than others.²⁷⁶ And credit decisions made on the basis of factors other than income will have the salutary effect of reducing interest rates for poor-but-creditworthy consumers (and of increasing the interest rates for higher-income-but-unreliable debtors).²⁷⁷ It may be that some information collected by web-tracking technologies is the sort of intimate data that should be shielded from the scrutiny and use of others through the right to seclusion.²⁷⁸ But the problems anticipated by privacy advocates have not been carefully thought through.

Similarly, the government may (as it has²⁷⁹) prohibit health insurers from using information about gun ownership when pricing health policy premiums. Such a law makes the deliberate choice to force non-gun owners to subsidize the health insurance of gun owners. Policymakers are free to craft these discrete restrictions on information use, and then take the political consequences of the assumptions and priorities that are reflected in the laws.

But data privacy regulations that attempt to hamper the pace of information accumulation favor the noisy signals of yesterday's data environment over the more precise signals of tomorrow out of an unfounded assumption that better information will produce *worse* results for society. If the state has reason to think this is true in some circumstances, it can regulate conduct within those circumstances. But the state should not maintain an information status quo for its own sake.

There is another underutilized arrow in the legislature's quiver, too. Even if lawmakers are constrained from banning the collection or dissemination of data, they can always employ their considerable power to compel disclosures. For example, a law requiring anybody who wishes to record another person to provide notice of the recording will not, under usual circumstances, interfere with a person's right to record.

276. See Dana Mattioli, *On Orbitz, Mac Users Steered to Pricier Hotels*, WALL ST. J. (Aug. 23, 2012, 6:07 PM EDT), <http://online.wsj.com/article/SB10001424052702304458604577488822667325882.html>.

277. For a discussion about how the collection of personal data has democratized the availability of consumer credit, see J. Howard Beales, III & Timothy J. Muris, *Choice or Consequences: Protecting Privacy in Commercial Information*, 75 U. CHI. L. REV. 109, 115-17 (2008).

278. I have suggested so myself, though not through the lens of constitutional scrutiny. Bambauer, *supra* note 24, at 244-53.

279. Patient Protection and Affordable Care Act § 10101(e), 42 U.S.C. § 300gg-17(c)(4) (2011).

C. *Concern 3: Scrutiny Is Subjective at Best, Political at Worst*

Another concern is that judges will not be able to keep their personal or political beliefs from improperly influencing the scrutiny analysis.

The *IMS Health* decision exacerbates this concern. After all, when the Vermont legislature passed the law restricting the transfer and use of prescription data, it did so under the belief that the targeted marketing practices of pharmaceutical companies put indirect pressure on patients, through their doctors, to pay for and use more expensive new drug treatments when cheaper options were available. Whether or not the legislature's theory was correct, it was supported to some extent by public health research. Justice Kennedy found that the state could not regulate speech on the basis that such speech is persuasive.²⁸⁰ But Justice Kennedy's analysis is facile, and underdeveloped. Even if the state does not have an interest in protecting *doctors* from speech that they may find persuasive, the state's interests in protecting patients from the indirect influence of detailing practices is, at least arguably, a different matter. If the state had reason to believe that pharmaceutical detailing practices succeed in manufacturing demand running against consumer interests, the state would presumably have an interest in correcting the distortions. Justice Kennedy's opinion in *IMS Health*, like other First Amendment opinions that came before it, provides no guidance for determining which types of state interests are sufficiently important, and how the significance can be proved.²⁸¹

But these problems are not peculiar to data. Every time courts apply constitutional scrutiny, they sit in judgment about which state interests are sufficiently compelling and which are not. Human bias and error are intrinsic to judicial scrutiny, but this is the least worst of our options.

CONCLUSION

When privacy scholars argue that data should be treated differently from traditional forms of communication (utterances, journals, movies, and the like), they often do so for entirely rational and admirable reasons. If the First Amendment is too strong, and obstructs regulations that target low-value and negative-value speech, then the First Amendment will pose massive inefficiencies in our self-governance. Skepticism about the Supreme Court's maximalist approach to First Amendment law is completely logical: the First Amendment is quite literally getting in the way of regulations that might improve social welfare.

280. *Sorrell v. IMS Health Inc.*, 131 S. Ct. 2653, 2670 (2011).

281. In *IMS Health*, this may have been because the Vermont law was not designed to guard the state's stated interest in avoiding unwanted pressure, since doctors had the power to opt their patients into the detailers' databases. Also, Vermont put much less emphasis on this purported state interest than it did on others, related to patient and doctor privacy. *See id.* 2668-70.

The authors of the Bill of Rights could not have foreseen an avalanche of data when they drafted the First Amendment. But however they would have felt about Big Data, the restraints that they created were expected to frustrate the government, even when speech regulations are well intentioned. The First Amendment is, in many ways, an experiment that hinders the government from deciding what speech, and what thoughts, are good, even if most levelheaded people could agree on the matter. After all, a benevolent dictator is still a dictator.

But there is another, less pessimistic, explanation of the First Amendment's restrictions. Our predictions about the negative effects of speech and information are frequently proven to be wildly off the mark. Every new innovation provokes a flurry of fear and draft legislation, leaving behind the remnants of technopanic.²⁸² To take just one example, Caller ID, a service that is now taken for granted, was once the center of such heated privacy debate that the FCC still requires telephone companies to block any phone number if the caller requests it.²⁸³ The regulation was promulgated over the objections of privacy advocates, who urged the FCC to set the default as blocking the reporting of telephone numbers, and requiring callers to opt in if they were willing to let their phone numbers be displayed.²⁸⁴

This FCC rule looks much less critical in hindsight than it did at the time.²⁸⁵ This is why the First Amendment should, and does, create a strong presumption in favor of access to information. Very often the most sensible-seeming restrictions on information turn out to be flawed by status quo bias and fear. These biases and fears are overrepresented among public intellectuals today. Jeffrey Rosen has predicted that the Internet will be an existential threat to

282. The consequences are worse when legislation *is* passed. Donald Lively's account shows that "[h]istory actually suggests that the undoing of misconceived or dated regulation presents a much more vexing challenge." Donald E. Lively, *The Information Superhighway: A First Amendment Roadmap*, 35 B.C. L. REV. 1067, 1098 (1994); *see also* Donald E. Lively, *Fear and the Media: A First Amendment Horror Show*, 69 MINN. L. REV. 1071, 1076 (1985).

283. Delivery Requirements and Privacy Restrictions, 47 C.F.R. § 64.1601(b) (2013).

284. Calling Number Identification Service—Caller ID, 60 Fed. Reg. 29,489, 29,489 (June 5, 1995) (discussing privacy objections); *see* Marc Rotenberg, Director, Computer Prof'ls for Social Responsibility, Telephone Privacy in the 1990's Before the United States Telephone Association (Sept. 13, 1989).

285. Today the Caller ID debate looks especially odd since, we would say, the caller is imposing a call into the recipient's home, and thus should expect significantly less privacy. Even if the caller, too, is sitting in his own home, we see the placement of the phone call as an obvious and intentional transmission *out* of the home. But this goes to show that "the home" is a social construct that changes with increased technological education. The next generation may look at ours and wonder why a person sitting in their home but surfing the Internet would consider anything they do online to be "private."

our identities and individuality,²⁸⁶ and Viktor Mayor-Schönberger has argued that increased information will *decrease* our ability to learn, forget, and forgive.²⁸⁷ These theories have so little support from the history of the written word and the printing press, both information technological shocks of their times, that they are best forgiven and forgotten.²⁸⁸

Justice Holmes's dissents, which over time have become seminal to modern First Amendment law, show a desire to craft speech rights that will not bend to accommodate a bad idea that looks deceptively good. According to David Rabban's historical account, Justice Holmes may have been motivated to push for robust speech rights because of his own, personal realization that some of his strongly held beliefs might be wrong.²⁸⁹

In 1919, over the course of four opinions, *Schenck v. United States*,²⁹⁰ *Sugarman v. United States*,²⁹¹ *Frohwerk v. United States*,²⁹² and *Debs v. United States*,²⁹³ the Supreme Court upheld the convictions of several socialists for antiwar speech under the Espionage Act of 1917. Ironically, Justices Holmes and Brandeis were responsible for the majority opinions in all four of these cases—decisions that Justice Holmes described as cruel, but correct.²⁹⁴ But after these prosecutions, as the dreary negotiations of the Treaty of Versailles were playing out, many Americans began to question whether World War I had achieved the goals that had justified American intervention in Europe.²⁹⁵ Justice Holmes may have been among them.²⁹⁶ During this time of reflection, the justifications for quashing antiwar speech began to look hollow, and the

286. JEFFREY ROSEN, *THE UNWANTED GAZE: THE DESTRUCTION OF PRIVACY IN AMERICA* 8-9 (2000).

287. MAYER-SCHÖNBERGER, *supra* note 208, at 196-97.

288. Nate Silver recounts the history of the printing press in his new book. SILVER, *supra* note 171. The printing press caused such a rapid increase in the dissemination of books that Europeans of the fifteenth and sixteenth centuries had to toil through the information overload of their day, selecting among a wide library that was dominated by heretical religious texts and pseudoscience. *Id.* at 2-3. At the time, broad distribution of books would have seemed counterproductive for the pursuit of knowledge, but nobody would doubt the connection between the printing press and increased human knowledge today.

289. RABBAN, *supra* note 187, at 249-55.

290. 249 U.S. 47 (1919).

291. 249 U.S. 182 (1919).

292. 249 U.S. 204 (1919).

293. 249 U.S. 211 (1919).

294. G. Edward White, *Justice Holmes and the Modernization of Free Speech Jurisprudence: The Human Dimension*, 80 CALIF. L. REV. 391, 420-21 (1992).

295. RABBAN, *supra* note 187, at 342. For more information on the political fracturing caused by the Treaty of Versailles, see LLOYD E. AMBROSIUS, *WOODROW WILSON AND THE AMERICAN DIPLOMATIC TRADITION: THE TREATY FIGHT IN PERSPECTIVE* (1987).

296. See RABBAN, *supra* note 187, at 351; JAY STANLEY, AM. CONSTITUTION SOC'Y, *ISSUE BRIEF: THE CRISIS IN FOURTH AMENDMENT JURISPRUDENCE* 11-12 (2010), available at <http://www.acslaw.org/files/ACS%20Issue%20Brief%20-%20Stanley%204th%20Amendment.pdf>.

government's increased use of the Espionage Act to prosecute socialists looked equally wrong-headed.

Before the year 1919 came to a close, Justice Holmes had written one of his many famous dissents in the case of *Abrams v. United States*.²⁹⁷ The other Justices continued to write opinions, as they had when Justice Holmes was among them, that accused the antiwar radicals of manipulating, even abusing, the First Amendment by invoking it "to justify the activities of anarchy or of the enemies of the United States."²⁹⁸ Justice Holmes, on the other hand, used his dissent to walk back some of the exceptions to free speech that he had himself created. Though he never admitted to having a change of heart about his earlier opinions, the stance he took in *Abrams* shows unequivocally that he had.²⁹⁹

Justice Holmes may not have come to the defense of the persecuted Bolshevik pamphleteers if he had continued to believe that their message was wrong. That is, Justice Holmes's metamorphosis was not entirely driven by sympathy for political minorities. Equally important was the fact that his previously held assumptions about the Great War conflicted with the raw evidence he came to perceive as it wrapped up. So when Justice Holmes wrote "[p]ersecution for the expression of opinions seems to me perfectly logical,"³⁰⁰ he was not conceding that an expansive First Amendment is illogical. Rather, the persecution of disfavored expressions only *seems* logical. Time and the flushing out of conflicting evidence may wind up proving otherwise.

Information about the war led Justice Holmes to have a special, powerful experience: the changing of the mind. The sanctity of a freely made mind requires protection not only for speech, but also for the digestion of raw facts.

297. 250 U.S. 616, 624-31 (Holmes, J., dissenting).

298. *Schaefer v. United States*, 251 U.S. 466, 477 (1920); *see also* *Gilbert v. Minnesota*, 254 U.S. 325, 333 (1920).

299. RABBAN, *supra* note 187, at 346. The explanation Justice Holmes gave to preserve consistency was that the pamphlets advocated noninterference with the Bolshevik Revolution, and since the United States was not at war with Russia, the message could not cause a clear danger. *See Abrams*, 250 U.S. at 628-29 (Holmes, J., dissenting).

300. *Abrams*, 250 U.S. at 629 (Holmes, J., dissenting).