THE ASPEN INSTITUTE
CONGRESSIONAL PROGRAM

INTERNET, BIG DATA, AND ALGORITHMS:
THREATS TO PRIVACY AND FREEDOM OR GATEWAY TO A NEW FUTURE

May 10 – 13, 2019 | Cambridge, Massachusetts

INTERNET, BIG DATA & ALGORITHMS: GATEWAY TO A NEW FUTURE OR A THREAT TO PRIVACY AND FREEDOM

The Aspen Institute Congressional Program
May 10-13, 2019
Cambridge, Massachusetts

TABLE OF CONTENTS

RAPPORTEUR'S SUMMARY

Grace Abuhamad

Graduate student, Technology and Policy Program, MIT

Under the auspices of the Aspen Institute Congressional Program, a bipartisan group of twelve members of Congress convened from May 10—13, 2019, at the Massachusetts Institute of Technology to discuss implications and policy options regarding the Internet, big data, and algorithms. The members of Congress deliberated with scholars and practitioners to acquire a better understanding of artificial intelligence technologies, their current and future applications, and possible threats to consumer privacy and freedom.

The participants were mindful that artificial intelligence is a new source of wealth, but also a new source of inequality among nations and within nations. Today's "arms race" is one where countries such as China have directed national strategies and aim to claim technological supremacy within a decade. Given the scope and scale of artificial intelligence, the nation that will shape the future of these technologies will shape the future of the world. Whether or not the United States may be the only nation able to leverage its resources and win such a race remains to be seen.

Defining Success in Artificial Intelligence

Artificial intelligence is the ability for machines to learn without being explicitly programmed. Like humans, these machines learn from past data to predict future outcomes. When the input data is limited, machines produce biased and harmful results that tend to have disparate impact on disempowered groups.

Algorithm designers can mitigate these results by recognizing limitations and changing their definition of success. Currently, success is measured by an algorithm's overall or aggregate performance at a defined task, such as matching faces to names. Research indicates that algorithms can have high aggregate accuracy, and yet, when results are disaggregated by racial or ethnic groups, can show significant disparities among these groups. Applications of such algorithms can automate inequality and discrimination that existed in past data on which these algorithms are trained.

In most cases, designers are not aware of the data limitations and their unintended consequences in artificial intelligence applications. This challenge is not unique to artificial intelligence. For example, there used to be more female than male fatalities in automobile accidents since automobiles were designed and tested according to male-form crash test dummies. Once this limitation was recognized and corrected, fatalities equalized across genders. The definition of successful design

and testing expanded to include gender equality. As awareness around algorithmic bias increases, there may also be an expansion of the definition of success for these algorithms.

Awareness, context, and transparency are three ways by which to expand the definition of success. Given that artificial intelligence has the potential to impact every sector of the economy and aspect of American lives, there needs to be more widespread training to increase awareness of both the benefits and risks of artificial intelligence. Once aware, Americans can democratize artificial intelligence by bringing diverse experiences to recognize and address limitations.

Context plays an important role in artificial intelligence, since some applications have more limitations than others. Participants recognized, for example, that export controls needed to be more precise: instead of limiting artificial intelligence as a whole, limits could be applied specifically to kinetic applications. Contexts that are highly-regulated today, such as healthcare and national security, will have higher thresholds for safety and accuracy of artificial intelligence applications. Where determining precise regulations or thresholds may not yet be possible, increasing transparency is another way to expand discourse around success in artificial intelligence applications.

Transparency can help align artificial intelligence with public trust. Artificial intelligence presents a number of opportunities from autonomy, speed, and endurance that exceed human capacities and could serve as a power system for national security. Even as these applications may deliver lethal capacity in a more targeted way, there is a need for legal checks, and perhaps a "human-in-the-loop" process in order for these systems to be trustworthy.

Explanations are a form of transparency that develop this human-machine collaboration. These too, are context specific, and each context may require different gradients of explanations, raising questions such as: for whom is the explanation for? What is its purpose? Can the explanation be contested? Is the explanation feasible technically and financially? In the medical context, for example, a machine learning system developed to reduce the instances of sepsis was not explainable, but brought down the instance of sepsis by 60%. Participants agreed that this example and others make the debate about explanation requirements more nuanced.

Threats to Privacy and Democratic Freedoms

Algorithmic harms are perhaps less noticeable, though no less threatening to civil liberties. In the context of facial recognition technology, an estimated 130 million people in the United States already have images of their faces in government databases and can be subject to unwarranted searches. While these images may have been lawfully collected through driver's license registries and other government services, it is not clear that any searches respect the context in which the image was collected..

Private companies engage in more pervasive data collection since individuals voluntarily upload and identify personal photographs, without full awareness as to how these data will be used. During the Black Lives Matter protests, law enforcement officials identified some individuals using data sourced from both government databases and social media

platforms. Such uses of data could have chilling effects on civil liberties. There are no federal laws regulating the use of facial recognition technology and individual face prints.

Like the facial recognition example above, certain uses of data are "lawful but awful," in the sense that they do not promote democratic values. At scale, these uses can undermine democracy and election integrity through surveillance, misinformation, disinformation, and manipulation. About 70% of American adults use social media today, yet only about 5% did in 2005. Social media networks have evolved to play a role in trust and civic engagement, though perhaps not always as a positive force. A recent study indicated that, following mass shootings, notoriety can be propagated by and within social media networks: after a request not to name the perpetrator of the attacks in Christchurch, New Zealand, only 15% of the American press did, yet those articles represented 45% of the articles shared on social media.

Social media platforms were perhaps not designed to protect democratic values, though this does not necessarily prevent a change of purpose. Participants discussed Newt Minow's leadership as Chairman of the Federal Communications Commission in the late 1960s to evolve television programming beyond what he viewed as a "vast wasteland," launching the Public Broadcasting Service and creating a role for television in informing the public. These actions did not eliminate television, but instead created an additional form of television aligned with public purpose. Similar to the "4th Estate" role of the press, today's influence of social media networks suggests perhaps a public role in protecting democratic values, in addition to, or along with, a role in protecting individual privacy.

Negotiating the Boundaries of Privacy and Control

The legal and social boundaries of privacy have changed over time, and are based on different assumptions in different cultures and societies. In our modern world, data is key. But who actually owns the data and when or how one consents to having their data collected are disputable topics. For example, once an individual's data has been harvested and processed, through either voluntarily or involuntarily online interactions, it can be put to use in targeted consumer marketing campaigns, campaign advertisements, and individualized sales pitches.

In debating how and whether to protect privacy, policymakers face a revealed preferences contrast: individuals claim privacy is important to them, yet behave as though it is not. One may therefore suggest that protecting privacy is not necessary. However, there are other ways to understand consumer behavior that lead to the opposite conclusion. Consumers have less information than collectors about how their data can be used and, at the same time, they are overloaded with too much information that they cannot easily process. Researchers suggest that the opportunity cost of reading privacy terms of service amounts to over $700 billion in a consumer's lifetime. As a result, consumers tend to select default options, which, depending on how they are set, influence the outcomes. For example, with organ donation, in countries that presume donation, there are higher donations, and vice versa. Overall, this suggests that consumers exert less of a choice than is assumed with regard to protecting their information and privacy.

Most privacy policies today are developed on a flawed "notice and consent" model. For one, privacy policies change often, yet they do not require new consent for every change. They assume consent. Second, this model depends on the construct of information as property that consumers can control, yet it is increasingly difficult for consumers to exert such control. Often, information that is used is not the same as the purpose for which it was collected. Another issue with control is that consumers cannot always control others' sharing information about them, such as through a social network. Privacy is viewed as an individual choice, but it is increasingly communal. It may be that privacy needs to be defined not by information control, but rather a negotiation of boundaries between public and private spaces.

Privacy protection has put too much burden on consumers to manage their information. There are social norms about privacy in public that have been broken by social networks. Consumers may feel coerced into handing over information in order to use a service. Participants agreed that privacy should not be a privilege: consumers need options and cannot be cut out of processes simply because they do not agree to terms, especially as services are increasingly online. There was interest in privacy-enhancing tools and technologies, such as encryption, differential privacy, and secure multi-party computing, though concern about whether consumers would bear the cost of deployment. Participants also considered how to operationalize the principle of data minimization. Another suggestion was a do-not-track registry for consumers to opt-out of data sharing, though there was concern that offering opt-out options would create a selection bias among consumers.

One of the challenges is assessing the value of personal data and the impact of stronger protections. In discussing whether there should be some form of insurance policy against personal data breaches, participants agree that it is difficult to quantify the consumer costs of violations. While individuals' comfort with these techniques varies, one thing is certain: marketing will never be the same. The explosive power of artificial intelligence is being harnessed for commercial advantage, which can be either advantageous or disadvantageous to the consumer depending on what perspective is held.

Platform Regulation as Information Fiduciaries

The major digital companies spent over $60 million in 2018 in lobbying. Consolidation in the digital industry has raised questions about the power of dominant major players. Participants discussed whether the economies of scale serve consumer interest, or are to the detriment of consumer choices and costs. The digital market is dominated by advertising technology, with consumer data as the value driver. Large platforms dominate the space since their products are designed to collect user information as soon as the user is signed-in and identified. Third parties are struggling to compete by collecting data through other online activity.

To hold companies accountable for their power, regulators need to identify harms to consumers. Data breach harms have galvanized some action, and this action might be best focused on increasing consumer confidence. While there are many criticisms of the General Data Protection Regulation in Europe, the law gave Europeans a sense of confidence and control over their privacy. At the same time,

some view this law as a "digital trade war" and a way for Europe to extract rents from American companies since they have yet to develop their own innovative environment. Of the 20 most valuable companies in the world, 11 are in the United States, 9 are in China, and none are in Europe. Despite very different approaches, both China and the United States have designed markets that allow for companies to prosper and create value. Participants agreed that regulation in the United States needs to better balance consumer protection with innovation, keeping in mind smaller companies and startups that could bear a heavier compliance burden than large incumbents.

One suggestion for platform regulation emerges in the common law tradition of fiduciaries. For example, medical and legal professionals need sensitive information about their patients and clients in order to perform a service, yet these professionals also have a responsibility not to disclose this information. Platforms could be considered "information fiduciaries." This means that companies would be responsible for acting in the best interest of consumers when it comes to their data. These companies would have a duty of care, a duty of confidentiality, and a duty of loyalty. Such duties would foster an environment of trust and increase consumer confidence.

These duties would still allow companies to continue advertising in ways that provide information to consumers, but would hold them responsible when actions are discriminatory, manipulative, or predatory. Some companies have already started to take on a fiduciary role with regard to "news," by defining the term and choosing to boost rankings for more reliable sources of information. In developing the terms of their moderation role, some may argue that companies lose the intermediary liability protections awarded to them in Section 230 of the Communications Decency Act (also known as Title V of the 1996 Telecommunications Act). Participants discussed the intent of Section 230, suggesting that its goal was to protect companies when they acted as editors, assuming that companies would engage rather than not. Participants also discussed possible First Amendment challenges, citing, as an example, the 2018 repeal of the Department of Labor's financial fiduciary rule.

Takeaways

Though not specified explicitly in the Constitution, privacy has emerged as an individual right. Consumers have lost control over their personal information and the ability to think about privacy as consumers rather than products. Companies have taken some action, but need further direction and clarity as to their roles and responsibilities, especially as the European Union's General Data Privacy Regulation, and now the state of California, have imposed greater privacy protections for online behavior than previously required. From an innovation perspective, countries such as China are investing heavily in research and development of technologies to shape our future. The time to act is now. The public interest is at stake, perhaps even beyond the borders of the United States.

Participants suggested shifting the burden of responsibility off consumers, focusing remedies to measurable harms, being aware of the balance between incumbents and entrants, and encouraging experimentation in a "sandbox" type regulatory environment. While the Federal Trade Commission has the ability to police bad behavior, participants did not settle the question of whether extending rulemaking authority would be beneficial. Some of the

proposals require more research and impact assessments prior to decisions, which policymakers could facilitate by mandating more transparency and access to data. Overall, participants agreed that threats to privacy and democratic values should be viewed through a lens of continuous and evolving risk mitigation, as opposed to total eradication.

OPENING REMARKS BY MIT PRESIDENT

L. Rafael Reif

President, Massachusetts Institute of Technology

As you may have heard, MIT recently launched the MIT Stephen A. Schwarzman College of Computing. It's a $1.1 billion dollar initiative that will add 50 new faculty—and it represents the largest restructuring of MIT in 70 years.

As a university president, let me tell you: Orchestrating that amount of change is really difficult! You definitely would not try it without some very good reasons!

So, what inspired us to start the new College?

Artificial Intelligence is an enabling technology. You may even have heard it called, "the new electricity." That means it has the power to decisively reshape how all of us live and work together. It will help humanity learn more, waste less, work smarter, live longer – and better understand, predict and make decisions about almost anything that can be measured.

As a result, in the not-too-distant future, AI will be a dominant source of new wealth—for those nations equipped to make the right commitments now. There are not many of those countries! So we should not be surprised if this new source of wealth also becomes a new source of inequality, both between nations, and within them.At the same time, however, the promises and benefits of AI and related technologies clearly come with risks, the kind you have all come here to learn about: threats to privacy, public safety, jobs, the security of nations – and more.

All of this is to say: the opportunities are immense, very few people are prepared to seize them—and society is simply not armed with practical strategies for making sure these technologies are responsibly deployed for the common good.

Now let me frame this in terms of what we're doing with the MIT Schwarzman College of Computing.

Given the power and pervasiveness of AI and related technologies, and given their potent mix of opportunity and threats, I believe that those of us in higher education share a pressing responsibility:

We need to reshape our institutions, so we can equip our students to shape the future.

At MIT, our students have been quietly leading this revolution for some time. Computer science has long been our most popular major, and in the last few years, the interest has been explosive. Roughly 40% of our students now major in computer science alone or paired with another subject. On their own, our students are making sure they graduated ready to bring computing-intensive approaches to

fields from molecular biology to economics to urban planning.

In effect, our students are telling us that educational institutions like MIT must deliberately equip tomorrow's leaders to be "bilingual" in computing and whatever else interests them. Graduates entering every field will need to be fluent in AI strategies to advance their own work. And technologists will also need to be fluent in the cultural values and ethical principles that should ground and govern the use of these tools, for the good of all.

We aim to equip our students to be leaders in making sure that AI can flourish in, and support, a society that values individual rights and freedoms. And we want them to be actively engaged with how to help American workers compete and succeed, as AI transforms the very nature of work.

In short, we have come to believe that it's time to educate a new generation of technologists in the public interest.

Fortunately, MIT is not alone. Other institutions across the country are responding to this new reality too, in various ways—and that is good news for the nation.

But no matter how well we teach our students, and no matter how "bilingual" they become, these efforts must be matched by a national effort to sustain our nation's technology leadership. So let me close on that note.

Because you signed up for this remarkable program, I do not need to tell anyone here that our nation is globally engaged in a technological race to the horizon. Other nations, such as China, have been advancing aggressively to assert technological supremacy in critical fields of science and technology. And they are doing this by pursuing a systematic, long-term, highly funded national strategy.

I believe that America needs to respond urgently and deliberately to the scale and intensity of this challenge. We may not do so, however. In which case, we should expect that, in fields from personal communications to business, health and security, China is likely to become the world's most advanced technological nation, and the source of the most cutting-edge technological products in not much more than a decade.

Fortunately, this scenario is not inevitable.

The United States has tremendous assets, including the immense global strength of our technology sector today. This is the result, in part, of a unique formula that no other country has been able to copy: the large number of first-rate American universities, pursuing advanced research—with long-term federal support. This relationship is rooted in a national culture of opportunity and entrepreneurship. It is inspired by an atmosphere of intellectual freedom. It is supported by the rule of law. And, most importantly, it enables new creative heights by uniting brilliant talent from every sector of our society and every corner of the world.

For decades, these factors have helped make our nation the most powerful scientific and technological engine on Earth. Every American can take pride in this distinctive system.

If we want to secure our nation's future, and its technological pre-eminence in this technology race, we need a highly visible, focused, sustained federal effort to fund research and development in key

science and technology areas. And we need incentives to make sure universities, industry and government are working together to capitalize on them.

In the coming decades, it will feel as though the opportunities, disruptions and progress of the Industrial Revolution are playing out at time-lapse speed. Responding to the magnitude of this challenge will require a strategic effort across society. Whichever nation acts now to shape the future of AI will shape the future for us all. I hope and believe it will be ours.

I am impressed and grateful that all of you made this journey in search of greater understanding. As the nation confronts the challenges of the technological future, I hope you will continue to see MIT as a resource. And I wish you many wonderful discussions over the next two days!

ARTIFICIAL INTELLIGENCE & PUBLIC POLICY: THE BEGINNING OF A CONVERSATION

R. David Edelman

Director, Project on Technology, Economy and National Security, MIT

Whether in the headlines or our news feeds, Board Rooms or the Situation Room, discussion about artificial intelligence (AI) seems to have reached a fever pitch — and with good reason. Like cybersecurity a decade ago, a combination of optimism, intrigue, concern, and technical complexity are catapulting this once-obscure set of technological tools to the center of dialogues far beyond computer science.

The United States is blessed with exceptional expertise in computing, and continues to be at the forefront of many technological advances in AI algorithms, applications, and hardware to run them. What is perhaps newer is a growing sense of discussion, urgency, and for many profound purpose associated with ensuring that the benefits of these advances of computing are widely distributed and the substantial (and increasingly visible) harms they might exacerbate. The challenges are myriad but varied across applications of AI; the benefits are potentially substantial but not automatic; and the government's role in any of these areas is deeply unsettled. This discussion will hopefully provide a foundation for both what AI and its constituent technology of machine learning (ML) truly is — and, crucially, what it is not yet capable of — and a sense of how it is and will continue to be applied. The goal is singular: to inform and help guide sound public policy amidst a context of continual technological change.

As we have seen so spectacularly, perhaps even tragically over the last couple of years, gone are the days when technologists have their world, when the governments had another, completely separate world and ne'er the two shall meet. The conversations we will have over these next few days, between those from the technology world and those in the policy world, are proof that the nation's top policymakers and engineers are beginning to understand two key insights: that technology policy is now increasingly just policy, and data ethics are increasingly just ethics. No part of the business of government is unaffected by the changes ushered in by the digital world, and ethical decisions are increasingly going to be data-driven.

Our task is to work to align the benefits of Artificial Intelligence with the obligations of public trust.

Now what does that effort require? First, it takes technical rigor, insights from the people who are in the labs, in academia and in industry, pioneering these innovations in machine learning robotics. Second, it takes policy awareness, the

perspective of those who know how to turn hard problems that have inevitable tradeoffs (there are always tradeoffs) into public policy that communities and countries can get behind. And third, it takes cross pollination, which is recognition that neither side, technical, policy or anyone in between, has a monopoly on good ideas, nor all the answers — and, frankly, a recognition that none of us in a single community have the luxury of deference to the other completely anymore.

The days in which engineers can simply ship a tool and assume that someone else will handle the consequences are gone. When it comes to AI in the area of public trust, the era of moving fast and breaking everything is coming to a close. There is simply too much at stake for us not to, collectively, have a say.

The topics that we will cover in our discussion are intentionally broad. We will discuss how AI is in use across a range of industries, in a range of applications, creating a range of cross-functional issues. We are taking that approach to help illustrate both the interconnections among them, and crucially, the differences. The use of AI to serve better ads on social media might be technologically similar to AI in aviation or autonomous vehicles, but arguably only one has the real if not immediate potential to save hundreds, thousands, or even tens of thousands of lives every year — if we can build and sustain enough trust in these systems to get inside of them ourselves.

In that context of transportation, for instance, we might rightly ask: what is the proper threshold of safety? Why is it that we have, for instance, almost zero tolerance for safety risk as airline passengers but getting in my car for the commute home is probably the most dangerous thing I will do all day, if not all year? How specifically should engineers be building and validating systems to make them worthy of public confidence, the sort of confidence we have when we get in our cars or airplanes today? Who gets to decide what the proper threshold for that trust is? And how do we do all of that while, as you will hear, we are still perfecting these machines? To use the crudest metaphor: how can we build the plane while we're flying it?

These implications differ from those at the intersection of AI and manufacturing, including with its dramatic implications for the future of work. Perhaps more-so than many are aware, AI is already changing manufacturing, and along with it, potentially bringing profound shifts in the U.S. labor market. But this has also been an area rife with speculation and anxiety, with wild prognostications about the "end of work" striking fear in the hearts of both sides of the political spectrum. If there is one benefit to this uncertainty, though, it has been to motivate leading research — including here at MIT, from scholars like David Autor — to develop real data about what the future of work might look like, and what is working to manage these transitions where they present themselves.

AI is also changing the face of healthcare, an area with immense potential to improve lives and to save them, but is also one of the most heavily regulated industries in the United States and around the world. And so here, too, there are no easy answers. The argument, for instance, that AI systems should be able to explain themselves — as some European jurisdictions are now demanding of the tools wherever they manifest themselves — sounds like a very appealing concept on its face. But if patients are given the choice between a perfect cancer diagnosis system and another that is very imperfect, but

better at explaining itself, it is not hard to imagine which most will prefer.

In reality the American medical system accommodates innovation that works in ways that elude us — permitting drugs and devices that are proven effective and safe, even if their exact operation remains partially mysterious. By analogy, we might know that patching a tire prevents it from deflating, even if we lack a profound understanding of the physics of rubber. Might we want to build our regulations of these AI-enabled medical systems to accommodate that sort of ambiguity? If we do, it would make "regulation of AI transparency" writ large much more difficult — it would instead by context-specific. Might it be, then, that the concept of "AI regulation" is no more meaningful than "regulation of C++" the programming language, or plastic, or steel? If AI is today simply a series of machine learning techniques, does it make sense to treat them as a distinct technology at all? The answer to this question is core to understanding how policymakers, and Congress in particular, might approach these concerns.

Time-permitting, we will also discuss the use of AI for national security and defense. The notion of autonomous killer robots is certainly evocative. But more discussion can help us separate the fact from the fiction here, and unpack some proposed ideas of how to deal with these issues. For instance, the U.S. government alone has proposed the notion of export controls or arms control regimes for machine learning and computer vision — in other words, core AI technologies. But there are deep questions there. Can you have controls of any kind on systems that are, at their core in many cases, open source? How does that work? How can you even control it? Is the genie out the bottle? And what role, importantly as you've seen in the front pages, should engineers play, if any, in keeping their innovations from being used in the conduct of war?

The one thing that we do know and that is that we cannot wait. As policymakers, we cannot wait for the engineers to design the perfectly fair, perfectly accountable, perfectly transparent system. We have to help them understand how. And we have to help them understand what they need to do to meet the burden of public trust. As engineers, we also cannot wait for policymakers to have a single, clear, uniform, perfect instruction manual for how to design systems that are better, fairer, more accountable and consistent with law. In other words, we have to do both of those together.

This discussion will be for some a continuation, for many a start, but for none the final word on issues that will no doubt occupy us all for the duration of our careers, whether in public service, research, or in private industry. Our hope is that you will leave with a sense of information, engagement, and empowerment; to know where AI is and where it's headed, to understand the opportunities and be on the watch for the warning signs of it going awry.

Elected officials, policymakers, engineers, lawyers, teachers, advocates — and all of us, as citizens — have a role to play in shaping a future suffused with AI. Policy prescriptions may differ, but foundational understanding of the technology should not. That foundation, which we hope to build here together, is the only way that we can ensure that this technology is used for us — and not against us. We look forward to the conversation.

ALGORITHMS ARE REPLACING NEARLY ALL BUREAUCRATIC PROCESSES

Cathy O'Neil

CEO, O'Neil Risk Consulting and Algorithmic Auditing

Wherever there was once a complex, sticky human question, we now have a slick black box called artificial intelligence (AI) and that is presented as a foolproof scientific instrument. We should not trust these AI tools blindly.

For example, algorithms have taken over what was once the work of corporate Human Resources departments. Who gets interviewed, who gets hired or fired, who gets a bonus—these decisions were once made by humans, fairly or not. Now they are increasingly being made by machines, or at least supplemented by data.

Algorithms have been picked up by the justice system. Where to send the police, how long to sentence, whether to incarcerate pre-trial, or whether to grant parole are decisions that used to place humans at the center. Now we have scoring systems that do that for us.

In financial matters, algorithms are wholeheartedly embraced by companies competing with each other for the best customers. Who gets a credit card, and what their offered interest rate is, and for that matter which credit card ads they see when they go to the company's website, all determined by AI scoring systems that evaluate a person by their social media presence, browsing history, location, and the make and model of the device they're using. Similarly, life insurance, car insurance, and to some extent health insurance companies are using all kinds of "big data" and AI techniques to decide on someone's risk profile and premiums.

There's more. College admissions algorithms, child abuse risk scores, automated drone strikes, facial recognition in cities and stadiums, and suicide risk scores on social media to name a few. They're proliferating as the data about us makes those algorithms cheap to build and profitable or efficient to use.

That's a lot of power. The question is, how well made are these tools? Can we trust them to work?

To give two examples where the answer is no, think of the Volkswagen emissions scandal – an algorithm in automobiles trained to game emissions tests – and the recent Boeing disaster, which looks to be an algorithm that takes control of the airplane under certain conditions, which erred tragically in the presence of a malfunctioning sensor.

If you ask the owner of one of these algorithms whether it works, they'll undoubtably say "yes." And if you ask them, "what do you mean by that?", they'll suggest that it's either more efficient than

the old system or that it's more accurate than humans, who are notoriously biased. In other words, they'll explain how they work for them, the owners.

They will also probably suggest that the algorithm is too scientifically sophisticated to really explain or understand, and that we should trust them that they've got things under control.

What they probably will not have measured, however, is for whom the algorithms fail, and what is meant by failure for those folks.

Are the job application algorithms filtering out perfectly good candidates? Are they filtering out more qualified candidates from protected classes? They probably haven't tested that question.

In other words, the makers and the owners of these algorithms have overly narrow views of success, and insufficiently thoughtful approaches to what could go wrong. They often avoid thinking about worst case scenarios, first because it's against their commercial interest to do so, and second because they honestly haven't been forced to. So far people have trusted them.

Algorithms are often unaccountable. This is a threat to our constitutional rights.

In the context of livelihood or the justice system, people are not typically allowed to appeal their scores, and often don't even know what their scores are.

In a recent court case in Houston, six teachers who had been fired based solely on the basis of a "teacher accountability" scoring algorithm, called the teacher value-added model, sued and won. The judge ruled that, although their scores were low, their due process rights had been violated because nobody could explain to them what the scores actually meant.

That's a single example of how scoring systems can be flagged for accountability failures. We're seeing proposed legislation for cities and states to conduct "audits" of algorithms used by government. More generally, we can expect class action lawsuits to determine what exactly our constitutional rights are for accountability as consumers, workers, and citizens.

Powerful algorithms are sometimes invisible to the target. Especially in the context of online advertising, people often don't even know they're being scored.

And while that's fine when we're being evaluated online for our propensity to buy a sofa versus a loveseat, it becomes less clear when we're being evaluated for whether we'd like to be shown an ad for a STEM job in Silicon Valley or a daycare provider job in Sacramento.

We've recently seen a spate of lawsuits against Facebook for these very problems, both in employment and housing advertising.

But what's even messier about this particular problem is that, in general and outside of Facebook, online advertising ecosystems often do not collect information on protected class status like gender, race, age, or disability status. So it's unclear how to clean up the situation even if the intention was to do so, without entirely revamping the advertising ecosystem.

Regulators are not trained to solve these problems.

We are no longer in the age where a regulator could find a "smoking gun" email of a boss telling a hiring manager not to hire people based on their race. Instead, we

have opaque, black box algorithms that might have a disparate impact, but it could well be unintentional. How do we even discover that problem?

It's not impossible. Reuters recently reported that Amazon built a hiring algorithm for engineers, but decided not to use it after discovering it was sexist. That's good news, both because Amazon bothered to test that question and because they acted on the result. It also means that, if Amazon can do that, then so can regulators, at least once they've learned how.

Problems of algorithmic accountability are not going away.

They're just proliferating, because algorithms are indeed quite good at making things more profitable, more efficient, and less accountable for their owners.

The risks are too high to ignore the potential for algorithms to do harm. The government should figure out ways to prioritize public good, fairness, and accountability for these tools, whether that means training regulators to enforce current laws, passing new laws that enlarge the concept of algorithmic accountability, or forming a new regulator that acts like an "FDA" for algorithms.

How to Exercise the Power You Didn't Ask For*

Jonathan Zittrain

Professor of International Law, Harvard Law School

I used to be largely indifferent to claims about the use of private data for targeted advertising, even as I worried about privacy more generally. How much of an intrusion was it, really, for a merchant to hit me with a banner ad for dog food instead of cat food, since it had reason to believe I owned a dog? And any users who were sensitive about their personal information could just click on a menu and simply opt out of that kind of tracking.

But times have changed.

The digital surveillance economy has ballooned in size and sophistication, while keeping most of its day-to-day tracking apparatus out of view. Public reaction has ranged from muted to deeply concerned, with a good portion of those in the concerned camp feeling so overwhelmed by the pervasiveness of their privacy loss that they're more or less reconciled to it. It's long past time not only to worry but to act.

Advertising dog food to dog owners remains innocuous, but pushing payday loans to people identified as being emotionally and financially vulnerable is not. Neither is targeted advertising that is used to exclude people. Julia Angwin, Ariana Tobin, and Madeleine Varner found that on Facebook targeting could be used to show housing ads only to white consumers. Narrow targeting can also render long-standing mechanisms for detecting market failure and abuse ineffective: State attorneys general or consumer advocates can't respond to a deceitful ad campaign, for instance, when they don't see it themselves. Uber took this predicament to cartoon villain extremes when, to avoid sting operations by local regulators, it used data collected from the Uber app to figure out who the officials were and then sent fake information about cars in service to their phones.

These are relatively new problems. Originally, our use of information platforms, whether search engines or social media, wasn't tailored much to anything about us, except through our own direct choices. Your search results for the query "Are vaccinations safe?" would be the same as mine or, for a term like "pizza," varied in a straightforward way, such as by location, offering up nearby restaurants. If you didn't like what you got, the absence of tailoring suggested that the search platform wasn't to blame; you simply were seeing a window on the web at large. For a long time that was a credible, even desirable, position for

* This article first appeared in the *Harvard Business Review*, September 19, 2018

content aggregators to take. And for the most part they themselves weren't always good at predicting what their own platforms would offer up. It was a roulette wheel, removed from any human agent's shaping.

Today that's not true. The digital world has gone from pull to push: Instead of actively searching for specific things, people read whatever content is in the feeds they see on sites like Facebook and Twitter. And more and more, people get not a range of search results but a single answer from a virtual concierge like Amazon's Alexa. And it may not be long before such concierges rouse themselves to suggest it's time to buy a gift for a friend's birthday (perhaps from a sponsor) or persistently recommend Uber over Lyft when asked to procure a ride (again, thanks to sponsorship).

Is it still fair for search platforms to say, "Don't blame me, blame the web!" if a concierge provides the wrong directions to a location or the wrong drug interaction precautions? While we tend not to hold Google and Bing responsible for the accuracy of every link they return on a search, the case may be different when platforms actively pluck out only one answer to a question — or answer a question that wasn't even asked.

We've also moved to a world where online news feeds — and in some cases concierges' answers to questions — are aggressively manipulated by third parties trying to gain exposure for their messages. There's great concern about what happens when those messages are propaganda — that is, false and offered in bad faith, often obscuring their origins. Elections can be swayed, and people physically hurt, by lies. Should the platforms be in the business of deciding what's true or not, the way that newspapers are? Or does that open the doors to content control by a handful of corporate parties — after all, Facebook has access to far more eyeballs than a single newspaper has ever had — or by the governments that regulate them?

Companies can no longer sit this out, much as they'd like to. As platforms provide highly curated and often single responses to consumers' queries, they're likely to face heated questions — and perhaps regulatory scrutiny — about whom they're favoring or disfavoring. They can't just shrug and point to a "neutral" algorithm when asked why their results are the way they are. That abdication of responsibility has led to abuse by sophisticated and well-funded propagandists, who often build Astroturf campaigns that are meant to look as if they're grassroots.

So what should mediating platforms do?

An answer lies in recognizing that today's issues with surveillance and targeting stem from habit and misplaced trust. People share information about themselves without realizing it and are unaware of how it gets used, passed on, and sold. But the remedy of allowing them to opt out of data collection leads to decision fatigue for users, who can articulate few specific preferences about data practices and simply wish not to be taken advantage of.

Restaurants must meet minimum standards for cleanliness, or (ideally) they'll be shut down. We don't ask the public to research food safety before grabbing a bite and then to "opt out" of the dubious dining establishments. No one would rue being deprived of the choice to eat food contaminated with salmonella. Similar intervention is needed in the digital universe.

Of course, best practices for the use of personal information online aren't nearly as clear cut as those for restaurant cleanliness. After all, much of the personalization that results from online surveillance is truly valued by customers. That's why we should turn to a different kind of relationship for inspiration: one in which the person gathering and using information is a skilled hired professional helping the person whose data is in play. That is the context of interactions between doctors and patients, lawyers and clients, and certified financial planners and investors.

Yale Law School's Jack Balkin has invoked these examples and proposed that today's online platforms become "information fiduciaries." We are among a number of academics who have been working with policymakers and internet companies to map out what sorts of duties a responsible platform could embrace. We've found that our proposal has bipartisan appeal in Congress, because it protects consumers and corrects a clear market failure without the need for heavy-handed government intervention.

"Fiduciary" has a legalese ring to it, but it's a long-standing, commonsense notion. The key characteristic of fiduciaries is loyalty: They must act in their charges' best interests, and when conflicts arise, must put their charges' interests above their own. That makes them trustworthy. Like doctors, lawyers, and financial advisers, social media platforms and their concierges are given sensitive information by their users, and those users expect a fair shake — whether they're trying to find out what's going on in the world or how to get somewhere or do something.

A fiduciary duty wouldn't broadly rule out targeted advertising — dog owners would still get dog food ads — but it would preclude predatory advertising, like promotions for payday loans. It would also prevent data from being used for purposes unrelated to the expectations of the people who shared it, as happened with the "personality quiz" survey results that were later used to psychometrically profile voters and then attempt to sway their political opinions.

This approach would eliminate the need to judge good from bad content, because it would let platforms make decisions based on what their users want, rather than on what society wants for them. Most users want the truth and should be offered it; others may not value accuracy and may prefer colorful and highly opinionated content instead — and when they do, they should get it, perhaps labeled as such. Aggregators like Google News and Facebook are already starting to make such determinations about what to include as "news" and what counts as "everything else." It may well be that an already-skeptical public only digs in further when these giants offer their judgments, but well-grounded tools could also inform journalists and help prevent propaganda posted on Facebook from spreading into news outlets.

More generally, the fiduciary approach would bring some coherence to the piecemeal privacy protections that have emerged over the years. The right to know what data has been collected about you, the right to ask that it be corrected or purged, and the right to withhold certain data entirely all jibe with the idea that a powerful company has an obligation to behave in an open, fair way toward consumers and put their interests above its own.

While restaurant cleanliness can be managed with readily learned best practices

(keep the raw chicken on a separate plate), doctors and lawyers face more complicated questions about what their duty to their patients and clients entails (should a patient with a contagious and dangerous disease be allowed to walk out of the office without treatment or follow-up?). But the quandaries of online platforms are even less easy to address. Indeed, one of the few touchstones of data privacy — the concept of "personally identifiable information," or PII — has become completely blurry, as identifying information can now be gleaned from previously innocuous sources, making nearly every piece of data drawn from someone sensitive.

Nevertheless, many online practices will always be black-and-white breaches of an information fiduciary's duty. If Waze told me that the "best route" somewhere just so happened to pass by a particular Burger King, and it gave that answer to get a commission if I ate there, then Waze would be putting its own interests ahead of mine. So would Mark Zuckerberg if hypothetically he tried to orchestrate Facebook feeds so that Election Day alerts went only to people who would reliably vote for his preferred candidate. It would be helpful to take such possibilities entirely off the table now, at the point when no one is earning money from them or prepared to go to bat for them. As for the practices that fall into a grayer area, the information fiduciary approach can be tailored to account for newness and uncertainty as the internet ecosystem continues to evolve.

Ideally, companies would become fiduciaries by choice, instead of by legal mandate. Balkin and I have proposed how this might come about — with, say, U.S. federal law offering relief from the existing requirements of individual states if companies opt in to fiduciary status. That way, fiduciary duties wouldn't be imposed on companies that don't want them; they could take their chances, as they already do, with state-level regulation.

In addition, firms would need to structure themselves so that new practices that raise ethical issues are surfaced, discussed internally, and disclosed externally. This is not as easy as establishing a standard compliance framework, because in a compliance framework the assumption is that what's right and wrong is known, and managers need only to ensure that employees stay within the lines. Instead the idea should be to encourage employees working on new projects to flag when something could be "lawful but awful" and congratulate — rather than retaliate against — them for calling attention to it. This is a principle of what in medical and some other fields is known as a "just culture," and it's supported by the management concept of "psychological safety," wherein a group is set up in a way that allows people to feel comfortable expressing reservations about what they're doing. Further, information fiduciary law as it develops could provide some immunity not just to individuals but to firms that in good faith alert the public or regulators to iffy practices.

Instead of having investigations into problems by attorneys general or plaintiffs' lawyers, we should seek to create incentives for bringing problems to light and addressing them industrywide. That suggests a third touchstone for an initial implementation of information fiduciary law: Any public body chartered with offering judgments on new issues should be able to make them prospectively rather than retroactively. For example, the IRS can give taxpayers a "private letter ruling" before they commit to one tax strategy or another.

On truly novel issues, companies ought to be able to ask public authorities — whether the Federal Trade Commission or a new body chartered specifically to deal with information privacy — for guidance rather than having to make a call in unclear circumstances and then potentially face damages if it turns out to be the wrong one.

Any approach that prioritizes duty to customers over profit risks trimming margins. That's why we need to encourage a level playing field, where all major competitors have to show a baseline of respect. But the status quo is simply not acceptable. Though cleaning up their data practices will increase the expenses of the companies who abuse consumers' privacy, that's no reason to allow it to continue, any more than we should heed polluters who complain that their margins will suffer if they're forced to stop dumping contaminants in rivers.

The problems arising from a surveillance-heavy digital ecosystem are getting more difficult and more ingrained. It's time to try a comprehensive solution that's sensitive to complexities, geared toward addressing them as they unfold, and based on duty to the individual consumers whose data might otherwise be used against them.

BEYOND THE VAST WASTELAND*

Ethan Zuckerman

Director, Center for Civic Media, MIT Media Lab

In 1961, the newly appointed chairman of the Federal Communications Commission, Newt Minow, addressed the National Association of Broadcasters in Washington, D.C. Minow's speech demanded that broadcasters take seriously the idea to serve the public interest – and distinguished the public interest from simply what interests the public. And Minow coined an unforgettable phrase to explain what a poor job broadcasters were doing. Challenging executives to watch a day of their own programming without anything to distract or divert them, Minow declared, "I can assure you that what you will observe is a vast wasteland."[1]

There have been hundreds of articles written over the past two years about social media that might have been better titled "a vast wasteland". This flood of articles argues that social media often doesn't work the way we think it should, that partisan manipulation of Facebook may be swaying elections, and that extremism on YouTube may be contributing to a wave of ethnonationalist violence. It's a thoroughly appropriate moment to evaluate whether social media is making our society and our democracy stronger, or pulling it apart. From Cambridge Analytica to Comet Ping Pong to the massacre in New Zealand, alarm bells are sounding that not all is well in our online public spaces.

But Minow's speech didn't end with a condemnation of the sorry state of broadcasting in 1961. Instead, Minow articulated a vision for television to inform, enlighten and entertain, a future he hoped to achieve without censorship, without replacing private companies with government entities, and mostly through voluntary compliance. And, with 1967's Public Broadcasting Act, the founding of the Public Broadcasting Service in 1969 and National Public Radio in 1970, a surprising amount of Minow's vision came to pass.

It's important that we consider the real and potential harms linked to the rise of social media, from increasing political polarization, the spread of mis-, dis- and malinformation[2] to trolling, bullying and online abuse. But much as television was in

* Adapted from "Six or Seven Things Social Media Can Do for Democracy", May 30, 2018, https://www.media.mit.edu/articles/six-or-seven-things-social-media-can-do-for-democracy/

[1] Newt Minow, "Television and the Public Interest", address delivered 9 May 1961, National Association of Broadcasters, Washington, DC.

[2] Clare Wardle and Hossein Derakshan, "Information disorder: definitions", in "Understanding and Addressing the Disinformation Ecosystem", conference publication. https://firstdraftnews.org/wp-content/uploads/2018/03/The-Disinformation-Ecosystem-20180207-v4.pdf

its teenage years in the early 1960s, social media isn't going away any time soon. It's essential that we have a positive vision for what social media can be as well as a critical take on mitigating its harms.

I'm interested in what social media should do for us as citizens in a democracy. We talk about social media as a digital public sphere, invoking Habermas and coffeehouses frequented by the bourgeoisie. Before we ask whether the internet succeeds as a public sphere, we ought to ask whether that's actually what we want it to be.

I take my lead here from journalism scholar Michael Schudson, who took issue with a hyperbolic statement made by media critic James Carey: "journalism as a practice is unthinkable except in the context of democracy; in fact, journalism is usefully understood as another name for democracy." For Schudson, this was a step too far. Journalism may be necessary for democracy to function well, but journalism by itself is not democracy and cannot produce democracy. Instead, we should work to understand the "Six or Seven Things News Can Do for Democracy", the title of an incisive essay Schudson wrote to anchor his book, *Why Democracies Need an Unlovable Press*[3].

The six things Schudson sees news currently doing for democracy are presented in order of their frequency – as a result, the first three functions Schudson sees are straightforward and unsurprising. The news informs us about events, locally and globally, that we need to know about as citizens. The news investigates issues that are not immediately obvious, doing the hard work of excavating truths that someone did not want told. News provides analysis, knitting reported facts into complex possible narratives of significance and direction.

Schudson wades into deeper waters with the next three functions. News can serve as a public forum, allowing citizens to raise their voices through letters to the editor, op-eds and (when they're still permitted) through comments. The news can serve as a tool for social empathy, helping us feel the importance of social issues through careful storytelling, appealing to our hearts as well as our heads. Controversially, Schudson argues, news can be a force for mobilization, urging readers to take action, voting, marching, protesting, boycotting, or using any of the other tools we have access to as citizens.

His essay closes with a seventh role that Schudson believes the news should fill, even if it has yet to embrace it. The news can be a force for the promotion of representative democracy. For Schudson, this includes the idea of protecting minority rights against the excesses of populism, and he sees a possible role for journalists in ensuring that these key protections remain in force.

This is perhaps not an exhaustive list, nor is the news required to do all that Schudson believes it can do. Neither does the list include things that the news tries to do that aren't necessarily connected to democracy, like providing an advertising platform for local businesses, providing revenue for publishers, or entertaining audiences. And Schudson acknowledges that these functions can come into conflict – the more a news organization engages in mobilization, the more likely it is that it will compromise its ability to inform impartially.

[3] Schudson, Michael. Why Democracies Need an Unlovable Press. Polity, 2008.

In this same spirit, I'd like to suggest six or seven things social media can do for democracy. As with Schudson's list, these functions are not exhaustive – obviously, social media entertains us, connects us with family, friends and any advertiser willing to pay for the privilege, in addition to the civic functions I outline here. Furthermore, as with news media, these civic purposes are not always mutually reinforcing and can easily come into conflict. (And because I'm much less learned than Schudson, my list may be incomplete or just plain wrong.)

Social media can inform us.

Many of us have heard the statistic that a majority of young people see Facebook as a primary source for news[4], and virtually every newsroom now considers Facebook as an important distributor of their content (sometimes to their peril.) But that's not what's most important in considering social media as a tool for democracy. Because social media is participatory, it is a tool people use to create and share information with friends and family, and potentially the wider world. Usually this information is of interest only to a few people – it's what you had for lunch, or the antics of the squirrel in your backyard. But sometimes the news you see is of intense importance to the rest of the world.

When protesters took to the streets of Sidi Bouzid, Tunisia, they were visible to the world through Facebook even though the Tunisian government had prevented journalists from coming to the town. Videos from Facebook made their way to Al Jazeera through Tunisian activists in the diaspora, and Al Jazeera rebroadcast footage, helping spread the protests to Tunis and beyond. The importance of social media in informing us is that it provides a channel for those excluded by the news – whether through censorship, as in Tunisia, or through disinterest or ignorance – to have their voices and issues heard.

Places don't need to be as far away as Tunisia for social media to be a conduit for information – when Michael Brown was killed in Ferguson, Missouri, many people learned of his death, the protests that unfolded in the wake, and the militarized response to those protests, via Twitter. (And as news reporters were arrested for covering events in Ferguson, they turned to Twitter to share news of their own detention.) Social media is critically important in giving voice to communities who've been systemically excluded from media – people of color, women, LGBTQIA people, poor people. By giving people a chance to share their under-covered perspectives with broadcast media, social media has a possible role in making the media ecosystem more inclusive and fair.

Finally, social media may be helping replace or augment local information, as people connect directly with their children's schools or with community organizations. This function is increasingly important as local newspapers shed staff or close altogether, as social media may become the primary conduit for local information.

Social media can amplify important voices and issues.

In traditional (broadcast or newspaper) media, editors decide what topics are worth the readers' attention. This

[4] Amy Mitchell, Katerina Eva Masta and Jeffrey Gottfried, "Facebook Top Source for Political News Among Millennials", June 1, 2015. https://www.journalism.org/2015/06/01/facebook-top-source-for-political-news-among-millennials/

"agenda setting" function has enormous political importance – as Max McCombs and Donald Shaw observed in 1972[5] , the news doesn't tell us what to think, but it's very good at telling us what to think about.

That agenda-setting power takes a different shape in the era of social media. Instead of a linear process from an editor's desk through a reporter to the paper on your front porch, social media works with news media through a set of feedback loops[6]. Readers make stories more visible by sharing them on social media (and help ensure invisibility by failing to share stories). Editors and writers respond to sharing as a signal of popularity and interest, and will often write more stories to capitalize on this interest. Readers may respond to stories by becoming authors, injecting their stories into the mix and competing with professional stories for attention and amplification.

Amplification has become a new form of exercising political power. In 2012, we watched Invisible Children use a carefully crafted campaign, built around a manipulative video and a strategy of sharing the video with online influencers. Within an few days, roughly half of American young people had seen the video, and U.S. funding for the Ugandan military – the goal of the campaign – was being supported by powerful people in the U.S. Congress and military[7]. (That the organization's director had a nervous breakdown, leading to the group's implosion, was not a coincidence – Invisible Children managed to amplify an issue to a level of visibility where powerful backlash was inevitable.)

Amplification works within much smaller circles than those surrounding U.S. foreign policy. By sharing content with small personal networks on social media, individuals signal the issues they see as most important and engage in a constant process of self-definition. In the process, they advocate for friends to pay attention to these issues as well. Essentially, social media provides an efficient mechanism for the two-step flow of communication, documented by Paul Lazarsfeld and Elihu Katz[8], to unfold online. We are less influenced by mass media than we are by opinion leaders, who share their opinions about mass media. Social media invites all of us to become opinion leaders, at least for our circles of friends, and makes the process entertaining, gamifying our role as influencers by rewarding us with up to the second numbers on how our tweets and posts have been liked and shared by our friends.

Social media can be a tool for connection and solidarity.

The pre-web internet of the 1980s and 1990s was organized around topics of interest, rather than offline friendships, as social networks like Facebook organize. Some of the most long-lasting communities that emerged from the Usenet era of the internet were communities of interest that connected people who had a hard time

[5] McCombs, Maxwell E., and Donald L. Shaw. "The agenda-setting function of mass media." Public opinion quarterly 36.2 (1972): 176-187.

[6] Ethan Zuckerman, "Four Problems for Media and Democracy". April 1, 2018. https://medium.com/trust-media-and-democracy/we-know-the-news-is-in-crisis-5d1c4fbf7691

[7] Josh Kron and J. David Goodman, "Online, a Distant Conflict Soars to Topic No. 1", New York Times, March 8, 2012.

[8] Katz, Elihu, Paul F. Lazarsfeld, and Elmo Roper. Personal influence: The part played by people in the flow of mass communications. Routledge, 2017.

finding each other offline: young people questioning their sexuality, religious and ethnic minorities, people with esoteric or specialized interests. The spirit of the community of interest and identity continued through Scott Hefferman's meetup.com, which helped poodle owners or Bernie Sanders supporters in Des Moines find each other, and now surfaces again in Facebook Groups, semi-private spaces designed to allow people to connect with likeminded individuals in safe, restricted spaces.

Social critics, notably Robert Putnam[9], have worried that the internet is undermining our sense of community and lessening people's abilities to engage in civic behavior. Another possibility is that we're forming new bonds of solidarity based on shared interests than on shared geographies. I think of Jen Brea, whose academic career at Harvard was cut short by myalgic encephalomyelitis[10], who used the internet to build an online community of fellow disease sufferers, a powerful documentary film that premiered at Sundance, and a powerful campaign calling attention to the ways diseases that disproportionately affect women are systemically misdiagnosed. Brea's disease makes it difficult for her to connect with her local, physical community, but social media has made it possible to build a powerful community of interest that is working on helping people live with their disease.

One of the major worries voiced about social media is the ways in which it can increase political polarization. Communities of solidarity can both exacerbate and combat that problem. We may end up more firmly rooted in our existing opinions, or we may create a new set of weak ties to people who we may disagree with in terms of traditional political categories, but with whom we share powerful bonds around shared interests, identities and struggles.

Social media can be a space for mobilization

The power of social media to raise money for candidates, recruit people to participate in marches and rallies, to organize boycotts of products or the overthrow of governments is one of the best-documented – and most debated – powers of social media. From Clay Shirky's examination of group formation and mobilization in Here Comes Everybody to endless analyses of the power of Facebook and Twitter in mobilizing youth in Tahrir Square or Gezi Park, including Zeynep Tufekçi's Twitter and Tear Gas, the power of social media to both recruit people to social movements and to organize actions offline has been well documented. It's also been heartily critiqued, from Malcolm Gladwell, who believes that online connections can never be as powerful as real-world strong ties for leading people to protest, or by thinkers like Tufekçi, who readily admit that the ease of mobilizing people online is an Achilles heel, teaching leaders like Erdogan to discount the importance of citizens protesting in the streets.

[9] Putnam, Robert D. "Bowling Alone: America's Declining Social Capital." Culture and politics. Palgrave Macmillan, New York, 2000. 223-234.

[10] Kaitlyn Tiffany, "Jen Brea documented her Chronic Fatigue Syndrome on an iPhone so that doctors would believe other women", September 21, 2017. https://www.theverge.com/2017/9/21/16163950/unrest-documentary-sundance-creative-distribution-fellowship-interview

It's worth noting that mobilization online does not have to lead to offline action to be effective. A wave of campaigns like Sleeping Giants, which has urged advertisers to pull support from Breitbart, or #metoo, where tens of thousands of women have demonstrated that sexual harassment is a pervasive condition, not just the product of a few Harvey Weinsteins, have connected primarily online action to real-world change. What's increasingly clear is that online mobilization – like amplification – is simply a tool in the contemporary civic toolkit, alongside more traditional forms of organizing.

Social media can be a space for deliberation and debate.

Perhaps no promise of social media has been more disappointing than hope that social media would provide us with an inclusive public forum. Newspapers began experimenting with participatory media through open comments fora, and quickly discovered that online discourse was often mean, petty, superficial and worth ignoring. Moving debate from often anonymous comment sections onto real-name social networks like Facebook had less of a mediating effect that many hoped. While conversations less often devolve into insults and shouting, everyone who's shared political news online has had the experience of a friend or family member ending an online friendship over controversial content. It's likely that the increasing popularity of closed online spaces, like Facebook groups, has to do with the unwillingness of people to engage in civil deliberation and debate, and the hope that people can find affirmation and support for their views rather than experiencing conflict and tension.

Yet it is possible to create spaces for deliberation and debate within social media. Wael Ghonim was the organizer of the We Are All Khaled Said Facebook page, one of the major groups that mobilized "Tahrir youth" to stand up to the Mubarak regime in Egypt, leading to the most dramatic changes to come out of the Arab Spring. After the revolution, Ghonim was deeply involved with democratic organizing in Egypt. He became frustrated with Facebook, which was an excellent platform for rallying people and harnessing anger, but far less effective in enabling nuanced debate about political futures. Ghonim went on to build his own social network, Parlio, which focused on civility and respectful debate, featuring dialogs with intellectuals and political leaders rather than updates on what participants were eating for lunch or watching on TV. The network had difficulty scaling, but was acquired by Quora, the question-answering social network, which was attracted to Parlio's work in building high-value conversations that went beyond questions and answers[11].

Parlio suggests that the dynamics of social networks as we understand them have to do with the choices made by their founders and governing team. Facebook and Twitter can be such unpleasant places because strong emotions lead to high engagement, and engagement sells ads. Engineer a different social network around different principles, and it's possible that the deliberation and debate we might hope for from a digital public sphere could happen within a platform.

[11] Josh Constantine, "Quora's first acquisition is Arab Spring instigator's Q&A site Parlio", https://techcrunch.com/2016/03/30/quora-parlio/

Social media can be a tool for showing us a diversity of views and perspectives.

The hope that social media could serve as a tool for introducing us to people we don't already know – and particularly to people we don't agree with – may seem impossibly cyberutopian. Indeed, I wrote a book, Rewire, that argues that social media tends to reinforce homophily, the tendency of birds of a feather to flock together. Given the apparent track record of social media as a space where ethnonationalism and racism thrive, skepticism that social media can introduce us to new perspectives seems eminently reasonable.

Contemporary social networks have an enormous amount of potential diversity, but very little manifest diversity. In theory, you can connect with 2 billion people from virtually every country in the world on Facebook. In practice, you connect with a few hundred people you know offline, who tend to share your national origin, race, religion and politics. But a social network that focused explicitly on broadening your perspectives would have a tremendous foundation to build upon: networks like Facebook know a great deal about who you already pay attention to, and have a deep well of alternative content to draw from.

Projects like FlipFeed from MIT's Laboratory for Social Machines and gobo.social from my group at the MIT Media Lab explicitly re-engineer your social media feeds to encourage encounters with a more diverse set of perspectives. If a network like Twitter or Facebook concluded that increased diversity was a worthy metric to manage to, there's dozens of ways to accomplish the goal, and rich questions to be solved in combining increased diversity with a user's interests to accomplish serendipity, rather than increased randomness.

Social media can be a model for democratically governed spaces.

Users in social networks like Twitter and Facebook have little control over how those networks are governed, despite the great value they collectively create for platform owners. This disparity has led Rebecca MacKinnon to call for platform owners to seek Consent of the Networked, and Trebor Scholz to call us to recognize participation in social networks as Digital Labor. But some platforms have done more than others to engage their communities in governance.

Reddit is the fourth most popular site on the U.S. internet and sixth most popular site worldwide, as measured by Alexa Internet, and is a daily destination for at least 250 million users. The site is organized into thousands of "subreddits", each managed by a team of uncompensated, volunteer moderators, who determine what content is allowable in each community. The result is a wildly diverse set of conversations, ranging from insightful conversations about science and politics in some communities, to ugly, racist, misogynistic, hateful speech in others. The difference in outcomes in those communities comes in large part to differences in governance and to the partipants each community attracts.

Some Reddit communities have begun working with scholars to examine scientifically how they could govern their communities more effectively. /r/science, a community of 18 million subscribers and over a thousand volunteer moderators, has worked with communications scholar Nathan Matias to experiment with ways of enforcing their rules to maximize positive

discussions and throw out fewer rulebreakers[12]. The ability to experiment with different rules in different parts of a site and to study what rulesets best enable what kinds of conversations could have benefits for supporters of participatory democracy offline as well as online.

Beyond the vast wasteland

It's fair to point out that the social media platforms we use today don't fulfill all these functions. Few have taken steps to increase the diversity of opinions users are exposed to, and though many have tried to encourage civil discourse, very few have succeeded. It's likely that some of these goals are incompatible with current ad supported business models. Political polarization and name-calling may well generate more pageviews than diversity and civil deliberation.

Some of these proposed functions are likely incompatible. Communities that favor solidarity and subgroup identity, or turn that identity into mobilization, aren't the best ones to support efforts for diversity or for dialog.

Finally, it's also fair to note that there's a dark side to every democratic function I've listed. The tools that allow marginalized people to report their news and influence media are the same ones that allow fake news to be injected into the media ecosystem. Amplification is a technique used by everyone from Black Lives Matter to neo-Nazis, as is mobilization, and the spaces for solidarity that allow Jen Brea to manage her disease allow "incels" to push each other towards violence. While I feel comfortable advocating for respectful dialog and diverse points of view, someone will see my advocacy as an attempt to push politically correct multiculturalism down their throat, or to silence the exclusive truth of their perspectives through dialoge. The bad news is that making social media work better for democracy likely means making it work better for the Nazis as well. The good news is that there's a lot more participatory democrats than there are Nazis.

My aim in putting forward seven things social media could do for democracy is two-fold. As we demand that Facebook, Twitter and others do better – and we should – we need to know what we're asking for. I want Facebook to be more respectful of my personal information, more dedicated to helping me connect with my friends than marketing me to advertisers, but I also want them to be thinking about which of these democratic goals they hope to achieve.

The most profound changes Newt Minow inspired in television happened outside of commercial broadcasting, in the new space of public broadcasting. I believe we face a similar public media moment for social media. Achieving the democratic aims for social media outlined here requires a vision of social media that is plural in purpose, public in spirit and participatory in governance. Rather than one social network that fills all our needs, we need thousands of different social networks that serve different communities, meeting their needs for conversation with different rules, norms and purposes. We need tools that break the silos of contemporary social media, allowing a citizen to follow conversations in dozens of different spaces with a single tool. Some of these spaces will be ad or subscription supported, while some might be run by local governments with taxpayer funds, but some subset of social media needs to consciously serve the public interest as its

[12] See civilservant.io

primary goal. Finally, farming the management of online spaces to invisible workers half a world away from the conversations they're moderating isn't a viable model for maintaining public discussions. Many of these new spaces will be experiments in participatory governance, where participants will be responsible for determining and enforcing the local rules of the road.

We accept the importance of a free and vibrant press to the health of our democracy. It's time to consider the importance of the spaces where we deliberate and debate that news, where we form coalitions and alliances, launch plans and provide support to each other. The free press had defenders like Thomas Jefferson, who declared that if he had to choose between "a government without newspapers or newspapers without a government, I should not hesitate a moment to prefer the latter". The health of our digital public spheres is arguably as important, and worth our creative engagement as we imagine and build spaces that help us become better citizens. Social media as a vast wasteland is not inevitable, and it should not be acceptable. Envisioning a better way in which we interact with each other online is one of the signature problems of modern democracy and one that demands the attention of anyone concerned with democracy's health in the 21st century.

PRIVACY AND CONSUMER CONTROL

J. Howard Beales III and Timothy J. Muris

Professor of Strategic Management and Public Policy,
George Washington University

George Mason University Foundation Professor of Law at Antonin Scalia Law School, Senior Counsel at Sidley Austin LLP

We address the role of consumer control in protecting privacy. Our focus is commercial interactions, such as secondary uses of information originally collected for another purpose or the widespread practice of tracking consumers across websites. Social media, which attempt to mirror a non-commercial setting, raise different issues. In particular, when information is shared with friends on (or off of) social media, the friend also knows the information and can share it with others or use if for a different purpose.

We start by discussing policy approaches that in fact do little to protect most consumers, based in part on defining privacy in commercial transactions as the property of one side of the transaction. We then turn to a more practical policy approach based on the adverse consequences of information use.

Is Personal Information Property?

Many discussions of consumer control begin by asserting that information about a consumer is the property of the consumer. Because commercial information is in fact the joint product of an interaction between an individual consumer and another entity, property is of limited utility. Consider, for example, the online payment service Venmo, which lets consumers make and share payments with friends. By default, transactions on Venmo are public, but either the sender or the recipient of funds can change those settings. Which of us should control the information that you and I engaged in a transaction? The only sensible answer is that we both have access to that information, and can use it or reveal it as we see fit. Or consider genetic information. It is undoubtedly an individual's genetic profile, but it also belongs to children, parents, and siblings. Yet it makes no sense to say that we need the permission of all our relatives to obtain our own genetic profiles and use them as we desire.

Consistent with the notion that information is property, the traditional (and European) approach to privacy has long been based on the so-called Fair Information Practices ("FIPs"). The quintessential feature of FIPs is the seemingly attractive idea of notice and choice—tell consumers about information practices and let them choose whether to allow that use of information or not.

Unfortunately, however, FIPs is fatally flawed for protecting consumer privacy.

Consider first the problems of notice. Privacy policies are everywhere, but they are seldom read and even less likely to be seriously considered in deciding whether to interact with a website. The reason is obvious: One study estimated that the opportunity cost of actually reading online privacy notices would be $781 billion.[1] And, of course, the cost of reading the myriad of other privacy policies that surround us, from HIPPA to Gramm Leach Bliley notices and many others in between, is not even a part of this substantial cost estimate.

The costs of simply reading online notices greatly exceed what is at stake. The entire online advertising market in 2017 was $88 billion[2], just over a tenth of the cost of reading the notices. Moreover, many consumers see the mere existence of a privacy policy as meaning that their privacy is protected, when the policy itself may offer no protection at all.[3] Simpler privacy notices could help, but even if we could reduce the cost of reading privacy policies by half—a herculean undertaking—the costs would be far disproportionate to the stakes.

The principle of allowing consumer choice fares no better. Critical information systems function only because consumers lack choice about including their information. Credit reporting, for example, is critical for lenders to assess risks and avoid loans to people who will likely not repay. If consumers could choose whether their information is reported, however, consumers who are bad risks would likely opt out of having their payment history reported. The system would be much less able to distinguish good credit risks from bad, because many high risk consumers would simply not be reported. Responsible consumers with thin credit histories would be less able to get credit. As another example, consider the property recordation system, which records property ownership and any liens against the property. A creditor has a perfectly legitimate interest in knowing whether a consumer willing to pledge a house as collateral has already made that same promise to other lenders, but if consumers could opt out of having their information reported, they would be far less able to do so.

One could, of course, argue that these are exceptions to a general rule. But a general rule that applies only in certain unspecified circumstances is not a general rule at all, and thus it is hardly a guide for sound regulatory policy.

A privacy protection regime that relies on consumers deciphering elaborate privacy policies to determine with which service providers they are willing to interact is not consumer protection at all. It places a burden on consumers that is entirely unreasonable. Instead, it is, as the Europeans admit, about data protection. Data protection, however, is not an end in itself. Using data in ways that harm consumers may protect the data if consumers have consented, but it is hardly

[1] McDonald, Aleecia M., and Lorrie Faith Cranor. The Cost of reading privacy policies, ISJLP 4 (2008):43.

[2] https://www.iab.com/news/digital-ad-spend-reaches-all-time-high-88-billion-2017-mobile-upswing-unabated-accounting-57-revenue/

[3] Martin, K. (2015). Privacy Notices as Tabula Rasa: An Empirical Investigation into how Complying with a Privacy Notice is Related to Meeting Privacy Expectations Online. Journal of Public Policy & Marketing, 34(2), 210–227. https://doi.org/10.1509/jppm.14.139.

consumer protection. It is the privacy equivalent of giving consumers a long list of carcinogenic food additives they should carefully avoid, without a practical guide to avoidance.

Property rights are an attractive way to organize society when transactions costs are relatively low, because they can be reassigned. If the property is more valuable to someone other than its current owner, it can easily be sold, or the parties can bargain about how best to deploy the asset. When transactions costs are high, however, negotiations to rearrange rights can cost more than the potential gain. In a simple world, product liability cases would involve contract law and parties could bargain about the desired degree of safety precautions. As complexity of both manufacturing and distribution arrangements increase, however, product liability involves tort law: The manufacturer has a duty to take reasonable safety precautions, without the need (or practical ability) to negotiate the details.

Although transactions costs may not seem high between the consumer and the website collecting information, as noted above they are significant. And, unlike many other examples in the legal and economic literatures, benefits are small. For most consumers, there is very little at stake in considering how a website or another commercial entity might use information about a visit or a transaction. Because the probability of some adverse event occurring from secondary information use is remote, the question of how the information might be used is simply not worth much attention. Although a relatively small number of consumers are extremely concerned about protecting their privacy, most are not so concerned, and are therefore unwilling to incur the costs of even thinking about the issue.

In these circumstances, default rules about whether information can, or can not, be shared are therefore likely to determine the outcome. As Richard Posner observed about the clear importance of default rules for organ donations, "the probability that one's organs will be harvested for use in transplantation must be very slight—so slight that it doesn't pay to think much about whether one wants to participate in such a program. When the consequences of making a 'correct' decision are slight, ignorance is rational, and therefore one expects default rules to have their greatest effect on behavior …"[4] And that is what experimental studies of opting in versus opting out have consistently shown – the default rule controls for most consumers.

Of course, those who care more are more likely to be willing to think about the issue, whatever the default rule. Experimental evidence, although limited, indicates that those who care most about privacy make more consistent choices when the default rule changes.[5] That finding argues for an "opt out" rule, if we must choose between opt in and opt out, because the people who care about the issue are willing to take the time necessary to consider it. Those who are not concerned do not have to face the costs of thinking about it. That is an appropriate allocation of effort, because those who are

[4] Richard Posner, Organ Sales – Posner's Comment, The Becker-Posner Blog (Jan. 1, 2006), available at http://www.becker-posner-blog.com/2006/01/page/2/.

[5] Yee-Lin Lai & Kai-Lung Hui, Internet Opt-in and Opt-out: Investigating the Roles of Frames, Defaults and Privacy Concerns, Proceedings of the 2006 ACM SIGMIS CPR Conference on Computer Personnel Research, 253 (2006).

not concerned are happy to defer to the default rule. As discussed below, however, this default rule is crucial to the support of online advertising markets, and in turn to the principal funding mechanism for the internet content we all enjoy.

Are the Consequences of Information Use a Better Focus for Privacy Regulation?

Rather than property rights and default rules, a far superior way to develop privacy policy is to consider the consequences of information use and misuse. The reason we care about commercial information use or sharing is that something bad might happen to consumers, and the goal should be to avoid those adverse consequences. There is little reason for concern when information is used to benefit a consumer, such as when information is exchanged to facilitate a transaction, or when a vendor uses information that was originally collected for a completely different purpose to reduce the risk of fraud. There is reason for concern, however, when information is used in harmful ways. The harm, however, not the information, should provide the focal point for regulation.

The consequences of inappropriate information use may be physical, if use enables stalking, or locating children online. They may be economic, in the form of identity theft. They may be annoyances, in the form of unwanted telemarketing calls or irritating robocalls. And they may be the kinds of more subjective harms that have long been actionable as privacy torts: Intrusion upon seclusion, putting someone in a false light, or publicizing private information in a manner highly offensive to a reasonable person.

Since we implemented the harm-based approach to privacy regulation at the Federal Trade Commission in 2001, it has been extremely productive. It led directly to the National Do Not Call Registry, which worked well until it was overwhelmed by developments in robocall technology. (Like spam, robocalls will likely be solved by technology, and the regulators and phone companies are actively pursuing solutions.) It led as well to a series of cases to protect the security of information from thieves who would use it to do harm to consumers.

Focusing regulation on harm is particularly important because the internet has enabled substantial benefits from the information sharing economy. Fraud control tools that look for consistency in how an identity is used rely on information originally collected for far different purposes, including even magazine subscriptions, to help assess the risk that a particular transaction is fraudulent. Without such tools, identity theft would likely be an even more serious problem. Location data has enabled real-time navigation aids that can help avoid traffic problems and ease the daily commute. With the continued rapid growth of internet connected devices, more information will likely be available, and entrepreneurs will find new ways to use this information to enhance our lives.

Targeted advertising is a crucial use of online commercial information sharing. Advertising is the predominant mechanism for financing the internet content we all enjoy. This is not surprising: For centuries, advertising has been vital to financing news and entertainment, whether it is newspapers, magazines, radio, or television. Although pure subscription models exist, where consumers pay directly for content without advertising, most such content is advertiser supported. Consumer behavior

has made clear that most consumers most of the time will not pay enough to avoid the commercials that support much of our favorite programming.

In the digital advertising economy, what advertisers will pay for an advertisement depends on what they know about the person who will see that advertisement. Just as some audiences are more valuable than others in conventional media, the characteristics of the viewer are an important determinant of the price of online advertising. In turn, the price advertisers will pay determines the revenue available to support online content. Anonymity may be attractive to individual viewers, but it reduces the value of advertising and the revenue available to support the content that the viewer enjoys for "free." It is, in effect, a subtle form of free riding on the contributions of others.

The primary source of information about viewers in online advertising markets is data derived from tracking cookies from which advertisers develop a profile of the user's browsing behavior and the kind of sites likely to be of interest. That information in turn helps predict a viewer's likely interest in a particular advertisement.

The effect of information on advertising prices is relatively large. In two separate studies, one of us has examined the impact of better information on the price of digital advertising. A 2010 study of advertising networks examined prices for behaviorally targeted advertising on a cost per thousand basis. (Behavioral targeting uses browsing history to make better predictions about likely interests.) The study found that such advertising sold for nearly three times the price of "run of network" advertising that might appear anywhere in the advertising network.[6]

A second study, in 2013, examined the impact of additional information on the price of advertising exposures in two real-time advertising auction markets, finding that more information led to a significant price premium. In particular, if a cookie was available with the impression, the price was roughly three times higher than without a cookie. The longer the cookie had been in place, the higher the price of the advertisement. Moreover, the study found that advertising revenue derived from such third party sales was particularly important to smaller web publishers. Even the largest websites sold almost half of their advertising through these channels, while the smaller websites, sold more than two thirds of advertising through third parties.[7]

Third party advertising intermediaries are an important part of the online marketplace, and the most likely competition for the companies that today dominate online advertising, Google and Facebook. These smaller firms, however, likely are more vulnerable to adverse effects of regulatory intervention, particularly if privacy legislation follows the GDPR model of enhanced consent. Well-known consumer facing companies have an inherent advantage in obtaining consent – not because they are necessarily more trustworthy, but simply because they are known. Consumers are less likely to grant consent to companies they have never heard of—for example, 33across, Accuen,

[6] Howard Beales, "The Value of Behavioral Targeting," published online by Network Advertising Initiative, available at http://www.networkadvertising.org/pdfs/Beales_NAI_Study.pdf, March, 2010.

[7] J. Howard Beales & Jeffrey A. Eisenach, "An Empirical Analysis of the Value of Information Sharing in the Market for Online Content," published online by Digital Advertising Alliance, available at http://www.aboutads.info/resource/fullvalueinfostudy.pdf, January, 2014.

Acuity, or Adara—which happen to be the first four members listed for the Network Advertising Initiative, a self-regulatory organization for third party advertising providers. As with other areas of regulation, large and well known companies have incentives to support privacy rules that insulate them from competitive pressures. That outcome harms consumers.

Current privacy discussions sometimes promote the value of "transparency." Why? Most of us have no idea what programs or files are pre-loaded on our newly purchased computer, and we should not need to care. We do not know how the anti-lock brakes or the anti-skid features on our cars actually operate, and we should not need to care. Most of us have no idea of the number of intermediaries that handle a simple credit card transaction, we certainly have no idea who those intermediaries might be, and they may well be different in the transactions in which we engage. We should not have to care.

A more sensible goal of privacy protection, however, is not pursuing transparency; rather, it is making transparency unnecessary. We should be able to rely on the fact that our computer does not include malicious software that will destroy our data, that our automotive safety features will not kill us, and that our credit card transactions will not lead to identity theft. That should be the goal of privacy regulation as well.

Transparency is a means to an end, not an end in itself. When the government is an actor, transparency is often critical, because numerous potentially affected parties can, and likely will, scrutinize the information the government is seeking and how it intends to use it. That scrutiny is an important protector of our liberties. How a commercial enterprise uses data is far less susceptible to influence via transparency. Tellingly, if policing of commercial practice by those who understand the details of information technology is the goal, we need to reverse public policy toward privacy policies. To enable "advocate overseers" of commercial privacy practices, we need privacy policies that provide more detail, not less, and that are likely far more difficult for those who are not technologically sophisticated to understand.

Conclusion

Relying on consumer control to protect the privacy of commercial information is a chimera. Information about commercial interactions necessarily, and appropriately, belongs to both parties to the transaction. Either party may need the information, and may need to use it (or benefit from using it) in ways that are difficult to anticipate at the time of the transaction. There is no basis for assigning sole ownership to one party or the other. Information often has significant value when it is used for purposes different than those for which it was originally collected. Attempts to limit use to "agreed upon" uses will inevitably preclude valuable secondary uses of data that have yet to be developed. As noted above, fraud control is a clear example, because models are often built on data that were collected for entirely different purposes.

Privacy protection should focus on consumers, not data. We should seek to identify, and prevent, harmful uses of data that likely harm consumers, rather than relying on consumers to understand the nuances of information sharing and information use to protect themselves. Notice and choice, or its close cousin transparency, is a distraction, not a solution

REVIEW

Privacy and human behavior in the age of information

Alessandro Acquisti,[1]* Laura Brandimarte,[1] George Loewenstein[2]

This Review summarizes and draws connections between diverse streams of empirical research on privacy behavior. We use three themes to connect insights from social and behavioral sciences: people's uncertainty about the consequences of privacy-related behaviors and their own preferences over those consequences; the context-dependence of people's concern, or lack thereof, about privacy; and the degree to which privacy concerns are malleable—manipulable by commercial and governmental interests. Organizing our discussion by these themes, we offer observations concerning the role of public policy in the protection of privacy in the information age.

If this is the age of information, then privacy is the issue of our times. Activities that were once private or shared with the few now leave trails of data that expose our interests, traits, beliefs, and intentions. We communicate using e-mails, texts, and social media; find partners on dating sites; learn via online courses; seek responses to mundane and sensitive questions using search engines; read news and books in the cloud; navigate streets with geotracking systems; and celebrate our newborns, and mourn our dead, on social media profiles. Through these and other activities, we reveal information—both knowingly and unwittingly—to one another, to commercial entities, and to our governments. The monitoring of personal information is ubiquitous; its storage is so durable as to render one's past undeletable (*1*)—a modern digital skeleton in the closet. Accompanying the acceleration in data collection are steady advancements in the ability to aggregate, analyze, and draw sensitive inferences from individuals' data (*2*).

Both firms and individuals can benefit from the sharing of once hidden data and from the application of increasingly sophisticated analytics to larger and more interconnected databases (*3*). So too can society as a whole—for instance, when electronic medical records are combined to observe novel drug interactions (*4*). On the other hand, the potential for personal data to be abused—for economic and social discrimination, hidden influence and manipulation, coercion, or censorship—is alarming. The erosion of privacy can threaten our autonomy, not merely as consumers but as citizens (*5*). Sharing more personal data does not necessarily always translate into more progress, efficiency, or equality (*6*).

Because of the seismic nature of these developments, there has been considerable debate about individuals' ability to navigate a rapidly evolving privacy landscape, and about what, if anything, should be done about privacy at a policy level. Some trust people's ability to make self-interested decisions about information disclosing and withholding. Those holding this view tend to see regulatory protection of privacy as interfering with the fundamentally benign trajectory of information technologies and the benefits such technologies may unlock (*7*). Others are concerned about the ability of individuals to manage privacy amid increasingly complex trade-offs. Traditional tools for privacy decision-making such as choice and consent, according to this perspective, no longer provide adequate protection (*8*). Instead of individual responsibility, regulatory intervention may be needed to balance the interests of the subjects of data against the power of commercial entities and governments holding that data.

Are individuals up to the challenge of navigating privacy in the information age? To address this question, we review diverse streams of empirical privacy research from the social and behavioral sciences. We highlight factors that influence decisions to protect or surrender privacy and how, in turn, privacy protections or violations affect people's behavior. Information technologies have progressively encroached on every aspect of our personal and professional lives. Thus, the problem of control over personal data has become inextricably linked to problems of personal choice, autonomy, and socioeconomic power. Accordingly, this Review focuses on the concept of, and literature around, informational privacy (that is, privacy of personal data) but also touches on other conceptions of privacy, such as anonymity or seclusion. Such notions all ultimately relate to the permeable yet pivotal boundaries between public and private (*9*).

We use three themes to organize and draw connections between streams of privacy research that, in many cases, have unfolded independently. The first theme is people's uncertainty about the nature of privacy trade-offs, and their own preferences over them. The second is the powerful context-dependence of privacy preferences: The same person can in some situations be oblivious to, but in other situations be acutely concerned about, issues of privacy. The third theme is the malleability of privacy preferences, by which we mean that privacy preferences are subject to influence by those possessing greater insight into their determinants. Although most individuals are probably unaware of the diverse influences on their concern about privacy, entities whose interests depend on information revelation by others are not. The manipulation of subtle factors that activate or suppress privacy concern can be seen in myriad realms—such as the choice of sharing defaults on social networks, or the provision of greater control on social media—which creates an illusion of safety and encourages greater sharing.

Uncertainty, context-dependence, and malleability are closely connected. Context-dependence is amplified by uncertainty. Because people are often "at sea" when it comes to the consequences of, and their feelings about, privacy, they cast around for cues to guide their behavior. Privacy preferences and behaviors are, in turn, malleable and subject to influence in large part because they are context-dependent and because those with an interest in information divulgence are able to manipulate context to their advantage.

Uncertainty

Individuals manage the boundaries between their private and public spheres in numerous ways: via separateness, reserve, or anonymity (*10*); by protecting personal information; but also through deception and dissimulation (*11*). People establish such boundaries for many reasons, including the need for intimacy and psychological respite and the desire for protection from social influence and control (*12*). Sometimes, these motivations are so visceral and primal that privacy-seeking behavior emerges swiftly and naturally. This is often the case when physical privacy is intruded—such as when a stranger encroaches in one's personal space (*13*–*15*) or demonstratively eavesdrops on a conversation. However, at other times (often including when informational privacy is at stake) people experience considerable uncertainty about whether, and to what degree, they should be concerned about privacy.

A first and most obvious source of privacy uncertainty arises from incomplete and asymmetric information. Advancements in information technology have made the collection and usage of personal data often invisible. As a result, individuals rarely have clear knowledge of what information other people, firms, and governments have about them or how that information is used and with what consequences. To the extent that people lack such information, or are aware of their ignorance, they are likely to be uncertain about how much information to share.

Two factors exacerbate the difficulty of ascertaining the potential consequences of privacy behavior. First, whereas some privacy harms are tangible, such as the financial costs associated with identity theft, many others, such as having strangers become aware of one's life history, are intangible. Second, privacy is rarely an unalloyed good; it typically involves trade-offs (*16*). For example, ensuring the privacy of a consumer's

[1]H. John Heinz III College, Carnegie Mellon University, Pittsburgh, PA, USA. [2]Dietrich College, Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, USA.
*Corresponding author. E-mail: acquisti@andrew.cmu.edu

purchases may protect her from price discrimination but also deny her the potential benefits of targeted offers and advertisements.

Elements that mitigate one or both of these exacerbating factors, by either increasing the tangibility of privacy harms or making trade-offs explicit and simple to understand, will generally affect privacy-related decisions. This is illustrated by one laboratory experiment in which participants were asked to use a specially designed search engine to find online merchants and purchase from them, with their own credit cards, either a set of batteries or a sex toy (*17*). When the search engine only provided links to the merchants' sites and a comparison of the products' prices from the different sellers, a majority of participants did not pay any attention to the merchants' privacy policies; they purchased from those offering the lowest price. However, when the search engine also provided participants with salient, easily accessible information about the differences in privacy protection afforded by the various merchants, a majority of participants paid a roughly 5% premium to buy products from (and share their credit card information with) more privacy-protecting merchants.

A second source of privacy uncertainty relates to preferences. Even when aware of the consequences of privacy decisions, people are still likely to be uncertain about their own privacy preferences. Research on preference uncertainty (*18*) shows that individuals often have little sense of how much they like goods, services, or other people. Privacy does not seem to be an exception. This can be illustrated by research in which people were asked sensitive and potentially incriminating questions either point-blank, or followed by credible assurances of confidentiality (*19*). Although logically such assurances should lead to greater divulgence, they often had the opposite effect because they elevated respondents' privacy concerns, which without assurances would have remained dormant.

The remarkable uncertainty of privacy preferences comes into play in efforts to measure individual and group differences in preference for privacy (*20*). For example, Westin (*21*) famously used broad (that is, not contextually specific) privacy questions in surveys to cluster individuals into privacy segments: privacy fundamentalists, pragmatists, and unconcerned. When asked directly, many people fall in the first segment: They profess to care a lot about privacy and express particular concern over losing control of their personal information or others gaining unauthorized access to it (*22*, *23*). However, doubts about the power of attitudinal scales to predict actual privacy behavior arose early in the literature (*24*). This discrepancy between attitudes and behaviors has become known as the "privacy paradox."

In one early study illustrating the paradox, participants were first classified into categories of privacy concern inspired by Westin's categorization based on their responses to a survey dealing with attitudes toward sharing data (*25*). Next, they were presented with products to purchase at a discount with the assistance of an anthropomorphic shopping agent. Few, regardless of the group they were categorized in, exhibited much reluctance to answering the increasingly sensitive questions the agent plied them with.

Why do people who claim to care about privacy often show little concern about it in their daily behavior? One possibility is that the paradox is illusory—that privacy attitudes, which are defined broadly, and intentions and behaviors, which are defined narrowly, should not be expected to be closely related (*26*, *27*). Thus, one might care deeply about privacy in general but, depending on the costs and benefits prevailing in a specific situation, seek or not seek privacy protection (*28*) .

This explanation for the privacy paradox, however, is not entirely satisfactory for two reasons. The first is that it fails to account for situations in which attitude-behavior dichotomies arise under high correspondence between expressed concerns and behavioral actions. For example, one study compared attitudinal survey answers to actual social media behavior (*29*). Even within the subset of participants who expressed the highest degree of concern over strangers being able to easily find out their sexual orientation, political views, and partners' names, 48% did in fact publicly reveal their sexual orientation online, 47% revealed their political orientation, and 21% revealed their current partner's name. The second reason is that privacy decision-making is only in part the result of a rational "calculus" of costs and benefits (*16*, *28*); it is also affected by misperceptions of those costs and benefits, as well as social norms, emotions, and heuristics. Any of these factors may affect behavior differently from how they affect attitudes. For instance, present-bias can cause even the privacy-conscious to engage in risky revelations of information, if the immediate gratification from disclosure trumps the delayed, and hence discounted, future consequences (*30*).

Preference uncertainty is evident not only in studies that compare stated attitudes with behaviors, but also in those that estimate monetary valuations of privacy. "Explicit" investigations ask people to make direct trade-offs, typically between privacy of data and money. For instance, in a study conducted both in Singapore and the United States, students made a series of hypothetical choices about sharing information with websites that differed in protection of personal information and prices for accessing services (*31*). Using conjoint analysis, the authors concluded that subjects valued protection against errors, improper access, and secondary use of personal information between $30.49 and $44.62. Similar to direct questions about attitudes and intentions, such explicit investigations of privacy valuation spotlight privacy as an issue that respondents should take account of and, as a result, increase the weight they place on privacy in their responses.

Implicit investigations, in contrast, infer valuations of privacy from day-to-day decisions in which privacy is only one of many considerations and is typically not highlighted. Individuals engage in privacy-related transactions all the time, even when the privacy trade-offs may be intangible or when the exchange of personal data may not be a visible or primary component of a transaction. For instance, completing a query on a search engine is akin to selling personal data (one's preferences and contextual interests) to the engine in exchange for a service (search results). "Revealed preference" economic arguments would then conclude that because technologies for information sharing have been enormously successful, whereas technologies for information protection have not, individuals hold overall low valuations of privacy. However, that is not always the case: Although individuals at times give up personal data for small benefits or discounts, at other times they voluntarily incur substantial costs to protect their privacy. Context, as further discussed in the next section, matters.

In fact, attempts to pinpoint exact valuations that people assign to privacy may be misguided, as suggested by research calling into question the stability, and hence validity, of privacy estimates. In one field experiment inspired by the literature on endowment effects (*32*), shoppers at a mall were offered gift cards for participating in a non-sensitive survey. The cards could be used online or in stores, just like debit cards. Participants were given either a $10 "anonymous" gift card (transactions done with that card would not be traceable to the subject) or a $12 trackable card (transactions done with that card would be linked to the name of the subject). Initially, half of the participants were given one type of card, and half the other. Then, they were all offered the opportunity to switch. Some shoppers, for example, were given the anonymous $10 card and were asked whether they would accept $2 to "allow my name to be linked to transactions done with the card"; other subjects were asked whether they would accept a card with $2 less value to "prevent my name from being linked to transactions done with the card." Of the subjects who originally held the less valuable but anonymous card, five times as many (52.1%) chose it and kept it over the other card than did those who originally held the more valuable card (9.7%). This suggests that people value privacy more when they have it than when they do not.

The consistency of preferences for privacy is also complicated by the existence of a powerful countervailing motivation: the desire to be public, share, and disclose. Humans are social animals, and information sharing is a central feature of human connection. Social penetration theory (*33*) suggests that progressively increasing levels of self-disclosure are an essential feature of the natural and desirable evolution of interpersonal relationships from superficial to intimate. Such a progression is only possible when people begin social interactions with a baseline level of privacy. Paradoxically, therefore, privacy provides an essential foundation for intimate disclosure. Similar to privacy, self-disclosure confers numerous objective and subjective benefits, including psychological

and physical health (*34*, *35*). The desire for interaction, socialization, disclosure, and recognition or fame (and, conversely, the fear of anonymous unimportance) are human motives no less fundamental than the need for privacy. The electronic media of the current age provide unprecedented opportunities for acting on them. Through social media, disclosures can build social capital, increase self-esteem (*36*), and fulfill ego needs (*37*). In a series of functional magnetic resonance imaging experiments, self-disclosure was even found to engage neural mechanisms associated with reward; people highly value the ability to share thoughts and feelings with others. Indeed, subjects in one of the experiments were willing to forgo money in order to disclose about themselves (*38*).

Context-dependence

Much evidence suggests that privacy is a universal human need (Box 1) (*39*). However, when people are uncertain about their preferences they often search for cues in their environment to provide guidance. And because cues are a function of context, behavior is as well. Applied to privacy, context-dependence means that individuals can, depending on the situation, exhibit anything ranging from extreme concern to apathy about privacy. Adopting the terminology of Westin, we are all privacy pragmatists, privacy fundamentalists, or privacy unconcerned, depending on time and place (*40*).

The way we construe and negotiate public and private spheres is context-dependent because the boundaries between the two are murky (*41*): The rules people follow for managing privacy vary by situation, are learned over time, and are based on cultural, motivational, and purely situational criteria. For instance, usually we may be more comfortable sharing secrets with friends, but at times we may reveal surprisingly personal information to a stranger on a plane (*42*). The theory of contextual "integrity" posits that social expectations affect our beliefs regarding what is private and what is public, and that such expectations vary with specific contexts (*43*). Thus, seeking privacy in public is not a contradiction; individuals can manage privacy even while sharing information, and even on social media (*44*). For instance, a longitudinal study of actual disclosure behavior of online social network users highlighted that over time, many users increased the amount of personal information revealed to their friends (those connected to them on the network) while simultaneously decreasing the amounts revealed to strangers (those unconnected to them) (Fig. 1) (*45*).

The cues that people use to judge the importance of privacy sometimes result in sensible behavior. For instance, the presence of government regulation has been shown to reduce consumer concern and increase trust; it is a cue that people use to infer the existence of some degree of privacy protection (*46*). In other situations, however, cues can be unrelated, or even negatively related, to normative bases of decision-making. For example, in one online experiment (*47*) individuals were more likely to reveal personal and even incriminating information on a website with an unprofessional and casual design with the banner "How Bad R U" than on a site with a formal interface—even though the site with the formal interface was judged by other respondents to be much safer (Fig. 2). Yet in other situations, it is the physical environment that influences privacy concern and associated behavior (*48*), sometimes even unconsciously. For instance, all else being equal, intimacy of self-disclosure is higher in warm, comfortable rooms, with soft lighting, than in cold rooms with bare cement and overhead fluorescent lighting (*49*).

Some of the cues that influence perceptions of privacy are one's culture and the behavior of other people, either through the mechanism of descriptive norms (imitation) or via reciprocity (*50*). Observing other people reveal information increases the likelihood that one will reveal it oneself (*51*). In one study, survey-takers were asked a series of sensitive personal questions regarding their engagement in illegal or ethically questionable behaviors. After answering each question, participants were provided with information, manipulated unbeknownst to them, about the percentage of other participants who in the same survey had admitted to having engaged in a given behavior. Being provided with information that suggested that a majority of survey takers had admitted a certain questionable behavior increased participants' willingness to disclose their engagement in other, also sensitive, behaviors. Other studies have found that the tendency to reciprocate information disclosure is so ingrained that people will reveal more information even to a computer agent that provides information about itself (*52*). Findings such as this may help to explain the escalating amounts of self-disclosure we witness online: If others are doing it, people seem to reason unconsciously, doing so oneself must be desirable or safe.

Other people's behavior affects privacy concerns in other ways, too. Sharing personal information with others makes them "co-owners" of that information (*53*) and, as such, responsible for its protection. Mismanagement of shared information by one or more co-owners causes "turbulence" of the privacy boundaries and, consequently, negative reactions, including anger or mistrust. In a study of undergraduate Facebook users (*54*), for instance, turbulence of privacy boundaries, as a result of having one's profile exposed to unintended audiences, dramatically increased the odds that a user would restrict profile visibility to friends-only.

Likewise, privacy concerns are often a function of past experiences. When something in an environment changes, such as the introduction of a camera or other monitoring devices, privacy concern is likely to be activated. For instance, surveillance can produce discomfort (*55*) and negatively affect worker productivity (*56*). However, privacy concern, like other motivations, is adaptive; people get used to levels of intrusion that do not

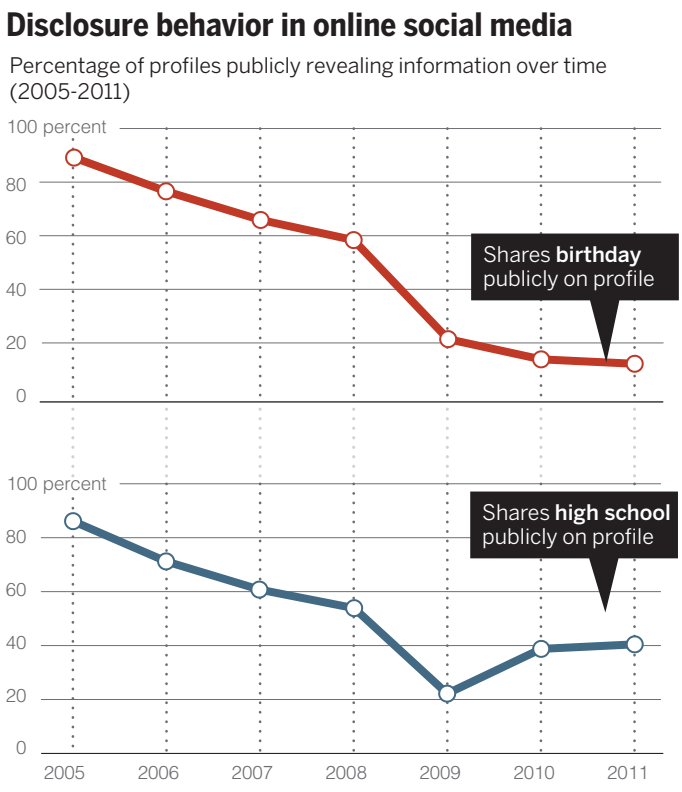

Fig. 1. Endogenous privacy behavior and exogenous shocks. Privacy behavior is affected both by endogenous motivations (for instance, subjective preferences) and exogenous factors (for instance, changes in user interfaces). Over time, the percentage of members in the Carnegie Mellon University Facebook network who chose to publicly reveal personal information decreased dramatically. For instance, over 80% of profiles publicly revealed their birthday in 2005, but less than 20% in 2011. The decreasing trend is not uniform, however. After decreasing for several years, the percentage of profiles that publicly revealed their high school roughly doubled between 2009 and 2010—after Facebook changed the default visibility settings for various fields on its profiles, including high school (bottom), but not birthday (top) (*45*).

change over time. In an experiment conducted in Helsinki (*57*), the installation of sensing and monitoring technology in households led family members initially to change their behavior, particularly in relation to conversations, nudity, and sex. And yet, if they accidentally performed an activity, such as walking naked into the kitchen in front of the sensors, it seemed to have the effect of "breaking the ice"; participants then showed less concern about repeating the behavior. More generally, participants became inured to the presence of the technology over time.

The context-dependence of privacy concern has major implications for the risks associated with modern information and communication technology (*58*). With online interactions, we no longer have a clear sense of the spatial boundaries of our listeners. Who is reading our blog post? Who is looking at our photos online? Adding complexity to privacy decision-making, boundaries between public and private become even less defined in the online world (*59*) where we become social media friends with our coworkers and post pictures to an indistinct flock of followers. With different social groups mixing on the Internet, separating online and offline identities and meeting our and others' expectations regarding privacy becomes more difficult and consequential (*60*).

Malleability and influence

Whereas individuals are often unaware of the diverse factors that determine their concern about privacy in a particular situation, entities whose prosperity depends on information revelation by others are much more sophisticated. With the emergence of the information age, growing institutional and economic interests have developed around disclosure of personal information, from online social networks to behavioral advertising. It is not surprising, therefore, that some entities have an interest in, and have developed expertise in, exploiting behavioral and psychological processes to promote disclosure (*61*). Such efforts play on the malleability of privacy preferences, a term we use to refer to the observation that various, sometimes subtle, factors can be used to activate or suppress privacy concerns, which in turn affect behavior.

Default settings are an important tool used by different entities to affect information disclosure. A large body of research has shown that default settings matter for decisions as important as organ donation and retirement saving (*62*). Sticking to default settings is convenient, and people often interpret default settings as implicit recommendations (*63*). Thus, it is not surprising that default settings for one's profile's visibility on social networks (*64*), or the existence of opt-in or opt-out privacy policies on websites (*65*), affect individuals' privacy behavior (Fig. 3).

In addition to default settings, websites can also use design features that frustrate or even confuse users into disclosing personal information (*66*), a practice that has been referred to as "malicious interface design" (*67*). Another obvious strategy that commercial entities can use to avoid raising privacy concerns is not to "ring alarm bells" when it comes to data collection. When companies do ring them—for example, by using overly fine-tuned personalized advertisements—consumers are alerted (*68*) and can respond with negative "reactance" (*69*).

Various so-called "antecedents" (*70*) affect privacy concerns and can be used to influence privacy behavior. For instance, trust in the entity receiving one's personal data soothes concerns. Moreover, because some interventions that are intended to protect privacy can establish trust, concerns can be muted by the very interventions intended to protect privacy. Perversely, 62% of respondents to a survey believed (incorrectly) that the existence of a privacy policy implied that a site could not share their personal information without permission (*40*), which suggests that simply posting a policy that consumers do not read may lead to misplaced feelings of being protected.

Control is another feature that can inculcate trust and produce paradoxical effects. Perhaps because of its lack of controversiality, control has

Box 1. Privacy: A modern invention?

Is privacy a modern, bourgeois, and distinctly Western invention? Or are privacy needs a universal feature of human societies? Although access to privacy is certainly affected by socioeconomic factors (*87*) [some have referred to privacy as a "luxury good" (*15*)], and privacy norms greatly differ across cultures (*65*, *85*), the need for privacy seems to be a universal human trait. Scholars have uncovered evidence of privacy-seeking behaviors across peoples and cultures separated by time and space: from ancient Rome and Greece (*39*, *88*) to preindustrialized Javanese, Balinese, and Tuareg societies (*89*, *90*). Privacy, as Altman (*91*) noted, appears to be simultaneously culturally specific and culturally universal. Cues of a common human quest for privacy are also found in the texts of ancient religions: The Quran (49:12) instructs against spying on one another (*92*); the Talmud (Bava Batra 60a) advises home-builders to position windows so that they do not directly face those of one's neighbors (*93*); the Bible (Genesis, 3:7) relates how Adam and Eve discovered their nakedness after eating the fruit of knowledge and covered themselves in shame from the prying eyes of God (*94*) [a discussion of privacy in Confucian and Taoist cultures is available in (*95*)]. Implicit in this heterogeneous selection of historical examples is the observation that there exist multiple notions of privacy. Although contemporary attention focuses on informational privacy, privacy has been also construed as territorial and physical, and linked to concepts as diverse as surveillance, exposure, intrusion, insecurity, appropriation, as well as secrecy, protection, anonymity, dignity, or even freedom [a taxonomy is provided in (*9*)].

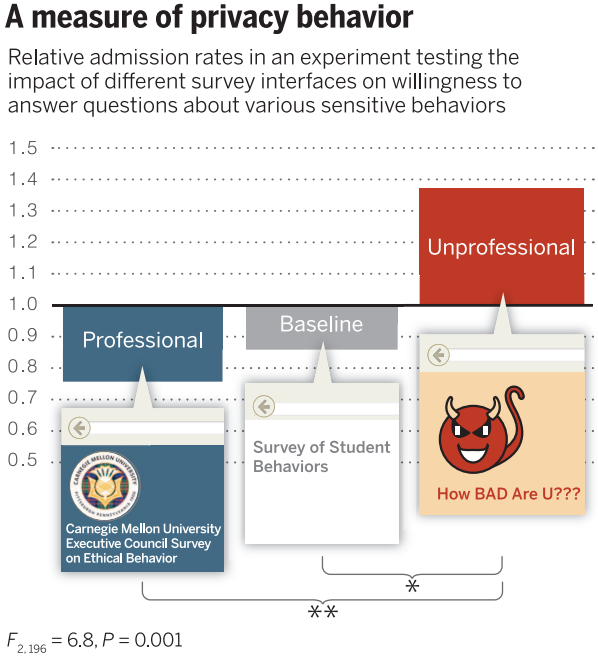

Fig. 2. The impact of cues on disclosure behavior. A measure of privacy behavior often used in empirical studies is a subject's willingness to answer personal, sometimes sensitive questions—for instance, by admitting or denying having engaged in questionable behaviors. In an online experiment (*47*), individuals were asked a series of intrusive questions about their behaviors, such as "Have you ever tried to peek at someone else's e-mail without them knowing?" Across conditions, the interface of the questionnaire was manipulated to look more or less professional. The *y* axis captures the mean affirmative admission rates (AARs) to questions that were rated as intrusive (the proportion of questions answered affirmatively) normed, question by question, on the overall average AAR for the question. Subjects revealed more personal and even incriminating information on the website with a more casual design, even though the site with the formal interface was judged by other respondents to be much safer. The study illustrates how cues can influence privacy behavior in a fashion that is unrelated, or even negatively related, to normative bases of decision-making.

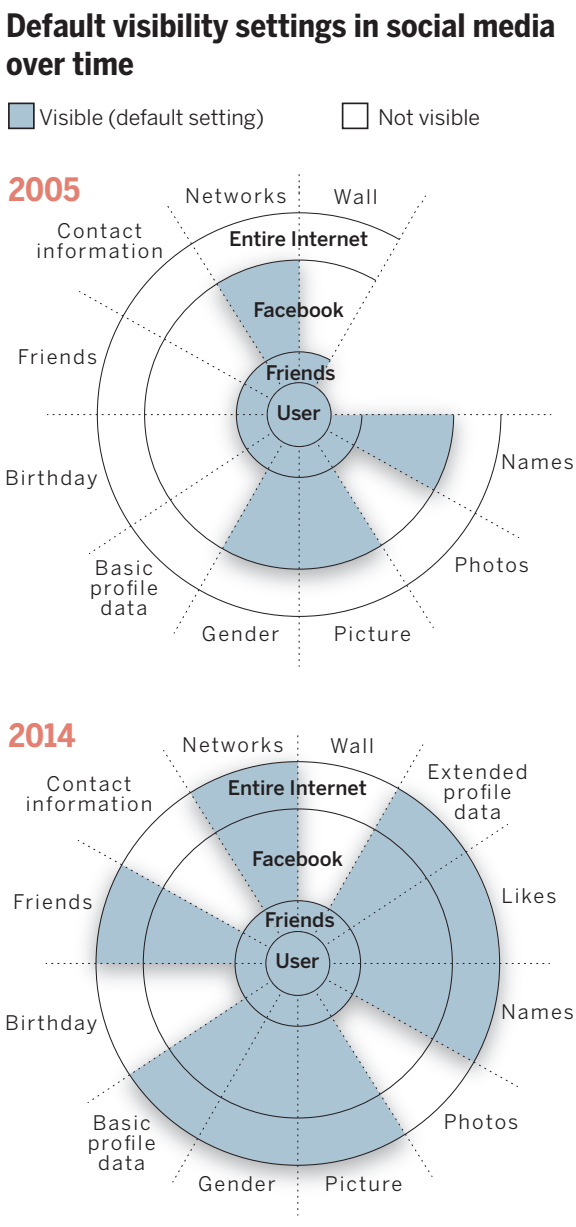

Fig. 3. Changes in Facebook default profile visibility settings over time (2005–2014). Over time, Facebook profiles included an increasing amount of fields and, therefore, types of data. In addition, default visibility settings became more revelatory between 2005 (top) and 2014 (bottom), disclosing more personal information to larger audiences, unless the user manually overrode the defaults (fields such as "Likes" and "Extended Profile Data" did not exist in 2005). "Basic profile data" includes hometown, current city, high school, school (status, concentration, secondary concentration), interested in, relationship, workplace, about you, and quotes. Examples of "Extended profile data" include life events such as new job, new school, engagement, expecting a baby, moved, bought a home, and so forth. "Picture" refers to the main profile image. "Photos" refers to the additional images that users might have shared in their account. "Names" refers to the real name, the username, and the user ID. This figure is based on the authors' data and the original visualization created by M. McKeon, available at http://mattmckeon.com/facebook-privacy.

been one of the capstones of the focus of both industry and policy-makers in attempts to balance privacy needs against the value of sharing. Control over personal information is often perceived as a critical feature of privacy protection (*39*). In principle, it does provide users with the means to manage access to their personal information. Research, however, shows that control can reduce privacy concern (*46*), which in turn can have unintended effects. For instance, one study found that participants who were provided with greater explicit control over whether and how much of their personal information researchers could publish ended up sharing more sensitive information with a broader audience—the opposite of the ostensible purpose of providing such control (*71*).

Similar to the normative perspective on control, increasing the transparency of firms' data practices would seem to be desirable. However, transparency mechanisms can be easily rendered ineffective. Research has highlighted not only that an overwhelming majority of Internet users do not read privacy policies (*72*), but also that few users would benefit from doing so; nearly half of a sample of online privacy policies were found to be written in language beyond the grasp of most Internet users (*73*). Indeed, and somewhat amusingly, it has been estimated that the aggregate opportunity cost if U.S. consumers actually read the privacy policies of the sites they visit would be $781 billion/year (*74*).

Although uncertainty and context-dependence lead naturally to malleability and manipulation, not all malleability is necessarily sinister. Consider monitoring. Although monitoring can cause discomfort and reduce productivity, the feeling of being observed and accountable can induce people to engage in prosocial behaviors or (for better or for worse) adhere to social norms (*75*). Prosocial behavior can be heightened by monitoring cues as simple as three dots in a stylized face configuration (*76*). By the same token, the depersonalization induced by computer-mediated interaction (*77*), either in the form of lack of identifiability or of visual anonymity (*78*), can have beneficial effects, such as increasing truthful responses to sensitive surveys (*79*, *80*). Whether elevating or suppressing privacy concerns is socially beneficial critically depends, yet again, on context [a meta-analysis of the impact of de-identification on behavior is provided in (*81*)]. For example, perceptions of anonymity can alternatively lead to dishonest or prosocial behavior. Illusory anonymity induced by darkness caused participants in an experiment (*82*) to cheat in order to gain more money. This can be interpreted as a form of disinhibition effect (*83*), by which perceived anonymity licenses people to act in ways that they would otherwise not even consider. In other circumstances, though, anonymity leads to prosocial behavior—for instance, higher willingness to share money in a dictator game, when coupled with priming of religiosity (*84*).

Conclusions

Norms and behaviors regarding private and public realms greatly differ across cultures (*85*). Americans, for example, are reputed to be more open about sexual matters than are the Chinese, whereas the latter are more open about financial matters (such as income, cost of home, and possessions). And even within cultures, people differ substantially in how much they care about privacy and what information they treat as private. And as we have sought to highlight in this Review, privacy concerns can vary dramatically for the same individual, and for societies, over time.

If privacy behaviors are culture- and context-dependent, however, the dilemma of what to share and what to keep private is universal across societies and over human history. The task of navigating those boundaries, and the consequences of mismanaging them, have grown increasingly complex and fateful in the information age, to the point that our natural instincts seem not nearly adequate.

In this Review, we used three themes to organize and draw connections between the social and behavioral science literatures on privacy and behavior. We end the Review with a brief discussion of the reviewed literature's relevance to privacy policy.

Uncertainty and context-dependence imply that people cannot always be counted on to navigate the complex trade-offs involving privacy in a self-interested fashion. People are often unaware of the information they are sharing, unaware of how it can be used, and even in the rare situations when they have full knowledge of the consequences of sharing, uncertain about their own preferences. Malleability, in turn, implies that people are easily influenced in what and how much they disclose. Moreover, what they share can be used to influence their emotions, thoughts, and behaviors in many aspects of their lives, as individuals, consumers, and citizens. Although such influence is not always or necessarily malevolent or dangerous, relinquishing control over one's personal data and over one's privacy alters the

balance of power between those holding the data and those who are the subjects of that data.

Insights from the social and behavioral empirical research on privacy reviewed here suggest that policy approaches that rely exclusively on informing or "empowering" the individual are unlikely to provide adequate protection against the risks posed by recent information technologies. Consider transparency and control, two principles conceived as necessary conditions for privacy protection. The research we highlighted shows that they may provide insufficient protections and even backfire when used apart from other principles of privacy protection.

The research reviewed here suggests that if the goal of policy is to adequately protect privacy (as we believe it should be), then we need policies that protect individuals with minimal requirement of informed and rational decision-making—policies that include a baseline framework of protection, such as the principles embedded in the so-called fair information practices (*86*). People need assistance and even protection to aid in navigating what is otherwise a very uneven playing field. As highlighted by our discussion, a goal of public policy should be to achieve a more even equity of power between individuals, consumers, and citizens on the one hand and, on the other, the data holders such as governments and corporations that currently have the upper hand. To be effective, privacy policy should protect real people—who are naïve, uncertain, and vulnerable—and should be sufficiently flexible to evolve with the emerging unpredictable complexities of the information age.

REFERENCES AND NOTES

1. V. Mayer-Schönberger, *Delete: The Virtue of Forgetting In the Digital Age* (Princeton Univ. Press, Princeton, 2011).
2. L. Sweeney, *Int. J. Uncert. Fuzziness Knowl. Based Syst.* **10**, 557–570 (2002).
3. A. McAfee, E. Brynjolfsson, *Harv. Bus. Rev.* **90**, 60–66, 68, 128 (2012).
4. N. P. Tatonetti, P. P. Ye, R. Daneshjou, R. B. Altman, *Sci. Transl. Med.* **4**, 125ra31 (2012).
5. J. E. Cohen, *Stanford Law Rev.* **52**, 1373–1438 (2000).
6. K. Crawford, K. Miltner, M. L. Gray, *Int. J. Commun.* **8**, 1663–1672 (2014).
7. R. A. Posner, *Am. Econ. Rev.* **71**, 405–409 (1981).
8. D. J. Solove, *Harv. Law Rev.* **126**, 1880–1903 (2013).
9. D. J. Solove, *Univ. Penn. L. Rev.* **154**, 477–564 (2006).
10. F. Schoeman, Ed., *Philosophical dimensions of privacy—An anthology* (Cambridge Univ. Press, New York, 1984).
11. B. M. DePaulo, C. Wetzel, R. Weylin Sternglanz, M. J. W. Wilson, *J. Soc. Issues* **59**, 391–410 (2003).
12. S. T. Margulis, *J. Soc. Issues* **59**, 243–261 (2003).
13. E. Goffman, *Relations in Public: Microstudies of the Public Order* (Harper & Row, New York, 1971).
14. E. Sundstrom, I. Altman, *Hum. Ecol.* **4**, 47–67 (1976).
15. B. Schwartz, *Am. J. Sociol.* **73**, 741–752 (1968).
16. R. S. Laufer, M. Wolfe, *J. Soc. Issues* **33**, 22–42 (1977).
17. J. Y. Tsai, S. Egelman, L. Cranor, A. Acquisti, *Inf. Syst. Res.* **22**, 254–268 (2011).
18. P. Slovic, *Am. Psychol.* **50**, 364–371 (1995).
19. E. Singer, H. Hippler, N. Schwarz, *Int. J. Public Opin. Res.* **4**, 256–268 (1992).
20. V. P. Skotko, D. Langmeyer, *Sociometry* **40**, 178–182 (1977).
21. A. Westin, Harris Louis & Associates, Harris-Equifax Consumer Privacy Survey (Tech. rep. 1991).
22. M. J. Culnan, P. K. Armstrong, *Organ. Sci.* **10**, 104–115 (1999).
23. H. J. Smith, S. J. Milberg, S. J. Burke, *Manage. Inf. Syst. Q.* **20**, 167–196 (1996).
24. B. Lubin, R. L. Harrison, *Psychol. Rep.* **15**, 77–78 (1964).
25. S. Spiekermann, J. Grossklags, B. Berendt, *E-Privacy in 2nd Generation E-Commerce: Privacy Preferences versus Actual Behavior* (Third ACM Conference on Electronic Commerce, Tampa, 2001), pp. 38–47.
26. P. A. Norberg, D. R. Horne, D. A. Horne, *J. Consum. Aff.* **41**, 100–126 (2007).
27. I. Ajzen, M. Fishbein, *Psychol. Bull.* **84**, 888–918 (1977).
28. P. H. Klopfer, D. I. Rubenstein, *J. Soc. Issues* **33**, 52–65 (1977).
29. A. Acquisti, R. Gross, in *Privacy Enhancing Technologies*, G. Danezis, P. Golle Eds. (Springer, New York, 2006), pp. 36–58.
30. A. Acquisti, *Privacy in Electronic Commerce and the Economics of Immediate Gratification* (Fifth ACM Conference on Electronic Commerce, New York, 2004), pp. 21–29.
31. I. Hann, K. Hui, S. T. Lee, I. P. L. Png, *J. Manage. Inf. Syst.* **24**, 13–42 (2007).
32. A. Acquisti, L. K. John, G. Loewenstein, *J. Legal Stud.* **42**, 249–274 (2013).
33. I. Altman, D. Taylor, *Social Penetration: The Development of Interpersonal Relationships* (Holt, Rinehart & Winston, New York, 1973).
34. J. Frattaroli, *Psychol. Bull.* **132**, 823–865 (2006).
35. J. W. Pennebaker, *Behav. Res. Ther.* **31**, 539–548 (1993).
36. C. Steinfield, N. B. Ellison, C. Lampe, *J. Appl. Dev. Psychol.* **29**, 434–445 (2008).
37. C. L. Toma, J. T. Hancock, *Pers. Soc. Psychol. Bull.* **39**, 321–331 (2013).
38. D. I. Tamir, J. P. Mitchell, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8038–8043 (2012).
39. A. Westin, *Privacy and Freedom* (Athenäum, New York, 1967).
40. C. J. Hoofnagle, J. M. Urban, *Wake Forest Law Rev.* **49**, 261–321 (2014).
41. G. Marx, *Ethics Inf. Technol.* **3**, 157–169 (2001).
42. J. W. Thibaut, H. H. Kelley, *The Social Psychology Of Groups* (Wiley, Oxford, 1959).
43. H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford Univ. Press, Redwood City, 2009).
44. d. boyd, *It's Complicated: The Social Lives of Networked Teens* (Yale Univ. Press, New Haven, 2014).
45. F. Stutzman, R. Gross, A. Acquisti, *J. Priv. Confidential.* **4**, 7–41 (2013).
46. H. Xu, H. H. Teo, B. C. Tan, R. Agarwal, *J. Manage. Inf. Syst.* **26**, 135–174 (2009).
47. L. K. John, A. Acquisti, G. Loewenstein, *J. Consum. Res.* **37**, 858–873 (2011).
48. I. Altman, *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding* (Cole, Monterey, 1975).
49. A. L. Chaikin, V. J. Derlega, S. J. Miller, *J. Couns. Psychol.* **23**, 479–481 (1976).
50. V. J. Derlega, A. L. Chaikin, *J. Soc. Issues* **33**, 102–115 (1977).
51. A. Acquisti, L. K. John, G. Loewenstein, *J. Mark. Res.* **49**, 160–174 (2012).
52. Y. Moon, *J. Consum. Res.* **26**, 323–339 (2000).
53. S. Petronio, *Boundaries of Privacy: Dialectics of Disclosure* (SUNY Press, Albany, 2002).
54. F. Stutzman, J. Kramer-Duffield, *Friends Only: Examining a Privacy-Enhancing Behavior in Facebook* (SIGCHI Conference on Human Factors in Computing Systems, ACM, Atlanta, 2010), pp. 1553–1562.
55. T. Honess, E. Charman, *Closed Circuit Television in Public Places: Its Acceptability and Perceived Effectiveness* (Police Research Group, London, 1992).
56. M. Gagné, E. L. Deci, *J. Organ. Behav.* **26**, 331–362 (2005).
57. A. Oulasvirta *et al*., *Long-Term Effects of Ubiquitous Surveillance in the Home* (ACM Conference on Ubiquitous Computing, Pittsburgh, 2012), pp. 41–50.
58. L. Palen, P. Dourish, *Unpacking "Privacy" For A Networked World* (SIGCHI Conference on Human Factors in Computing Systems, ACM, Fort Lauderdale, 2003), pp. 129–136.
59. Z. Tufekci, *Bull. Sci. Technol. Soc.* **28**, 20–36 (2008).
60. J. A. Bargh, K. Y. A. McKenna, G. M. Fitzsimons, *J. Soc. Issues* **58**, 33–48 (2002).
61. R. Calo, *Geo. Wash. L. Rev.* **82**, 995–1304 (2014).
62. E. J. Johnson, D. Goldstein, *Science* **302**, 1338–1339 (2003).
63. C. R. McKenzie, M. J. Liersch, S. R. Finkelstein, *Psychol. Sci.* **17**, 414–420 (2006).
64. R. Gross, A. Acquisti, *Information Revelation and Privacy in Online Social Networks* (ACM Workshop–Privacy in the Electronic Society, New York, 2005), pp. 71–80.
65. E. J. Johnson, S. Bellman, G. L. Lohse, *Mark. Lett.* **13**, 5–15 (2002).
66. W. Hartzog, *Am. Univ. L. Rev.* **60**, 1635–1671 (2010).
67. G. Conti, E. Sobiesk, *Malicious Interface Design: Exploiting the User* (19th International Conference on World Wide Web, ACM, Raleigh, 2010), pp. 271–280.
68. A. Goldfarb, C. Tucker, *Mark. Sci.* **30**, 389–404 (2011).
69. T. B. White, D. L. Zahay, H. Thorbjørnsen, S. Shavitt, *Mark. Lett.* **19**, 39–50 (2008).
70. H. J. Smith, T. Dinev, H. Xu, *Manage. Inf. Syst. Q.* **35**, 989–1016 (2011).
71. L. Brandimarte, A. Acquisti, G. Loewenstein, *Soc. Psychol. Personal. Sci.* **4**, 340–347 (2013).
72. C. Jensen, C. Potts, C. Jensen, *Int. J. Hum. Comput. Stud.* **63**, 203–227 (2005).
73. C. Jensen, C. Potts, *Privacy Policies as Decision-Making Tools: An Evaluation of Online Privacy Notices* (SIGCHI Conference on Human factors in computing systems, ACM, Vienna, 2004), pp. 471–478.
74. A. M. McDonald, L. F. Cranor, *I/S: J. L. Policy Inf. Society.* **4**, 540–565 (2008).
75. C. Wedekind, M. Milinski, *Science* **288**, 850–852 (2000).
76. M. Rigdon, K. Ishii, M. Watabe, S. Kitayama, *J. Econ. Psychol.* **30**, 358–367 (2009).
77. S. Kiesler, J. Siegel, T. W. McGuire, *Am. Psychol.* **39**, 1123–1134 (1984).
78. A. N. Joinson, *Eur. J. Soc. Psychol.* **31**, 177–192 (2001).
79. S. Weisband, S. Kiesler, *Self Disclosure On Computer Forms: Meta-Analysis And Implications* (SIGCHI Conference Conference on Human Factors in Computing Systems, ACM, Vancouver, 1996), pp. 3–10.
80. R. Tourangeau, T. Yan, *Psychol. Bull.* **133**, 859–883 (2007).
81. T. Postmes, R. Spears, *Psychol. Bull.* **123**, 238–259 (1998).
82. C. B. Zhong, V. K. Bohns, F. Gino, *Psychol. Sci.* **21**, 311–314 (2010).
83. J. Suler, *Cyberpsychol. Behav.* **7**, 321–326 (2004).
84. A. F. Shariff, A. Norenzayan, *Psychol. Sci.* **18**, 803–809 (2007).
85. B. Moore, *Privacy: Studies in Social and Cultural History* (Armonk, New York, 1984).
86. *Records, Computers and the Rights of Citizens* (Secretary's Advisory Committee, U.S. Dept. of Health, Education and Welfare, Washington, DC, 1973).
87. E. Hargittai, in *Social Stratification*, D. Grusky Ed. (Westview, Boulder, 2008), pp. 936-113.
88. P. Ariès, G. Duby (Eds.), *A History of Private Life: From Pagan Rome to Byzantium* (Harvard Univ. Press, Cambridge, 1992).
89. R. F. Murphy, *Am. Anthropol.* **66**, 1257–1274 (1964).
90. A. Westin, in *Philosophical Dimensions of Privacy: An Anthology*, F.D. Schoeman Ed. (Cambridge Univ. Press, Cambridge, UK, 1984), pp. 56–74.
91. I. Altman, *J. Soc. Issues* **33**, 66–84 (1977).
92. M. A. Hayat, *Inf. Comm. Tech. L.* **16**, 137–148 (2007).
93. A. Enkin, "Privacy," www.torahmusings.com/2012/07/privacy (2014).
94. J. Rykwert, *Soc. Res. (New York)* **68**, 29–40 (2001).
95. C. B. Whitman, in *Individualism and Holism: Studies in Confucian and Taoist Values*, D. J. Munro, Ed. (Center for Chinese Studies, Univ. Michigan, Ann Arbor, 1985), pp. 85–100.

ACKNOWLEDGMENTS

We are deeply grateful to the following individuals: R. Gross and F. Stutzman for data analysis; V. Marotta, V. Radhakrishnan, and S. Samat for research; W. Harsch for graphic design; and A. Adams, I. Adjerid, R. Anderson, E. Barr, C. Bennett, R. Boehme, R. Calo, J. Camp, F. Cate, J. Cohen, D. Cole, M. Culnan, R. De Wolf, J. Donath, S. Egelman, N. Ellison, S. Fienberg, A. Forget, U. Gasser B. Gellman, J. Graves, J. Grimmelmann, J. Grossklags, S. Guerses, J. Hancock, E. Hargittai, W. Hartzog, J. Hong, C. Hoofnagle, J. P. Hubaux, K. L. Hui, A. Joinson, S. Kiesler, J. King, B. Knijnenburg, A. Kobsa, B. Kraut, P. Leon, M. Madden, I. Meeker, D. Mulligan, C. Olivola, E. Peer, S. Petronio, S. Preibusch, J. Reidenberg, S. Romanosky, M. Rotenberg, I. Rubinstein, N. Sadeh, A. Sasse, F. Schaub, P. Shah, R. E. Smith, S. Spiekermann, J. Staddon, L. Strahilevitz, P. Swire, O. Tene, E. VanEpps, J. Vitak, R. Wash, A. Woodruff, H. Xu, and E. Zeide for enormously valuable comments and suggestions.

10.1126/science.aaa1465

THE SUMMER OF HATE SPEECH*

Larry Downes

Project Director, Georgetown Center for Business and Public Policy

Over the past few years, pressure has been building on online platforms to do something — anything — about increasingly hostile, misleading and distasteful Internet content. The 2016 election, Brexit and other polarizing events have brought out some of the worst in human nature, much of it amplified and rapidly disseminated on free digital services.

Growing conflict over who gets to say what to whom and where, of course, is not limited to the Internet. Even here in Berkeley, Calif., the free-speech capital of the world, we have been engulfed in fights — some violent — over which viewpoints should be heard on the UC campus. Berkeley Free Speech Movement founder Mario Savio would not be proud.

But this year, largely unregulated Internet companies have fallen into a black hole of disgruntled users, hyperventilating activists and an angry Congress. Facebook, Twitter, Instagram, YouTube and other social media companies are innovating wildly, implementing increasingly Rube Goldberg-like fixes to adjust their content policies and the technologies that enforce them.

"Users are calling on online platforms to provide a moral code," says Daphne Keller, director of the intermediary liability project at Stanford's Center for Internet and Society. "But we'll never agree on what should come down. Whatever the rules, they'll fail." Humans and technical filters alike, according to Keller, will continue to make "grievous errors."

Do not look to the Constitution to solve the problem. Contrary to popular belief, the First Amendment plays no role in determining when content hosts have gone too far, or not far enough. That is because, as I regularly explain to incredulous students, free-speech protections limit only censorship by governments and then only in the United States.

Some restrictions on foreign nationals — e.g., electioneering — are permitted. With very limited exceptions, private actors can press mute on whomever and whatever they want. Indeed, the Constitution protects the sites from government efforts to impose speech codes — moral or otherwise.

But while the First Amendment does not apply to the practices of Internet companies, the inevitable failure of platform providers to find the "Goldilocks zone" of just-right content moderation underscores the wisdom of the Founding Fathers. Picking and choosing among good and bad

* First published in *The Washington Post*, August 30, 2018

speech is a no-win proposition, no matter how good your intentions.

So here is my advice to tech CEOS: Don't try. Don't moderate, don't filter, don't judge. Allow opinions informed and ignorant alike to circulate freely in what Supreme Court Justice William O. Douglas famously called "the marketplace of ideas." Trust that, sooner or later, truth will prevail over lies and good over evil. Deny yourself the power to interfere, especially at those excruciating moments when the temptation is most irresistible — when the most detestable content is flowering malodorously.

Today, that solution may seem even more unpalatable than it was when the Bill of Rights was being debated nearly 250 years ago. But every day brings new evidence that the alternative of unaccountable private gatekeepers appointing themselves the task of deciding what violates virtual moral codes, especially in the chaos of messy and often ugly political and social disruption, is worse. Much worse.

A sobering report last week on Motherboard, for example, details the "impossible" effort of a beleaguered Facebook to reinvent its "community standards" — a daunting task given the billions of posts a week originating in over a hundred countries. Acceptable content rules are developed unilaterally by a policy team "made up of lawyers, public relations professionals, ex-public policy wonks and crisis management experts."

Enforcement, according to the report, is now the job of about 7,500 low-wage "moderators," deciding case by case whether to remove posts flagged by artificial intelligence software or by complaining users — with the latter assigned a "trustworthiness score." Flowcharts guide the review, asking, for example, whether the challenged post encourages violence, curses or uses slurs against a protected class or is guilty of "comparing them to animals, disease or filth."

National laws and local customs also have to be taken into consideration. The process and the rules are constantly and opaquely updated, often in response to the latest public relations crisis. No surprise few moderators last a year in the job, according to the report.

As one indication of just how fraught the complex system has become, moderators removed a July Fourth post quoting the Declaration of Independence. Why? A reference to "merciless Indian savages" was deemed hate speech.

Yet Facebook's face-plants seem almost trivial compared with the free-speech barbarism of other Internet giants. Consider the social news site Reddit, which three years ago announced a confusing set of changes to its "Content Policy" in an improvised effort to curb sexist posts. Forums dominated by such content were simply erased.

The deleted groups, said then-chief executive Ellen Pao, "break our Reddit rules based on their harassment of individuals," a determination made solely by the company. (Due process is also a government-only requirement.)

After users and volunteer editors revolted over both the policy change and its ham-handed implementation, Reddit's board of directors dismissed Pao and revised yet again the amendments to its policy. But Reddit founder and returning chief executive Steve Huffman still defended the changes. Neither he nor co-founder Alexis

Ohanian, Huffman said, had "created Reddit to be a bastion of free speech, but rather as a place where open and honest discussion can happen."

Except that Ohanian, in an earlier interview, said precisely the opposite, down to the same archaic phrasing. When asked what he thought the Founding Fathers would have made of the site's unregulated free-for-all of opinions, Ohanian boasted, "A bastion of free speech on the World Wide Web? I bet they would like it."

Even worse, consider the approach of website security provider Cloudflare, whose CEO, Matthew Prince, personally terminated the account of the neo-Nazi Daily Stormer after previously defending his company's decision to manage the site's traffic. Prince's reasoned explanation for the change of heart? "I woke up in a bad mood and decided someone shouldn't be allowed on the Internet," he wrote in an internal memo to employees.

In a supreme gesture of having his cake and censoring it too, Prince then condemned his own action, fretting "no one should have that power." But he does. (Activists for "net neutrality," which would prohibit blocking access to any website, notably want restrictions solely for ISPs.)

Refusing to moderate at all would certainly be easier. But could Internet users stomach it? The First Amendment, after all, is nearly absolute. The U.S. Supreme Court has carved out a few narrow exceptions, most of them irrelevant to the current debate over online speech. Discussions of current events and politics, for example, are considered the most protected category of all.

Even the most repulsive opinions are protected from government suppression. As First Amendment scholar Eugene Volokh reminds politicians, "There is no hate speech exception to the First Amendment."

IS THE TECH BACKLASH GOING ASKEW?*

Larry Downes and Blair Levin

Project Director, Georgetown Center for Business and Public Policy

As winter sets in, the dark days for technology companies have been getting longer.

We sympathize with the increased anxiety over the poor data hygiene practices of leading tech platforms. And we agree that legislation clarifying the duties of those who collect and use personal information is important, as is delineating enforcement responsibilities among agencies and jurisdictions.

We're concerned, however, by the tendency of some to shoehorn pet theories into the debate — notably the passionate but incomplete argument that it's time to jettison decades of antitrust policy that limits the government to intervening only when market concentration has, or could, cause higher prices for consumers.

The vague alternative, proposed by critics on the left and right, is a return to a failed framework that boils down to, at best, a general belief that "big is bad" and, at worst, to politically-based payback for companies on the wrong side of an election.

Writing recently in the New York Times, law professor Tim Wu urged antitrust enforcers to launch sweeping lawsuits against Facebook and other "Big Tech" platforms that would likely last a decade or more. Anything less, Wu says, amounts to giving "these companies a pass when it comes to antitrust enforcement, allowing them to dominate their markets and buy up their competitors."

The goal of Wu's approach is not to actually win so much as it is to distract the companies' leaders. The litigation is not a means but the end in itself. Paraphrasing Thomas Jefferson, Wu advocates spilling the corporate equivalent of the "blood of patriots," attacking relentlessly regardless of whether there's actually, you know, a sustainable case.

That logic is oddly aligned with the views of some, including President Trump and his former attorney general, Jeff Sessions, who believe they are justified in threatening companies they view as politically hostile on the fuzzy grounds that there is a "very antitrust situation."

Wu's best example of how this abuse of legal process works was the U.S. government's 13-year crusade in the 1970s to break up IBM. At the time, IBM was the undisputed leader of the computer business.

Though the government was never able to prove the company had, as accused, "undertaken exclusionary and predatory

* First published in *The Washington Post*, January 16, 2019

conduct with the aim and effect of eliminating competition," Wu believes the cost and uncertainty for the company of the extended legal fight saved the U.S. economy, giving personal computers an opening to proliferate and unseat IBM's mainframe computer "monopoly."

Never mind that IBM was the most successful seller of PCs and, through its relationship with Lenovo, still is. The company was certainly hurt by the ultimately abandoned case, as were, in later examples, Microsoft, Intel, Qualcomm and others.

But do antitrust jihads really help consumers more than it hurts them? Probably not. While well-founded prosecutions, such as those leading to the 1982 breakup of AT&T, did open critical markets, that success may not be duplicated elsewhere. In fact, Philip Verveer, a Visiting Fellow at the Harvard Kennedy School and the government's lead counsel in the AT&T case, recently concluded that unleashing antitrust against today's platform companies would amount to little more than "an act of faith that a successful prosecution would bring about benefits."

There's no need to gamble. The more effective regulator of digital markets has always been the happy confluence of engineering and business innovations in hardware, software and communications driving exponential improvements in speed, quality, size, energy usage and, of course, cost.

As computing continues to improve, markets become unsettled, innovation flourishes, and new leaders emerge. It's not the arbitrary release of the "blood of patriots" that best corrects market imbalances. It's the normal cycles of capitalism sped up by disruptive innovation, or what economist Joseph Schumpeter in 1942 famously called "creative destruction."

If the tech sector was immune to that process, as some allege, we would expect stagnant productivity and wage growth, with profits protected and funneled to shareholders.

But that view doesn't square with recent findings from Michael Mandel, chief economist at the Progressive Policy Institute. According to Mandel, who has been measuring the digital economy for decades, the technology sector broadly "accounted for almost half of non-health private sector growth between 2007 and 2017." Technology prices, at the same time, "fell by 15%, compared to a 21 percent increase in the rest of the non-health private sector."

Annual pay for tech workers (including hourly workers at e-commerce fulfillment centers) rose at more than twice the rate of other industries. Job growth in tech was four times faster.

Lower prices, higher pay and growing productivity: That doesn't suggest a problem, or at least not one requiring radical restructuring of the companies driving the gains.

Consider the alternative approach taken in Europe, which has ramped up an aggressive attack of U.S. technology companies, applying the kind of expansive view of competition law urged by Wu and others. European businesses are still largely no-shows in the digital revolution, the result not of monopolies but of the micromanagement of employment, investment and infrastructure by regulators. Rather than freeing up local innovators to benefit European consumers, the European

Union seems content simply to fine successful U.S. businesses.

The European approach highlights another problem with calls for U.S. antitrust enforcers to punish platform companies just for their size. Looking ahead to the technology drivers of the near future, such as artificial intelligence and autonomous vehicles, any hopes for the United States to lead internationally depend on heavy investment today in research and development. Many of the highest-risk bets are being placed by, you guessed it, today's "monopoly" companies.

So what should U.S. regulators do? The starting point is vigilance in applying tried-and-true tools to new harms. The Federal Trade Commission, for example, has already brought over 150 enforcement actions against tech companies in the last decade for violations of consumer protection laws, reaching settlements that in many cases include decades of ongoing oversight and reporting.

The trade agency is gearing up a broad review of Facebook to see whether the company's many embarrassing failures of the past few years amount to violations of a 2011 consent decree or, indeed, new violations. And the commissioners recently told Congress that they want additional resources and authority to better enforce existing law, joining a bipartisan call for targeted legislation, particularly on consumer data collection and use.

Tech's loudest critics argue that the gears of government are turning too slowly. But that's actually another reason why calls to simply throw out measured approaches to regulating competition are dangerous, despite their populist appeal. Even assuming new standards could be developed that wouldn't stall the innovation engine driving the U.S. economy, rewriting federal competition law, realistically, would take Congress and the courts decades to hammer out.

Fortunately, the next wave of disruptive technology is always coming. It won't fix everything. But if history is any guide, it will fix an awful lot — and do so without breaking everything that's actually working.

How More Regulation for U.S. Tech Could Backfire*

Larry Downes

Project Director, Georgetown Center for Business and Public Policy

If 2017 was the year that tech became a lightning rod for dissatisfaction over everything from the last U.S. presidential election to the possibility of a smartphone-driven dystopia, 2018 already looks to be that much worse.

Innovation and its discontents are nothing new, of course, going back at least to the 18th century, when Luddites physically attacked industrial looms. Hostility to the internet appeared the moment the Web became a commercial technology, threatening from the outset to upend traditional businesses and maybe even our deeply embedded beliefs about family, society, and government. George Mason University's Adam Thierer, reviewing a resurgence of books about the "existential threat" of disruptive innovation, has detailed what he calls a "techno-panic template" in how we react to disruptive innovations that don't fit into familiar categories.

But with the proliferation of new products and their reach ever-deeper into our work, home, and personal lives, the relentless tech revolt of the last year shouldn't really have come as any surprise, especially to those of us in Silicon Valley.

Still, the only solution critics can propose for our growing tech malaise is government intervention — usually expressed vaguely as "regulating tech" or "breaking up" the biggest and most successful Internet companies. Break-ups, which require a legal finding that the structure of a company is enabling anti-competitive behavior, seem now to have become a synonym for somehow crippling a successful enterprise.

Of course, nobody thinks technology companies should be left unregulated. Tech companies, like any other enterprise, are already subject to a complex tangle of laws, varying based on industry and local authority. They all pay taxes, report their finances, disclose significant shareholders, and comply with the full range of employment, health and safety, advertising, intellectual property, consumer protection and anti-competition laws, to name just a few.

There are also specialized laws for tech, including limits on how Internet companies can engage with children. In the U.S., commercial drones must be registered with the Federal Aviation Administration. Genetic testing and other health-related

* First published in the *Harvard Business Review*, February 9, 2018

devices must pass muster with the Food and Drug Administration. Increasingly, ride-sharing and casual rental services must meet some of the same standards and inspections as long-time transportation and hospitality incumbents.

There are growing calls, likewise, to regulate social media and video platforms as if they were traditional print or broadcast news sources, even though doing so would almost certainly run afoul of the very free speech protections proponents are hoping to preserve.

But perhaps what tech critics really want are more innovative rules. Traditional regulations, after all, were designed in response to earlier technologies and the market failures they generated. They don't cover largely speculative and mostly future-looking concerns.

What if, for example, artificial intelligence puts an entire generation out of work? What if genetic manipulations accidentally unravel the fabric of DNA, reversing evolution in one fell swoop? What if social media companies learn so much about us that they undermine—intentionally or otherwise—democratic institutions, creating a tyranny of "unregulated" big data controlled by a few unelected young CEOs?

The problem with such speculation is that it is just that. In deliberative government, legislators and regulatory agencies must weigh the often-substantial costs of proposed limits against their likely benefit, balanced against the harm of simply leaving in place the current legal status quo.

But there's no scientific method for estimating the risk of prematurely shutting down experiments that could yield important discoveries. There's no framework for pre-emptively regulating nascent industries and potential new technologies. By definition, they've caused no measurable harm.

In particular, breaking up the most successful Internet and cloud-based companies only looks like a solution. It isn't. Antitrust is meant to punish dominant companies that use their leverage to raise costs for consumers. Yet the services provided by technology companies are often widely available at little or no cost. Many of the products and services of Amazon, Apple, Google, Facebook and Microsoft — the internet giants referred to by the New York Times as "the frightful five" — are free for consumers.

More to the point, break-ups almost always backfire. Think of the former AT&T, which was regulated as a monopoly utility until 1982, when the government changed its mind and split the company into component long-distance and regional phone companies. The sum of the parts actually increased in value — except for the long-distance company, which faded in the face of unregulated new competitors.

Then, over the next 20 years, the regional companies put themselves back together, and, with economies of scale, reemerged as a mobile internet network and Pay TV provider, competing with cable companies and fast-growing internet-based video services including YouTube, Amazon and Netflix. What started as a regulatory punishment for AT&T led to an even bigger network of companies.

On the other hand, the constant threat of a forced divestiture can be disastrous for consumers and enterprise alike. IBM prevailed against multiple efforts to break it up along product lines, but was so shaken by the decades-long experience that the company became dangerously

timid about future innovations, missing the shifts first to client-server and then to Internet-based computing architectures, nearly bankrupting the business.

Microsoft, similarly, was so distracted by its multi-year fight to avoid break-up both by U.S. and European regulators that it lost essential momentum. It mostly missed out on the mobile revolution, and hesitated in responding to open-source alternatives to operating systems, desktop applications, and other software apps that seriously eroded the company's once-formidable competitive advantage. (The company is now growing a cloud services business, but is still far behind Google and Amazon.)

These examples hint at an alternative to random and unproven new forms of regulation for emerging technologies: simply waiting for the next generation of innovations and the entrepreneurs who wield them to disrupt the supposed monopolists right out of their disagreeable behaviors, sometimes fatally.

Today, it might seem that the companies in the frightful five have unbeatable leads in retailing and cloud services, social media, search, advertising, desktop operating systems and mobile devices. But the landscape of business history is littered with the corpses of supposedly invulnerable giants. In our research on wildly successful enterprises who fail to find a second act, Paul Nunes and I note that the average life span of companies on the Standard & Poor's 500 has fallen from 67 years in the 1920s to just 15 years today.

In the early years of the internet age, a half-dozen companies were serially crowned the victor in search, only to be unseated by more innovative technology soon after. Yahoo and others gave way to Google, just as Blackberry faded in response to the iPhone. MySpace (remember them?) collapsed at the introduction of Facebook, which, at the time, was little more than a bit of software from a college student. Napster lost in court (no new laws were needed for that), leaving Apple to define a working market for digital music. And who remembers the alarm bells rung in 2000 when then-dominant ISP America On-Line merged with content behemoth Time Warner?

The best regulator of technology, it seems, is simply more technology. And despite fears that channels are blocked, markets are locked up, and gatekeepers have closed networks that the next generation of entrepreneurs need to reach their audience, somehow they do it anyway — often embarrassingly fast, whether the presumed tyrant being deposed is a long-time incumbent or last year's startup darling.

That, in any case, is the theory on which U.S. policymakers across the political spectrum have nurtured technology-based innovation since the founding of the Republic. Taking the long view, it's clearly been a winning strategy, especially when compared to the more invasive, command-and-control approach taken by the European Union, which continues to lag on every measure of the Internet economy. (Europe's strategy now seems to be little more than to hobble U.S. tech companies and hope for the best.)

Or compared to China, which has built tech giants of its own, but only by limiting outside access to its singularly enormous local market. And always with the risk that too much success by Chinese entrepreneurs may one day crash headfirst

into a political culture that is deeply uncomfortable with the internet's openness.

That solution — to stay the course, to continue leaving tech largely to its own correctives — is cold comfort to those who believe tomorrow's problems, coming up fast in the rear-view mirror, are both unprecedented and catastrophic.

Yet, so far there's no evidence supporting shrill predictions of a technology-driven apocalypse. Or that existing safeguards — both market and legal — won't save us from our worst selves.

Nor have tech's growing list of critics proposed anything more specific than simply calling for "regulation" to save us. Perhaps that's because effective remedies are incredibly hard to design.

FIXING SOCIAL MEDIA'S GRAND BARGAIN*

Jack Balkin

Professor of Constitutional Law, Yale Law School

To regulate social media in the twenty-first century, we should focus on its political economy: the nature of digital capitalism and how we pay for the digital public sphere we have. Our digital public sphere is premised on a grand bargain: free communications services in exchange for pervasive data collection and analysis. This allows companies to sell access to end users to the companies' advertisers and other businesses.

The political economy of digital capitalism creates perverse incentives for social media companies. It encourages companies to surveil, addict, and manipulate their end users and to strike deals with third parties who will further manipulate them.

Treating social media companies as public forums or public utilities is not the proper cure. It may actually make things worse. Even so, social media companies, whether they like it or not, have public obligations. They play important roles in organizing and curating public discussion and they have moral and professional responsibilities both to their end users and to the general public.

A reinvigorated competition law is one important way of dealing with the problems of social media, as I will describe later on. This essay, however, focuses on another approach: new fiduciary obligations that protect end-user privacy and counteract social media companies' bad incentives.

How Do We Pay for the Digital Public Sphere?

How does the political and economic system pay for the digital public sphere in our Second Gilded Age?[1] In large part, it pays for it through digital surveillance and

* Balkin, "Fixing Social Media's Grand Bargin," Hoover Working Group on National Security, Technology, and Law, Aegis Series Paper No. 1814 (October 16, 2018), available at https://www.lawfareblog.com/advanced-persistent-manipulators-and-social-media-nationalism-national-security-world-audiences.

[1] During the First Gilded Age, which ran from the end of Reconstruction to the beginning of the twentieth century, technological innovation created huge fortunes in the hands of a small number of entrepreneurs and produced increasing inequalities of wealth and deep political corruption. Waves of immigration and increasing racial tensions led to the emergence of populist demagogues. American government was increasingly for sale, and many people despaired for the future of American democracy. The corruption and inequality of the First Gilded Age led to the reforms of the Progressive Era and, eventually, the New Deal. For a general history of the period, see H. W. Brands, American Colossus: The Triumph of Capitalism, 1865–1900 (New York: Anchor, 2010).

through finding ever new ways to make money out of personal data.

Twenty-first-century social media like Facebook or YouTube differ from twentieth-century mass media like broadcast radio and television in two important respects. First, they are participatory, many-to-many media. Twentieth-century broadcast media are few-to-many: they publish and broadcast the content of a relatively small number of people to large audiences. In the twentieth century, most people would never get to use these facilities of mass communication to speak themselves. They were largely relegated to the role of audiences.

Twenty-first-century social media, by contrast, are many-to-many: they depend on mass participation as well as mass audiences. They make their money by encouraging enormous numbers of people to spend as much time as possible on their platforms and produce enormous amounts of content, even if that contribution is something as basic as commenting on, liking, or repeating somebody else's contribution. Facebook and Twitter would quickly collapse if people didn't constantly produce fresh content. Search engines, which are key parts of the digital infrastructure, also depend on people creating fresh links and fresh content that they can collect and organize.

Second, twenty-first-century social media like Facebook, YouTube, and Instagram rely on far more advanced and individualized targeted advertising than was available to twentieth-century broadcast media. Television and radio attempted to match advertisers with viewers, but there were limits to how finely grained they could target their audiences. (And newspapers, of course, relied on very broad audiences to sell classified advertisements.)

What makes targeted advertising possible is the collection, analysis, and collation of personal data from end users. Digital communication leaves collectible traces of interactions, choices, and activities. Hence digital companies can collect, analyze, and develop rich dossiers of data about end users. These include not only the information end users voluntarily share with others, but their contacts, friends, time spent on various pages, links visited, even keystrokes. The more that companies know about their end users, the more they know about other people who bear any similarity to them, even if the latter spend less time on the site or are not even clients. In the digital age, we are all constantly informing, not only on ourselves, but on our friends and relatives and, indeed, on everyone else in society.

This is not only true of social media, but of a wide range of digital services. The publisher of a paperback book in the 1960s could tell little about the reading habits of the people who purchased it, while Amazon can tell a great deal about the reading habits of the people who use their Kindle service, down to the length of time spent, the pages covered, the text highlighted and shared, and so on. As the Internet of things connects more and more devices and appliances to digital networks, surveillance spreads to ever more features of daily interaction. In general, the more interactive and the more social the service, the greater the opportunities for data collection, data analysis, and individualized treatment.

Data collection and analysis allow targeted advertising, which allows more efficient advertising campaigns, which allow greater revenues. But data collection and analysis offer another advantage: in theory, they give social media opportunities to structure and curate content for end users

that they will find most engaging and interesting. That is important because advertising revenues depend on the amount of time and attention spent on the site. More engaging content means more time spent and more attention gained.

Social media companies have economic incentives to develop algorithms that will promote content that engages people. That is because companies' central goal is to gain attention share. This leads them to collect ever more data about their end users so that they can tailor content to individual end users to maximize their emotional engagement.[2]

This creates a problem. Often what engages people the most is material that produces strong emotional reactions—even if it is polarizing, false, or demagogic. Companies have economic incentives to expose people to this material. And unscrupulous actors, both domestic and foreign, have learned to take advantage of this feature of social media. As a result, the same business model that allows companies to maximize advertising revenues also makes them conduits and amplifiers for propaganda, conspiracy theories, and fake news.[3]

The Digital Grand Bargain and its Problems

Social media business models are a special case of the grand bargain that has made the digital public sphere possible in our Second Gilded Age. The bargain goes something like this: We will give you miraculous abilities. We will give you social media that allow you to connect with anyone, anywhere, anytime, in a fraction of a second. We will give you search engines that find anything you are looking for instantaneously. We will give you new forms of entertainment that are absorbing, engaging, outrageous, and amusing. We will give you ever more ways to measure yourself and express yourself to others.

We will give all of this to you, for free! And in return, you will let us surveil you. You will let us collect and analyze your habits, your locations, your links, your contacts with your friends, your mouse clicks, your keystrokes, anything we can measure. We will gladly take all of that and study it, and draw inferences from it, and monetize it, so that we can give you all these miraculous things. And we will use that data to perform experiments on you to figure out how to keep you even more focused on our sites and our products, so

[2] See Zeynep Tufecki, "Facebook's Surveillance Machine," New York Times, March 19, 2018, accessed September 27, 2018, https:// www .nytimes . com / 2018/ 03 / 19 / opinion/ facebook - cambridge- analytica .html. ("Facebook makes money, in other words, by profiling us and then selling our attention to advertisers, political actors and others. These are Facebook's true customers, whom it works hard to please.") These business models and the incentives they create are examples of what Shoshana Zuboff calls "surveillance capitalism." Shoshana Zuboff, "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization," Journal of Information Technology 30 (April 2015): 75 (defining "surveillance capitalism" as a "new logic of accumulation" and a "new form of information capitalism [that] aims to predict and modify human behavior as a means to produce revenue and market control").

[3] See, e.g., Paul Lewis, " 'Fiction Is Outperforming Reality': How YouTube's Algorithm Distorts Truth," Guardian, February 2, 2018, accessed September 27, 2018, https://www . theguardian . com / technology/ 201 8/ feb / 02 / how - youtubes - algorithm - distorts - truth (explaining how YouTube's algorithm to engage viewers promotes conspiracy theories).

that you can produce even more data for us, which we can monetize.

This is the grand bargain of the Second Gilded Age. This is twenty-first-century data capitalism. And this is also the irony of the digital age: an era that promised unbounded opportunities for freedom of expression is also an era of increasing digital surveillance and control. The same technological advances allow both results. The infrastructure of digital free expression is also the infrastructure of digital surveillance.

What is objectionable about this grand bargain? The most obvious objection is that we must surrender individual privacy in order to speak. We must make ever more detailed portraits of our lives available to social media companies and their business partners. Beyond this, however, lies a deeper concern: the potential for abuse of power. In particular, the digital grand bargain creates an increased danger of manipulation—both by social media companies and by those who use social media companies—that is of a different degree and kind than that which existed in the pre-digital era. By "manipulation" I mean techniques of persuasion and influence that (1) prey on another person's emotional vulnerabilities and lack of knowledge (2) to benefit oneself or one's allies and (3) reduce the welfare of the other person.[4] (Successful manipulation can also have ripple effects on third parties, such as family members and friends, or even fellow citizens.)

The problem with the current business models for social media companies such as Facebook, Twitter, and YouTube is that they give companies perverse incentives to manipulate end users—or to allow third parties to manipulate end users—if this might increase advertising revenues, profits, or both.

Manipulation is not a new problem. In the past, businesses have often appealed to people's emotions, desires, and weaknesses and have taken advantage of their relative ignorance. So have demagogues and political con artists. But the digital world of social media amplifies the opportunities for manipulation, both by social media companies and by those who use social media to reach end users.

The digital age exacerbates the twentieth-century problem of manipulation in several important respects. First, there is the issue of individual targeting. Twentieth-century influence campaigns were usually aimed at broad groups of individuals, with effects that were often hit-or-miss. With digital technologies it is now possible to tailor influence campaigns to individuals or to very small groups. Instead of appealing to the general emotional vulnerabilities of the public or the vulnerabilities of large demographic groups, digital companies can increasingly target the specific

[4] This definition of manipulation focuses on leveraging another's lack of knowledge and emotional vulnerability to benefit oneself at the expense of another's welfare. This is not the only way to define the concept. See, e.g., Cass R. Sunstein, The Ethics of Influence: Government in the Age of Behavioral Science (New York: Cambridge University Press, 2016), 82 (a technique of influence is "manipulative to the extent that it does not sufficiently engage or appeal to [a person's] capacity for reflection and deliberation"). That definition, however, raises the problem of how to distinguish manipulation from a wide range of ordinary techniques of marketing. A third approach would focus on real or objective interests: manipulation is persuasion that leverages lack of knowledge and emotional vulnerability to cause people to act against their real interests, however those are defined. This approach raises the question of how we know what people's real or objective interests are.

vulnerabilities and emotional hot buttons of individuals who may not be aware of precisely how they have been singled out.

Second, there are differences in scale, speed, and interactivity. Digital technologies allow individualized messages to be targeted to vast numbers of people simultaneously, something that was not possible with twentieth-century media. Moreover, end users' responses can be collected instantaneously, allowing companies to continually fine-tune their approaches, speeding up the Darwinian evolution of the most successful influence strategies. On top of this, digital companies now have the ability to perform interactive social science experiments on us to perfect their abilities to leverage and control our emotions. Facebook, for example, performed experiments to manipulate the emotional moods of 700,000 end users without their knowledge.[5] It has also experimented with ways of encouraging people to vote. But such techniques might also be used to discourage people from voting.[6] Moreover, these experiments can affect the behavior of not only end users but also those they come into contact with.[7]

Third, there is the problem of addiction. The more digital companies know about people's emotional vulnerabilities and predispositions, the more easily they can structure individual end-user experience to addict end users to the site.[8] Social media leverage the data they collect about end users to offer periodic stimulation that keeps users connected and constantly checking and responding to social media. Media have always been designed to draw people's attention, but the digital experience can be especially immersive and pervasive, and thus a more powerful lure than a billboard or magazine advertisement. Here once again, the digital age far outstrips the powers of twentieth-century media.

One might object that, despite all this, the digital grand bargain remains freely chosen and welfare-enhancing. End users are free to use or not to use social media, and thus they are free to decide whether they will subject themselves to experimentation and emotional manipulation. If the free service is sufficiently valuable to them, they will accept the bargain. But this overlooks three

[5] "Facebook Admits Failings over Emotion Manipulation Study," BBC News, October 3, 2014, accessed September 27, 2018, https:// www . bbc . com / news / technology - 29475019 ("the company was widely criticised for manipulating material from people's personal lives in order to play with user emotions or make them sad").

[6] Jonathan Zittrain, "Engineering an Election," Harvard Law Review Forum 127 (June 20, 2014): 335–36 (noting that experiment caused an additional 340,000 votes to be cast).

[7] Ibid., 336 (describing the "ripple effects" of experiments).

[8] See Mike Allen, "Sean Parker Unloads on Facebook: 'God Only Knows What It's Doing to Our Children's Brains,' " Axios, November 9, 2017, accessed September 27, 2018, https:// www . axios . com / sean - parker - unloads- on - facebook - god - only - knows - what - its - doing - to - our - childrens - brains - 1513306792 -f855e7b 4 - 4e99 - 4d60 - 8d51 - 2775559c2671 . html (quoting statement by former president of Facebook that social media applications are designed to "exploit a vulnerability in human psychology" using psychological methods to "consume as much of your time and conscious attention as possible" and keep users locked into the site); Paul Lewis, " 'Our Minds Can Be Hijacked': The Tech Insiders who Fear a Smartphone Dystopia," Guardian, October 6, 2017, accessed September 27, 2018, https:// www . theguardian. com / technology / 2017 / oct / 05 / smartphone - addiction - silicon- valley - dystopia (interviewing former employees at Google and Facebook who report that technologies are designed to addict users and monopolize their attention).

important features of the emerging system of digital surveillance that make the assumption of a mutually beneficial arm's-length bargain highly implausible.

First, we cannot assume that transactions benefit both parties when there is extreme asymmetry of knowledge, in which one party's behaviors, beliefs, and activities are known to the other party while the other party is essentially a black box.

Second, individuals suffer from privacy myopia, a characteristic feature of digital interactions.[9] Individuals constantly generate a broad range of information about themselves through digital interactions, much of which (for example location, social connections, timing of responses, and rate of keystrokes) they may be only dimly aware. Individuals have no way of valuing or assessing the risks produced by the collection of particular kinds of information about them or how that information might be employed in the future. That is because the value of such information is cumulative and connective. Information that seems entirely irrelevant or innocuous can, in conjunction with other information, yield surprisingly powerful insights about individual values, behavior, desires, weaknesses, and predispositions. Because individuals cannot assess the value of what they are giving up, one cannot assume that their decisions enhance their welfare. In this environment, the idea of relying on informed consumer choice to discipline social media companies is a fantasy.

Third, as noted above, information gathered from end users has significant external effects on third parties who are not parties to the bargain. As digital companies know more about you, they also can learn more about other people who are similar to you or connected to you in some respect.[10] In the digital age, we do not simply inform on ourselves; we inform on other people as well. And when a social media company experiments with social moods or engineers an election, it affects not only its end users but many other people as well.

For all these reasons, it is fatuous to compare the digital grand bargain to a mutually beneficial arm's-length economic transaction. If we can pay for digital

[9] See, e.g., Ryan Calo, "The Boundaries of Privacy Harm," Indiana Law Journal 86, no. 3 (October 4, 2011): 1131, 1149 ("Many consumers have little idea how much of their information they are giving up or how it will be used"); A. Michael Froomkin, "The Death of Privacy?" Stanford Law Review 52 (2000): 1461, 1502 ("Consumers suffer from privacy myopia: they will sell their data too often and too cheaply"); Daniel J. Solove, "Privacy and Power: Computer Databases and Metaphors for Information Privacy," Stanford Law Review 53 (2001): 1393, 1452 ("It is difficult for the individual to adequately value specific pieces of personal information").

[10] See Tufekci, "Facebook's Surveillance Machine," explaining that Facebook collects "shadow profiles" on nonusers: "even if you are not on Facebook, the company may well have compiled a profile of you, inferred from data provided by your friends or from other data. This is an involuntary dossier from which you cannot opt out in the United States." Social media users may unwittingly imperil each other's privacy. The Cambridge Analytica scandal revealed that when Facebook users logged in to a third-party app using their Facebook credentials, they shared the social graphs of all of their Facebook friends without the latter's consent. See Alexandra Samuel, "The Shady Data-Gathering Tactics Used by Cambridge Analytica Were an Open Secret to Online Marketers. I Know, Because I Was One," The Verge, March 25, 2018, accessed September 27, 2018, https:// www . theverge .co m/ 2018 / 3 / 25 / 17161726 / facebook - cambridge - analytica- data - online - marketers. ("The tactic of collecting friend data, which has been featured prominently in the Cambridge Analytica coverage, was a well-known way of turning a handful of app users into a goldmine.")

freedom of expression while reducing the dangers of digital manipulation, it is worth exploring alternatives.

Public Options

Proposals for reform of social media abound these days. One kind of proposal argues that we should counter the power of social media and search engines by treating them as state actors. Courts should apply standard First Amendment doctrine to them and treat them as public forums, which require complete content and viewpoint neutrality. If social media cannot choose what we see, they cannot manipulate us.

This solution fails to grapple with the central problems of the grand bargain. First, treating social media as public forums would only affect the ability of social media themselves to manipulate end users. It would do nothing to prevent third parties from using social media to manipulate end users, stoke hatred, fear, and prejudice, or spread fake news. And because social media would be required to serve as neutral public forums, they could do little to stop this. Second, even if social media do not curate feeds, they still collect end-user data. That end-user data, in turn, can be harvested and sold to third parties, who can use it on the site or elsewhere. (That is why, for example, requiring social media companies to offer a tiered service in which people pay not to receive commercial advertisements does not really deal with the underlying problem of surveillance, data collection, and manipulation.)

Perhaps equally important, the proposal is unworkable. Social media—and search engines— must make all sorts of editorial and curational judgments that the First Amendment forbids government entities to make.

For example, social media sites might want to require that end users use their real names or easily identifiable pseudonyms in order to limit trolling and abuse. They might decide to ban hate speech or dehumanizing speech, especially if they operate around the world. They might choose to ban graphic violence, nudity, or pornography. They might choose to ban advocacy of violence or illegal conduct, or the promotion of suicide. They might decide to ban certain types of harassment or incitement even if that harassment or incitement does not immediately lead to a breach of the peace.[11] They might ban certain forms of advertising. All of these regulations would be unconstitutional if a government imposed them in a public forum. More generally, we should accept that social media will have to make sometimes quite complicated decisions to discipline abusive trolls, maintain civility norms, demote the ranking of postings by conspiracy theorists and hate mongers, and, in cases of serial abuse, terminate accounts. Many of these policies would be unconstitutional if we applied the same standards to social media that the First Amendment applies to municipal streets and parks.

At a more basic level, it is impossible to manage a search engine or a social media site without curation, which involves a wide range of content-based judgments about what content to promote and what to

[11] For examples of what social media sites regulate, see Facebook, Community Standards, https:// www . facebook .com/ communitystandards; and Twitter, The Twitter Rules, https:// help . twitter . com/ en / rules - and - policies/ twitter - rules (both accessed September 27, 2018).

demote.[12] It is also impractical and self-defeating to manage a social media site without moderation, which requires the imposition of a wide range of civility rules that the First Amendment forbids governments from imposing in public discourse. Moreover, creating individualized social media feeds and search engine results inevitably requires content-based judgments. As described below, social media and search engines sometimes make bad decisions about these matters, but the solution is not to impose a set of doctrinal rules crafted for municipal streets and parks.

A second, related proposal argues that we should treat social media sites and search engines as public utilities because they perform what are clearly public functions. But public utility regulation—for example, of water and power utilities—generally focuses on two issues: access to essential services and fair pricing. Neither of these is particularly relevant. Social media and search engines want everyone to participate and they offer their services for free. If the goal of the public utility metaphor is to prevent content discrimination, it faces the same problems as treating digital media as state actors.

A third and quite different approach is public provisioning. Instead of treating existing private companies as arms of the state, governments could provide their own public options: government-run social media and search engines. For reasons stated above, these would not really work very well if they had to be organized as public forums and moderation was forbidden. There are potential solutions, however. The government could provide only a basic telecommunications system for social media messages and then allow various groups and businesses to create their own private moderation systems on top, from which individuals could choose. The government might also create an open system in which third parties could develop applications that allow people to design their own personalized feeds.

A government-provided search engine that is as efficient and effective as Google's is a somewhat harder lift and the cost of public provisioning for social media and search engines might be prohibitive. But public provisioning poses a far larger problem: state surveillance. Instead of Facebook and Google scooping up your personal data, the government would. The Fourth Amendment might not prohibit this under existing doctrines, because people willingly give the information to the public entity. Therefore any public provisioning system would have to be accompanied by very strict self-imposed restrictions on collection, analysis, and use. I am deeply skeptical that law enforcement and national security officials would willingly forgo access to all of this information.

Professional and Public-regarding Norms

We should not treat social media companies and search engines as state actors subject to the First Amendment. Yet we can still criticize them for arbitrariness and censorship. How is that possible if, as I have just explained, these companies must engage in content- and viewpoint-based judgments to do their jobs?

We can criticize social media companies in three ways, none of which requires us to treat them as state actors.

[12] Tarleton Gillespie, Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media (New Haven, CT: Yale University Press, 2018).

First, we can criticize them for being opaque and non-transparent and for denying basic norms of fair process. This happens when social media do not state their criteria for governance clearly in advance and do not offer reasoned explanations for their decisions.

Second, we can criticize them for being arbitrary—for not living up to their own community guidelines and terms of service. They should apply their own rules without fear or favor to the rich and to the poor, the high and low alike. Twitter and Facebook, to name two examples, have often been lax with violations of their terms of service by famous or well-known people and strict with violations by people who are not famous or well known.[13] This allows the more powerful and famous to abuse the less powerful with impunity and it creates blind spots in enforcement.

Third, and perhaps more important, we can criticize social media companies for failing to live up to norms of professionalism and expertise—that is, for failing to live up to the norms of the kind of entity they purport to be.

Here is an analogy. People criticize major newspapers and media outlets all the time. They criticize them for biased coverage, they criticize them for giving a platform to people who make stupid or evil arguments, and they criticize them for failing to educate the public about the issues of the day.

In most cases, people understand that these criticisms aren't supposed to lead to government regulation of newspapers and mass media. People understand that these companies have a First Amendment right to exercise editorial discretion as they see fit, even if they exercise it badly. Nevertheless, they hold these companies to a higher standard than ordinary individuals expressing their opinions. The public rightly assumes that media companies should live up to certain professional standards that are both public-regarding and connected to democratic life. These include, among other things, providing the public with important information necessary to self-government, striving to cover the news accurately and fairly, engaging in professional fact-checking, adhering to professional standards of journalistic ethics, and so on.

Many media organizations fail to live up to these standards, often spectacularly so. And some media organizations have essentially given up on professionalism, fairness, and accuracy. But people generally understand that this is a valid reason to criticize them, not to exculpate them. Media companies hold themselves out as adhering to professional and public-regarding norms. Therefore people in a democracy feel that they have a right to criticize them when, in their estimation, media fail to live up to those norms. Perhaps equally important, because the norms are public-regarding, citizens in a democracy feel that they have a right to debate what those professional norms should be, whether or not the media companies assume them or live up to them.

Social media companies and search engine companies are not newspapers. Even so, they are more than just run-of-the-mill companies. They do more than just serve ads or sell widgets. They perform a public service—three connected services, in fact. First, they facilitate public participation in art, politics, and culture. Second, they

[13] See Kate Klonick, "The New Governors: The People, Rules, and Processes Governing Online Speech," Harvard Law Review 131 (April 10, 2018): 1598, 1654–55 (2018) (noting that social media companies may disproportionately favor people with power over other end users).

organize public conversation so that people can easily find and communicate with each other. Third, they curate public opinion through individualized results and feeds and through enforcing terms-of-service obligations and community guidelines.

These digital companies are the twenty-first-century successors of twentieth-century mass media companies, even though their functions are somewhat different. The public, not surprisingly, has come to view them as having a public-oriented mission.

In fact, these companies encourage this understanding through the ways they talk about themselves. The Twitter Rules, for example, begin with the statement, "We believe that everyone should have the power to create and share ideas and information instantly, without barriers. In order to protect the experience and safety of people who use Twitter, there are some limitations on the type of content and behavior that we allow."[14] This is a statement of public-regarding, professional norms for facilitating public participation, organizing public discussion, and curating public opinion. Facebook and YouTube have made similar statements of purpose and justifications for their community guidelines, although their policies differ in some respects.[15]

Whether they imagined it or not at the outset, these companies have taken on a public function. People may therefore criticize them—and should criticize them—if they feel that these companies are acting contrary to appropriate professional norms.

Moreover, because these companies have taken on these three tasks—facilitating public participation, organizing public conversation, and curating public opinion—they may also impose basic civility norms against abuse, threats, and harassment. They may also ban hate speech or speech that denigrates people if they think that this kind of speech will undermine the public-regarding purposes of the site. Social media companies may do this even if the First Amendment would prevent the federal government from imposing the same civility norms on a government-operated social media site.

But if social media companies decide to govern their sites through imposing civility norms and regulating harassment and abuse, they should abide by the two other basic norms stated above. First, they should be transparent about what they are doing and why they are doing it. Second, they should not be arbitrary in their governance.

Social media companies have been only fitfully successful at meeting these obligations. Understood charitably, we might say that they are at the very beginning of a long process of learning how to be responsible professionals. They have been wildly successful as technology companies, but professionalism is more

[14] Twitter, the Twitter Rules.

[15] Facebook, Community Standards, "We recognize how important it is for Facebook to be a place where people feel empowered to communicate, and we take our role in keeping abuse off our service seriously. That's why we have developed a set of Community Standards that outline what is and is not allowed on Facebook. . . . The goal of our Community Standards is to encourage expression and create a safe environment," YouTube, Policies and Safety, accessed September 27, 2018, https:// www. youtub e .co m/ y t/ abou t/ policies/ #communit y- guidelines , "When you use YouTube, you join a community of people from all over the world. . . . Following the guidelines below helps to keep YouTube fun and enjoyable for everyone."

than technological expertise. Professional judgments may require the application of norms that do not scale well. Sometimes applying these norms will require judgment and individualized assessment as well as algorithmic sorting and bright-line rules. Doing this costs more in human resources and attention than purely technological solutions. To the extent that this is the case, social media companies should absorb the extra costs of being professionals and living up to professional norms. Although their efforts have been halting and often inadequate, social media companies are slowly beginning that arduous process. In the meantime, civil society can play an important role by continuing to criticize social media companies and by encouraging them to live up to their public responsibilities.

Reforming Social Media

I have already said that we should not use the law to force these companies to behave as public-regarding professionals any more than we can force major newspapers to adhere to proper journalistic standards. Does this mean that law has no role to play? No. The law may encourage these public-regarding norms in certain limited ways consistent with the First Amendment.

Instead of directly aiming at the editorial policies of social media companies, reform proposals should focus instead on the grand bargain that has turned the infrastructure of digital free expression into the infrastructure of digital surveillance and control. Social media companies will continue to cause a host of social problems as long as their business models cause them not to care about these problems.

There are two central ways to change their behavior. The first is to reshape the organization of social media companies. This is the task of antitrust and pro-competition law, which have grown moribund in the Second Gilded Age and need a serious rethinking.

Social media companies' perverse incentives derive from their business models—selling end users' information to advertisers and manipulating and addicting end users so that they spend more time on social media and are thus more accessible to advertisers. Because a small number of social media dominate end users' attention, they also have a stranglehold over digital advertising. People who wish to advertise online must operate primarily through Facebook's and Google's advertising networks. This reduces revenues for many news and media sites that are crucial to the health and vibrancy of the digital public sphere.

Increased enforcement of existing antitrust laws and a series of new pro-competition policies might have two salutary effects. First, these reforms might restructure how digital advertising operates, ameliorating the current bottleneck and freeing up revenues for a wider range of media companies. Second, reform of competition policy and stricter antitrust enforcement might break up the largest companies into smaller companies that can compete with each other or create a space for new competitors to emerge. (Facebook and Google have often bought up potential competitors before they could grow large enough to threaten them.)

More social media companies mean more platforms for innovation and more different software features and affordances. More companies might also make it more difficult for foreign hackers to disrupt the digital public sphere. All other things being equal, it may be harder to hack twelve

Facebooks than only one.[16] Finally, more different kinds of companies might also provide more models for social spaces and communities and a wider variety of speech policies.

This last point is especially important. I have just argued that social media companies must be allowed to enforce civility norms and regulate or even ban a wide range of speech that state actors may not touch. But modern democracies increasingly rely on social media to perform the public functions of organizing public opinion and facilitating public discussion. Therefore it is very important to ensure that there are many social media applications and businesses in order to prevent a small number of powerful for-profit companies from dominating how public opinion is organized and governed.

Moreover, social media companies often enforce their terms of service imperfectly and arbitrarily and they may make many questionable judgments. Some, like Facebook, attempt to impose the same standards around the world.[17] Finally, civil society organizations, mass media, politicians, and governments have and will put increasing pressure on social media to ban speech that they do not like and expel speakers who offend them. All of them, in various ways, will try to coax social media into serving their political or ideological agendas. These are all reasons for using pro-competition laws to ensure a healthy number of competing firms organizing public discourse. Precisely because people will demand that huge multinational corporations ban speech they do not like, it is important to have many Facebooks, not just one. If we expect social media sites to enforce civility norms, we also need multiple social media sites serving different values and different publics.

Information Fiduciaries

A second approach to reform is to make social media companies internalize the costs they impose on society through surveillance, addiction, and manipulation by giving them new social responsibilities. The short-term goal is to counteract the most egregious examples of bad behavior. The long-term goal is to create legal incentives for social media companies to develop professional cultures and public-oriented norms for organizing and curating public discussion. To do this, I propose reaching back to some very old ideas in the law that governs the professions: namely, the idea of fiduciary obligation.

We should treat social media companies—and many other digital media companies as well— as information fiduciaries toward their clients and end users.[18] As information fiduciaries, digital

[16] Sally Hubbard, "Fake News is a Real Antitrust Problem," CPI Antitrust Chronicle, December 2017: 5, accessed September 27, 2018, https:// www . competitionpolicyinternational . com / wp-content / uploads /201 7 / 12 / CPI -Hubbard. pdf .

[17] Klonick, "New Governors," 1642, describing Facebook's goal of applying its norms worldwide and the resulting compromises; Julia Angwin and Hannes Grassegger, "Facebook's Secret Censorship Rules Protect White Men from Hate Speech but Not Black Children," ProPublica, June 28, 2017, accessed September 27, 2018, https:// www . propublica. org /article/ facebook - hate - speech - censorship - internal - documents - algorithms (describing Facebook's attempts to enforce its hate speech rules worldwide and the arbitrariness of its categories).

[18] See Jack M. Balkin, "Information Fiduciaries and the First Amendment," UC Davis Law Review 49, no. 4 (April 2016): 1183; Jack M. Balkin, "The Three Laws of Robotics in the Age of Big Data," Ohio State Law Journal 78 (2017): 1217.

companies should have duties of care, confidentiality, and loyalty toward the people whose data they collect, store, and use. This reform is a natural outgrowth of the grand bargain that has enabled free expression in the digital age.

Because of digital companies' increasing capacities for surveillance and control, they must take on new legal responsibilities. Put simply, digital companies know a lot about us, and they can use that knowledge in many ways—but we don't know a lot about them. Moreover, people increasingly depend on a wide range of digital services that observe them and collect data about them. That makes people increasingly vulnerable to these companies. Because the companies' operations are not transparent, people have to trust that these services will not betray them or manipulate them for their own ends. Digital companies that create and maintain this dependence and vulnerability should be considered information fiduciaries toward their end users.

There is plenty of precedent for this idea. For centuries, the law has recognized that certain people hold power over others who are vulnerable to them, dependent on them, and have to trust them. It created the idea of fiduciary obligations for just these situations.[19] For example, the law has long maintained that the clients or patients of doctors and lawyers are in special relationships of dependence and vulnerability. We need to trust these professionals with sensitive personal information about ourselves, but the people we trust could use this same information to harm us and enrich themselves in many different ways. Therefore the law treats professionals like doctors, lawyers, accountants, and estate managers as fiduciaries. Fiduciary relationships require good faith and loyalty toward people whom the relationships place in special positions of vulnerability. Accordingly, fiduciaries have special duties of care, confidentiality, and loyalty toward their clients and beneficiaries.

Because social media companies collect so much data about their end users, use that data to predict and control what end users will do, and match them with third parties who may take advantage of end users, they are among the most important examples of the new information fiduciaries of the digital age. We should apply these traditional obligations to the changed conditions of a new technological era.

Facebook is not your doctor or lawyer. YouTube is not your accountant or estate manager. We should be careful to tailor the fiduciary obligations to the nature of the business and to the reasonable expectations of consumers. That means that social media companies' fiduciary duties will be more limited.

Social media companies and search engines provide free services in exchange for the right to collect and analyze personal data and serve targeted ads. This by itself does not violate fiduciary obligations. Nevertheless, it creates a perpetual conflict of interest between end users and social media companies. Companies will always be tempted to use the data they collect in ways that increase their profits to their end users' disadvantage. Unless we are to ban targeted advertising altogether (which I would oppose and which raises serious First Amendment problems) the goal should be to ameliorate or forestall conflicts of interest

[19] See generally Tamar Frankel, Fiduciary Law (New York: Oxford University Press, 2011).

and impose duties of good faith and non-manipulation. That means that the law should limit how social media companies can make money off their end users, just as the law limits how other fiduciaries can make money off their clients and beneficiaries.

As information fiduciaries, social media companies have three major duties: duties of care, duties of confidentiality, and duties of loyalty. The duties of care and confidentiality require fiduciaries to secure customer data and not disclose it to anyone who does not agree to assume similar fiduciary obligations. In other words, fiduciary obligations must run with the data. The duty of loyalty means that fiduciaries must not seek to advantage themselves at their end users' expense and they must work to avoid creating conflicts of interest that will tempt them to do so. At base, the obligations of loyalty mean that digital fiduciaries may not act like con artists. They may not induce trust on the part of their user base and then turn around and betray that trust in order to benefit themselves.

To see what these obligations would mean in practice, we can use the Cambridge Analytica scandal that propelled the issue of social media regulation to public attention in the spring of 2018. Although the facts are complicated, they essentially involved Facebook's decision to allow third parties to access its end users' data.[20] Facebook allowed researchers to do this for free and took a cut of the profits for business entities. This allowed it to leverage its central resource—consumer data—to increase profits.

Aleksandr Kogan, a data scientist, used a personality quiz to gain access to Facebook's end-user data. He thereby obtained not only the data of the 300,000 people who logged in using their Facebook credentials, but also all of their Facebook friends, an estimated 87 million people.[21] In fact, Kogan was actually working for Cambridge Analytica, a for-profit political consulting company. Cambridge Analytica used the end-user data to produce psychological profiles that, in turn, it would use to target political advertisements to unsuspecting Facebook users. In fact, these practices were only the tip of a far larger iceberg. Facebook made a series of unwise decisions to allow a range of business partners access to its end users' social

[20] See Carole Cadwalladr and Emma Graham-Harrison, "How Cambridge Analytica Turned Facebook 'Likes' into a Lucrative Political Tool," Guardian, March 17, 2018, accessed September 27, 2018, https:// www.theguardian.com / technology / 2018 / mar /17 / facebook -cambridge - analytica- kogan- data- algorithm; Carole Cadwalladr and Emma Graham-Harrison, "Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach," Guardian, March. 17, 2018, accessed September 27, 2018, https:// www . theguardian. com /news / 2018 /mar/1 7/ cambridg e- analytic a - facebook - influence-u s- election; Paul Lewis, " 'Utterly Horrifying': Ex-Facebook Insider Says Covert Data Harvesting Was Routine," Guardian, March 20, 2018, https://ww w .theguardian.com / news / 2018 / mar / 20/ facebook- data -cambridge- analytic a -sandy- parakilas.

[21] See Michael Riley, Sarah Frier, and Stephanie Baker, "Understanding the Facebook-Cambridge Analytica Story: QuickTake," Washington Post, April 9, 2018, accessed September 27, 2018, https://www. washingtonpost. com /business/ understanding -the - facebook - cambridge - analytica - story- quicktake /201 8/ 0 4/ 09 /0 f18d91c -3 c1c- 11 e8 - 955b - 7d2e19b79966_ story. html (estimating that 300,000 people participated and that 87 million users had their data harvested).

graphs and thus make them vulnerable to various kinds of manipulation.[22]

As an information fiduciary, Facebook violated all three of its duties of care, confidentiality, and loyalty. It did not take sufficient care to vet its academic and business partners. It did not ensure that it only gave access to data to entities that would promise to maintain the same duties of care, confidentiality, and loyalty as Facebook. It did not take sufficient steps to audit and oversee the operations of these third parties to ensure that they did not violate the interests of its end users. It allowed third parties to manipulate its end users for profit. And when it discovered what had happened, many years later, it did not take sufficient steps to claw back its end users' data and protect them from further breaches of confidentiality and misuse.

Fiduciary obligations matter most in situations in which social media companies have powerful market incentives not to protect their end users: for example, when social media companies give access to data to third-party companies without adequate safeguards to prevent these third parties from manipulating end users. Fiduciary obligations also matter when social media companies perform social science experiments on their end-user base. Social media companies are not part of universities and therefore are not bound by human-subjects research obligations. As information fiduciaries, however, they would have legal duties not to create an unreasonable risk of harm to their end users or to the public for their own advantage. They would have duties, just as university scientists do, to minimize harm and prevent overreaching and manipulation by their employees and contractors.

Finally, if social media companies are information fiduciaries, they should also have a duty not to use end-user data to addict end users and psychologically manipulate them. Social media companies engage in manipulation when end users must provide information in order to use the service and when companies use this information to induce end-user decision making that benefits the company at the expense of the end user and causes harm to the end user. Because this creates a conflict of interest between the company and its end users, it violates the duty of loyalty.

It may be useful to compare the fiduciary approach with the privacy obligations of the European Union's General Data Protection Regulation (GDPR). There is considerable overlap between the two approaches. But the most important difference is that the GDPR relies heavily on securing privacy by obtaining end-user consent to individual transactions. In many respects, it is still based on a contractual model of privacy protection. Contractual models will prove insufficient if end users are unable to assess the cumulative risk of granting permission and therefore must depend on the good will of data processors. The fiduciary approach to obligation does not turn on consent to particular transactions, nor is it bounded by the precise terms of a company's written privacy policy or terms of service, which are easy for companies to modify. Rather, the fiduciary approach holds digital fiduciaries to obligations of good faith and non-

[22] Lewis, "Covert Data Harvesting Was Routine" (quoting a former Facebook employee who explained that under the company's policies, "a majority of Facebook users" could have had their data harvested by app developers without their knowledge).

manipulation regardless of what their privacy policies say.

The fiduciary approach is also consistent with the First Amendment. That is because it aims at regulating the relationships of vulnerability and trust between information fiduciaries and those who must trust them.[23]

The First Amendment treats information gained in the course of a fiduciary relationship differently from other kinds of information. Tell a secret to a person in the street and he or she can publish it tomorrow and even use it against your interests. But when you reveal information to a fiduciary—a doctor, nurse, or lawyer—he or she has to keep it confidential and cannot use it against you. Information gained in the course of a fiduciary relationship— and that includes the information that social media companies collect about us—is not part of the public discourse that receives standard First Amendment protection. Instead, the First Amendment allows governments to regulate fiduciaries' collection, collation, use, and distribution of personal information in order to prevent overreaching and to preserve trust and confidentiality.[24] The same principle should apply to the new information fiduciaries of the digital age.

There may be close cases in which we cannot be sure whether a company really is acting as an information fiduciary. To deal with these situations, Jonathan Zittrain and I have proposed that Congress offer digital companies a different grand bargain to protect end users' privacy.[25] It would create a safe harbor provision for companies that agree to assume fiduciary obligations. The federal government would preempt state regulation if digital media companies accept the obligations of information fiduciaries toward their end users. Offering this exchange does not violate the First Amendment.

For the most part, the fiduciary approach leaves social media companies free to decide how they want to curate and organize public discussion, focusing instead on protecting privacy and preventing incentives for betrayal and manipulation. It affects companies' curation and organization of public discourse only to the extent that companies violate their duties of care, confidentiality, and loyalty.

The fiduciary approach has many advantages. It is not tied to any particular technology. It can adapt to technological change. It can be implemented at the state or the federal level, and by judges, legislatures, or administrative agencies.

The fiduciary approach also meshes well with other forms of consumer protection, and it does not exclude other reforms, like GDPR-style privacy regulation. In particular, it does not get in the way of new pro-competition rules or increased antitrust enforcement as described above. That is because it does not turn on the size of an organization (although Congress might choose to regulate only the larger sites in order to encourage innovation and avoid barriers to entry). It also does not turn on the presence or absence of monopoly power. It applies whether we

[23] On the First Amendment issues, see Balkin, "Information Fiduciaries and the First Amendment," 6.

[24] Ibid.

[25] Jack M. Balkin and Jonathan Zittrain, "A Grand Bargain to Make Tech Companies Trustworthy," Atlantic, October 3, 2016, accessed September 27, 2018, https:// www.theatlantic.com / technology / archive/ 2016 / 10 / information -fiduciary/ 502346.

have twelve Facebooks or only one. Indeed, even if we had a wide range of social media companies, all harvesting, analyzing, and using end-user data, there would still be a need for fiduciary obligations to prevent overreaching.

The fiduciary approach pays attention to deeper causes. It directs its attention to the political economy of digital media. It focuses on repairing the grand bargain that pays for the digital public sphere in the Second Gilded Age.

INTERNET, BIG DATA & ALGORITHMS: GATEWAY TO A NEW FUTURE OR A THREAT TO PRIVACY AND FREEDOM

The Aspen Institute Congressional Program
May 10-13, 2019
Cambridge, Massachusetts

CONFERENCE AGENDA

FRIDAY, MAY 10

Pre-Dinner Remarks

WELCOME TO MIT

Founded in 1861, Massachusetts Institute of Technology is one of America's premier institutions of higher education. With 7,000 graduate students and 5,000 undergrads, it is poised to make a significant mark in the fields of artificial intelligence (AI) and advancements of the digital age with its new $1 billion commitment to a College of Computing, set to open in September. The new College, with 50 new faculty positions, will work across MIT's existing five schools as part of a campus-wide effort to integrate computing and AI more deeply into the curriculum. MIT President Reif will welcome the group with this appropriate backdrop of MIT as the venue for our policy discussions.

***L. Rafael Reif**, President,*
Massachusetts Institute of Technology

Pre-Dinner Remarks
Remarks by **Doug Beck**, Vice President for the Americas and Northeast Asia, Apple, Inc.

Working Dinner

Seating is arranged to expose participants to a diverse range of views and provide the opportunity for a meaningful exchange of ideas. Scholars and lawmakers are rotated daily. Discussion will focus on the opportunities, challenges, and potential solutions regarding privacy and the Internet.

SATURDAY, MAY 11

Roundtable Discussion

THE BENEFITS AND HAZARDS OF ARTIFICIAL INTELLIGENCE ON TRANSPORTATION, HEALTH CARE, NATIONAL SECURITY, MANUFACTURING & THE WORKFORCE

Artificial Intelligence has the potential to have significant impact in numerous sectors of society. This session will survey the landscape of what machine learning can have for changes ahead brought about by utilization and expansion of this technology in wider and wider dimensions of everyday life.

***Hal Abelson**, Professor of Computer Science and Engineering, MIT*
***R. David Edelman**, Director, Project on Technology, Economy, and National Security, MIT*

Roundtable Discussion

ARTIFICIAL INTELLIGENCE, ALGORITHMS, FAIRNESS, AND THREATS TO PRIVACY

Today's online society is increasingly shaped by automated decision-making systems using algorithms and artificial intelligence learning models. These models are developed by individuals and companies from a particular subset of our society and may not represent a fully accurate or fair view of the world. Mathematical models that increasingly intersect citizens in their daily activities are developed by human beings and they can reflect hidden or deliberate biases. Machines, rather than humans, are making complex and morally difficult decisions on behalf of programmers, with consequences for free speech and nuanced thought. These same machines may even come to learn more about the individuals than the individuals know themselves. This unregulated new era of "Big Data" has implications for privacy and fairness that may require federal attention.

- How does this use of algorithms and Big Data impact citizens in areas such as hiring practices, job performance ratings, and credit scores, etc?
- Are there built-in inequities that should be taken into account?
- Does government have a role in alerting consumers to threats to their privacy?

***Joy Buolamwini**, Founder, Algorithmic Justice League & PhD student, MIT Media Lab*
***Cathy O'Neil**, Founder, ORCAA*

Roundtable Discussion

THREATS TO DEMOCRACY IN THE DIGITAL AGE

Four subtopics deserve focus: surveillance, election integrity, misinformation and disinformation, and digital manipulation for malevolent purposes. The explosion of public cameras, done for security purposes, has the potential to change the relationship between citizen and state. Nothing is more essential to the protection of democracy than fair and free elections. Yet, as the U.S. becomes more and more digitized and connected, as hackers take aim at our processes, and as foreign entities try to influence our elections, the integrity of the electoral process is jeopardized. The ease with which anyone can now manipulate information and images digitally opens up new realms of vulnerability with unknowable consequences.

- What actions can and should the U.S. Congress take to protect our freedoms and democratic rights with this explosive power of the Digital Age?
- Are citizen's rights infringed by the preponderance of public cameras?
- Will artificial intelligence enable a new era of state surveillance of citizens?
- Should online companies be subject to greater levels of liability, e.g., for defamation? If so, would these be onerous restrictions of a heavy-handed government limiting free speech or legitimate efforts to protect the public from harmful abuse?
- To what degree should governments be involved in monitoring or even regulating the spread of mis- and dis-information on the internet?
- What are the consequences for digitally spreading falsehoods?
- How do the boundaries of responsible free speech fit the Digital Age?
- Is freedom of expression in the digital world at odds with the maintenance of civic discourse?

***Jonathan Zittrain**, Professor of International Law, Harvard Law School*
***Ethan Zuckerman**, Director, Center for Civic Media, MIT*

SUNDAY, MAY 12

Roundtable Discussion

CONSUMER'S CONSENT AND CONTROL OF ONLINE INFORMATION

In our modern world, data is key. But who actually owns the data and when or how one consents to having their data collected are disputable topics. For example, once an individual's data has been harvested and processed, through either voluntarily or involuntarily online interactions, it can be put to use in targeted consumer marketing campaigns, campaign advertisements, and individualized sales pitches. While individuals' comfort with these techniques varies, one thing is certain: marketing will never be the same. The explosive power of artificial intelligence is being harnessed for commercial advantage, which can be either advantageous or disadvantageous to the consumer depending on what perspective is held.

- Does consumer use of social media expose them to the risk of exploitation?
- Is there a federal role to protect consumers from unwanted solicitations?

***Howard Beales**, Professor of Strategic Management and Public Policy, George Washington University*
***Alessandro Acquisti**, Professor of Information and Public Policy, Carnegie Mellon University*

Roundtable Discussion

PROTECTING THE DRIVE FOR INNOVATION WITHIN THE BOUNDARIES OF THE NEED FOR REGULATION

Our economy is increasingly dependent on the Internet. Social media entities are incentivized to increase their user base. The major digital companies spent over $60 million in 2018 in lobbying and consolidation in the digital industry has raised questions about the power of dominant major players. Do the practices of the economies of scale serve consumer interest, or is the potential of market dominance to the detriment of consumer choices and costs?

- What role does the federal government have in restraining the emergence of dominant major players in this industry?
- What can be done to enhance privacy protections?
- Are consumer concerns adequately taken into account by the industry?
- Do citizens have a right to conduct business online without leaving a digital footprint?

***Larry Downes**, Project Director,*
Georgetown Center for Business and Public Policy

Roundtable Discussion

BIG DATA'S END GAME: THE USE AND ABUSE OF CONSUMER DATA

Though not specified directly in the Constitution, privacy has emerged as a basic human right. Many feel that they have lost control over their personal information. They have. Those who collect information about their online users own it, not the customer. Some have called for personal ownership of the information about them. In Europe, there is a "right to be forgotten," which requires online search companies to delete information that a court decides should be forgotten. In the U.S. we have relied on the Federal Trade Commission to protect privacy against unfair practices and state law. But the European Union's General Data Privacy Regulation, and now the state of California, have imposed greater privacy protections for online behavior than previously required. (For example, Google was fined $57 million by French regulators for breaking the GDPR rules.) One solution is to require digital companies to be "information fiduciaries" with a duty of care not to harm users.

- Do citizens have a right to maintain and control publicly available data about themselves?
- Is there a need to delineate legal boundaries on data use to protect privacy?
- What controls should Congress allow users to retain?
- Is it time for a federal privacy law for the online world?

***Jack M. Balkin**, Professor of Constitutional Law,*
Yale University Law School
***Latanya Sweeney**, Professor of Government and Technology, Harvard University*

Working Lunch

EXPLORING PRIVACY IN THE PAST, PRESENT, AND FUTURE

The legal and social boundaries of privacy have changed over time, and are based on different assumptions in different cultures and societies. Concepts about privacy rooted in the Constitution may need updating in this era of widespread digital communications with implications for federal legislators.

***Daniel Weitzner**, Founding Director,*
MIT Internet Policy Research Initiative

MONDAY, MAY 13

All Participants Depart

INTERNET, BIG DATA & ALGORITHMS: GATEWAY TO A NEW FUTURE OR A THREAT TO PRIVACY AND FREEDOM

The Aspen Institute Congressional Program
May 10-13, 2019
Cambridge, Massachusetts
Massachusetts Institute of Technology

CONFERENCE PARTICIPANTS

MEMBERS OF CONGRESS:

Representative Ted Budd

Representative Jim Cooper

Representative John Curtis and Susan Curtis

Representative Susan Davis

Representative Ted Deutch

Representative John Garamendi and Patti Garamendi

Representative Tom Graves and Julie Graves

Representative Rick Larsen and Tiia Karlén

Senator Ed Markey and Susan Blumenthal

Representative Jan Schakowsky and Bob Creamer

Representative Peter Welch

Senator Roger Wicker and Gayle Wicker

SCHOLARS AND SPEAKERS:

Hal Abelson	Professor of Computer Science and Engineering, MIT
Alessandro Acquisti	Professor of Information Technology and Public Policy, Carnegie Mellon University
Jack M. Balkin	Knight Professor of Constitutional Law and the First Amendment, Yale University Law School
Howard Beales	Professor of Strategic Management and Public Policy, George Washington University
Doug Beck	Vice President, Americas and Northeast Asia, Apple, Inc.
Joy Buolamwini	Founder, Algorithmic Justice League & PhD student, MIT Media Lab
Larry Downes	Project Director, Georgetown Center for Business and Public Policy

R. David Edelman	Director, Project on Technology, Economy and National Security, MIT
Charlie Firestone	Executive Director, Aspen Institute Communications and Society Program
Kristine Gloria-Garcia	Associate Director, Aspen Institute Communications and Society Program
L. Rafael Reif	President, Massachusetts Institute of Technology
Cathy O'Neil	Founder, O'Neil Risk Consulting and Algorithmic Auditing
Daniel Weitzner	Founding Director, MIT Internet Policy Research Initiative
Jonathan Zittrain	George Bemis Professor of International Law, Harvard Law School
Ethan Zuckerman	Director, Center for Civic Media, MIT Media Lab

FOUNDATION REPRESENTATIVES:

Keesha Gaskins-Nathan	Director, Democratic Practice, Rockefeller Brothers Fund
Tom Glaisyer	Managing Director, Public Square Program, The Democracy Fund

RAPPORTEUR:

Grace Abuhamad	Graduate student, Technology and Policy Program, MIT

ASPEN INSTITUTE:

Dan Glickman and Rhoda Glickman	Executive Director, Congressional Program
Brian Hopkins	Program Development Coordinator
Lauren Kennedy	Manager of Congressional Engagement
Bill Nell	Deputy Director, Congressional Program
Carrie Rowell	Conference Director, Congressional Program