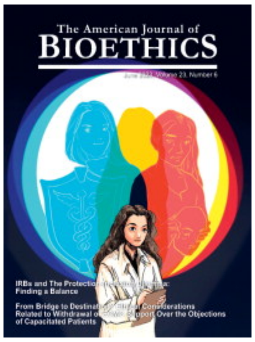

Taylor & Francis
Taylor & Francis Group

The American Journal of Bioethics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uajb20

Ethical Responsibilities for Companies That Process Personal Data

Matthew S. McCoy, Anita L. Allen, Katharina Kopp, Michelle M. Mello, D. J. Patil, Pilar Ossorio, Steven Joffe & Ezekiel J. Emanuel

To cite this article: Matthew S. McCoy, Anita L. Allen, Katharina Kopp, Michelle M. Mello, D. J. Patil, Pilar Ossorio, Steven Joffe & Ezekiel J. Emanuel (2023): Ethical Responsibilities for Companies That Process Personal Data, The American Journal of Bioethics, DOI: 10.1080/15265161.2023.2209535

To link to this article: https://doi.org/10.1080/15265161.2023.2209535

Published online: 01 Jun 2023.

THE AMERICAN JOURNAL OF BIOETHICS
https://doi.org/10.1080/15265161.2023.2209535

Taylor & Francis
Taylor & Francis Group

TARGET ARTICLE

Ethical Responsibilities for Companies That Process Personal Data

Matthew S. McCoy[a], Anita L. Allen[b], Katharina Kopp[c], Michelle M. Mello[d], D. J. Patil[e], Pilar Ossorio[f], Steven Joffe[a], and Ezekiel J. Emanuel[a]

[a]Perelman School of Medicine, University of Pennsylvania; [b]University of Pennsylvania Carey School of Law; [c]Center for Digital Democracy; [d]Stanford Law School and Stanford University School of Medicine; [e]Belfer Center, Harvard Kennedy School; [f]University of Wisconsin School of Law, Morgridge Institute for Research

ABSTRACT

It has become increasingly difficult for individuals to exercise meaningful control over the personal data they disclose to companies or to understand and track the ways in which that data is exchanged and used. These developments have led to an emerging consensus that existing privacy and data protection laws offer individuals insufficient protections against harms stemming from current data practices. However, an effective and ethically justified way forward remains elusive. To inform policy in this area, we propose the Ethical Data Practices framework. The framework outlines six principles relevant to the collection and use of personal data—minimizing harm, fairly distributing benefits and burdens, respecting autonomy, transparency, accountability, and inclusion—and translates these principles into action-guiding practical imperatives for companies that process personal data. In addition to informing policy, the practical imperatives can be voluntarily adopted by companies to promote ethical data practices.

KEYWORDS
Confidentiality & privacy; health policy; regulatory issues

INTRODUCTION

Private sector companies collect, aggregate, analyze, and sell vast amounts of information about us. They capture information about our web browsing, health habits, relationships, and even our daily movements (Pike 2020). The resultant troves of data have the potential to facilitate research and product development that improve public health and well-being, but the collection, use, and sale of personal data at such scale raise ethical concerns (Feathers et al. 2022; Khoury et al. 2018; McGraw and Mandl 2021; Mooney and Pejaver 2018).

It has become increasingly difficult for individuals to exercise meaningful control over the information they disclose to companies or to understand and track the ways in which other entities, including companies that have no direct relationships with patients or consumers, repackage, resell, and reuse that information (Allen 2021). Even scraps of seemingly insignificant personal data can be combined and analyzed using algorithmic methods to derive inferences about individuals' medical conditions, reproductive activities, sexual orientation, socioeconomic status, and other sensitive attributes (Wachter and Mittelstadt 2019). These inferences in turn allow companies to predict and influence individual behavior through targeted advertising and can inform decisions about whether to hire someone for a job or approve them for an insurance policy (Binns 2018).

These developments have led to an emerging consensus that existing privacy and data protection laws offer individuals insufficient protections against harms stemming from current data practices (Table 1). However, an effective and ethically justified way forward remains elusive (Fowler 2020). In recent years, several public and private sector bodies have sought to identify a set of ethical principles that should govern the collection and use of personal data. Unfortunately, many of these bodies fail both to define key principles and to spell out those principles' practical implications for organizations that collect and use personal data, thereby limiting the utility of their recommendations (Jobin, Ienca, and Vayena 2019). In a review of recently published guidelines, for instance, Jobin and colleagues find "widespread reference to 'responsible AI,'" but note that the underlying concepts of "responsibility and accountability are rarely defined (p. 394)." Across guidelines, they observe a "gap at the cross-section of principle

CONTACT Matthew S. McCoy mmcco@pennmedicine.upenn.edu Perelman School of Medicine, University of Pennsylvania, Pennsylvania, PA, USA.

Table 1. Ethical concerns raised by the collection and use of personal information.

Concern	Definition	Examples
Unwanted disclosure of sensitive information	Occurs when others observe or are able to infer information regarding medical conditions, sexual orientation, socioeconomic status, or other personal attributes that someone would not wish shared; may result in loss of dignity or "subjective injuries" such as embarrassment or shame (Price and Cohen 2019).	• Data broker MEDbase 200 compiled and sold lists of individuals with classifications including "erectile dysfunction sufferers," "alcoholism sufferers," "AIDS/HIV sufferers," and "rape sufferers" (Libert 2015).
Discrimination	The wrongful imposition of disadvantage or deprivation on people based on their actual or inferred membership in some salient social group (Altman 2020).	• Amazon's facial recognition tool "Rekognition" was found to falsely match people of color to a mugshot database at a higher rate than people of European origin (Redden, Brand, and Terzieva 2020). • Staples, Home Depot, and other companies were found to display higher online prices to potential customers in low-income communities "because those poorer areas had fewer local retail outlets competing with the online stores" (Newman 2014). • Algorithmic risk assessment tools used to predict recidivism have been found to falsely flag Black defendants as future offenders at nearly twice the rate of white defendants (Angwin et al. 2016). • Facebook suspended the accounts of Native Americans because "its algorithm did not recognize their names as real" (Redden, Brand, and Terzieva 2020).
Exploitation	Occurs when one party unfairly benefits in an interaction with another; may occur even in mutually beneficial, consensual interactions (McCoy, Joffe, and Emanuel 2020b).	• Without an open bidding process, Memorial Sloan Kettering Cancer Center licensed images of patients' tissue slides to a startup company in which several institutional leaders held a financial stake (Ornstein and Thomas 2018).
Online Manipulation	"The use of information technology to covertly influence another person's decision-making, by targeting and exploiting their decision-making vulnerabilities" (Susser, Roessler, and Nissenbaum 2019).	• Using data from social media use and online quizzes, Cambridge Analytica claimed to micro-target political advertisement based on individuals' "psychometric" traits (Susser, Roessler, and Nissenbaum 2019).

formulation and their implementation into practice (p. 396)."

To inform policy in this area, we propose the Ethical Data Practices framework. The framework lays out a set of foundational principles relevant to the collection, exchange, and use of personal data and translates them into a set of practical imperatives for data processors, which we illustrate with concrete examples. Thus, the framework offers an ethically defensible, actionable guide that can immediately inform policy and practice in this critical area.

While recent work has sought to generate ethical guidance for the processing of explicitly health-related data, a broader focus on *personal data*, defined as "any information that identifies, relates to, or could reasonably be linked with an individual or household" including "inferences from other personal information that could create a profile about [one's] preferences and characteristics," is necessary (California Consumer Privacy Act (CCPA) 2018). Although certain types of information are more sensitive than others and should be treated as such in assessments of particular data practices, this article's focus reflects the fact that lines between health-related and non-health-related data are increasingly blurred and that a wide range of personal information can contribute to revealing health-related inferences about individuals and groups (Grande et al. 2020).

Although a variety of actors, including those in the government and academic sectors, collect and use personal data, private companies warrant special attention due to their distinctive incentives and regulatory environment. While government agencies and academic institutions have mandates to act in the public interest, private sector companies' primary incentive is to generate financial returns (Pillar 2013). Additionally, government agencies are subject to official oversight activities and academic researchers are subject to institutional oversight and, in many cases, federal research regulations, but private sector companies' collection and use of personal data have comparatively limited oversight and regulation. These differences between private companies and other actors are reflected in disparate levels of public anxiety about what such companies may be doing with people's personal data. Though members of the public have general concerns about the use of their personal data, they are particularly concerned about the actions of private companies, with more than 80% of Americans saying that the risks of companies collecting data about them outweigh the benefits (Auxier et al. 2019).

Finally, while the foundational principles included in the framework are broadly applicable to the collection and use of personal data across jurisdictions, our discussion of how the principles could be operationalized focuses on the policy environment in the United States, which lacks comprehensive privacy and data protection legislation at the federal level.

The Ethical Data Practices framework has two purposes. It is intended both to guide self-regulation within industry and to inform government regulation of companies that handle personal data. First, in the near term, improved self-regulation can be a step in the right direction, a way to effectuate a modest shift toward more ethical data practices while awaiting government action. Even in the event that Congress passes federal privacy legislation, which is by no means guaranteed despite some bipartisan consensus on recent proposals (Kern 2022), ongoing industry self-regulation can play a complementary role by filling gaps in the government's power to monitor and enforce compliance with laws. In the final section of the paper, we discuss why profit-seeking companies have incentives to voluntarily adopt such ethical standards. Ultimately, however, history has shown that self-regulation alone cannot curb unethical data practices. Thus, the second purpose of the framework is to inform laws and regulations that will ultimately be needed to enforce ethical data practices.

LIMITATIONS OF THE CURRENT PRIVACY PROTECTION REGIME

Since the 1970s with the introduction of the Fair Information Practice Principles (FIPPs), a notice-and-consent regime has been the dominant approach to privacy regulation in the United States (Susser 2019). Under this approach, individuals are given "rights to notice, access, and consent regarding the collection, use, and disclosure of personal data," which, in theory, allow them to choose how, with whom, and under what terms they disclose their data (Solove 2013).

The notice-and-consent approach appeals to the principle of respect for individual autonomy—the notion that people should be able to determine how their personal data are collected and used. From this perspective, what justifies different forms of data collection and use is not their substantive value or fairness but the fact that people have agreed to them. Theoretically, consent-based approaches to privacy are attractive because they accommodate a wide range of individual preferences with respect to data sharing. In reality, however, they are flawed in both their underlying ethical justification and their practical application.

The notice-and-consent regime's grounding in respect for individual autonomy fails to account for the full range of ethical considerations raised by the collection and use of personal data. In the era of Big Data, a decision to disclose personal information is rarely purely "personal" because it can affect the privacy and the social standing of others (Fairfield and Engel 2015). Sætra refers to this dynamic as a "privacy externality," wherein "one individual's disclosed information can be used to infer information about other individuals" in the same social network or demographic category (Sætra 2020, 5). Indeed, technology companies' practices are increasingly "aimed primarily at deriving (and producing) population-level insights regarding how data subjects relate to others" (Viljoen 2021, 578). Insofar as they contribute to companies' ability to draw inferences about broader groups to which people belong, individuals' decisions to disclose personal information implicate a broad range of social interests and cannot be justified simply by appeal to individual consent.

Approaches to privacy regulation also have political consequences, as feminist critiques of digital consent practices have highlighted (Carmi 2021). Although notice-and-consent regimes appear to empower individuals to manage their privacy, in many cases they have the opposite effect. They empower companies to dictate the terms of data collection and use to people who cannot realistically negotiate, challenge, or opt out of data practices. Notice-and-consent regimes thus produce asymmetrical power relations wherein companies enjoy enormous discretion to collect and use data in ways with far-reaching social consequences while affected people lack meaningful opportunities to assert their interests.

In practice, notice-and-consent procedures frequently fail even to realize the value of individual autonomy that grounds them (Pike 2020; Susser 2019). First, because many services that collect personal data are integral to daily life, refusing to consent to data collection is often impossible. For example, it is absurd to say that people are free to opt out of privacy-invasive technologies if doing so requires them to forgo the opportunity to seek critical health information, access a government benefits portal, or apply for a job online (Acquisti, Taylor, and Wagman 2016; McCoy et al. 2020a).

Second, to the extent that people can exercise some voluntary choices about how their personal data are collected and used, those choices are rarely fully informed. Over 10 years ago, researchers estimated that it would take the average person roughly 250 h a year to read, let alone comprehend, all the privacy policies they encountered online (Pike 2020). Since then, the quantity and complexity of personal data collection and use and has grown dramatically. In this environment, individuals cannot possibly parse and make informed decisions on the basis of lengthy, legalistic privacy policies.

While there are tools, techniques, and even paid services that individuals can use to monitor and manage their privacy to some degree, they are not equally available to all people (Acquisti, Taylor, and Wagman 2016). Thus, another concerning feature of notice-and-consent regimes is that they exacerbate inequities by turning privacy into a "luxury good" more readily available to those with sufficient resources to protect themselves (Sætra 2020).

Cumulatively, these challenges suggest that a superior data ethics framework must improve upon the notice-and-consent paradigm in two fundamental ways. First, rather than focusing narrowly on the value of individual autonomy, it must be based on a broader set of ethical principles that reflect how companies' collection and use of personal data affect, and indeed shape, society. Second, it must translate these principles into actionable practices.

FOUNDATIONAL PRINCIPLES FOR RESPONSIBLE STEWARDSHIP OF PERSONAL DATA

Commentators have identified as many as 16 distinct substantive and procedural principles related to the collection and use of personal data (Xafis et al. 2019). Such an extensive list may be useful for fine-grained ethical analysis but may prove difficult to use as a practical guide for policy makers and decision makers in organizations that collect and use personal data, given the challenges inherent in balancing so many distinct considerations. We propose a set of foundational principles that capture relevant ethical considerations but are sufficiently parsimonious to understand, balance, and apply. The framework is grounded in three widely accepted substantive principles—minimizing harm, fairly distributing benefits and burdens, and respecting individual autonomy—along with three procedural principles—transparency, accountability, and inclusion.

The three substantive principles are similar to the basic ethical principles identified in the Belmont Report. In defending an ethical framework grounded in these familiar principles, we differ from recent commentators like Elizabeth Pike, who has argued that the Belmont Report is "generally ill-suited" to today's "complex, networked data landscape" (Pike 2020). While we agree that specific applications of these principles, such as the requirement for individual informed consent, are ill-suited to the collection and use of personal data, we argue that the basic principles address key ethical considerations raised by the collection and use of personal data. Moreover, the success of the Common Rule has shown that these principles can serve as an intelligible and enduring ethical basis for policy. Finally, anchoring the ethical framework in established principles helps to ensure its alignment with global efforts to promote ethical data use, such as the World Health Organization's recent guidance on the use of artificial intelligence for health, which is based on a similar set of basic ethical requirements (World Health Organization 2021).

Minimizing harm requires structuring the collection and use of personal data to facilitate socially valuable data practices while mitigating risks of discrimination, exploitation, and other harms to individuals and groups. Realizing this value requires identifying and weighing potential benefits and harms from particular data practices. While data practices whose harms outweigh their benefits are ethically unacceptable, even data practices that advance legitimate individual and social interests should be designed and carried out in a way that minimizes risks to individuals and groups.

Fairly distributing benefits and burdens addresses the question of how the beneficial and harmful impacts of various data practices are spread across different social groups. It requires particular attention to the ways in which data practices may exacerbate patterns of social disadvantage or exploit certain groups to the benefit of others. This principle entails both negative and positive obligations for companies that collect and use personal data. Specifically, it requires avoiding data practices that disproportionately burden or discriminate against particular groups while promoting equitable access to the benefits produced.

Respecting individual autonomy requires giving weight, within the constraints imposed by the obligations to minimize harm and to fairly distribute benefits and burdens, to people's considered choices about how their personal data are collected and used. As the limitations of the notice-and-consent paradigm illustrate, respecting individual autonomy cannot serve as the sole foundational value for data ethics, but balanced in conjunction with other substantive values it has an important role to play. Individuals should not be allowed to make autonomous decisions to participate in data practices that exploit or discriminate against others. However, there are a range of ethically permissible data practices in which individuals should be able to make an informed choice to engage—for example, allowing a consumer genetics company to share their data for vetted research projects. In short, the principle of respecting individual autonomy recognizes that, over the range of permissible practices, individuals have different preferences and can make different yet reasonable decisions about what information they wish to disclose and for which purposes. Presenting individuals with formal choices is not enough to realize the principle. Instead, individuals must have effective opportunities for voluntary and informed choices that companies in turn honor.

Respecting individual autonomy is also linked to dignity (Price and Cohen 2019). Individuals may suffer "dignitary harms" when others access or infer information that individuals wish to keep private (Price and Cohen 2019). Thus, respecting people's considered choices about which parts of their lives are revealed to others and for what reasons respects their dignity.

These three principles capture the key substantive ethical considerations that should guide decisions about companies' collection and use of personal data. But simply identifying these substantive principles is insufficient to ensure ethical practices, especially since the principles are framed in absolute terms, yet decisions about personal data collection and use often require balancing substantive principles against one another. For example, distributing harms in an equitable manner across groups in a population may mean not minimizing harm for a particular group. To guide future decisions about tradeoffs, procedural principles are needed. Procedural principles specify *how* organizations ought to make decisions about data collection and use—ensuring, for example, that their decision making is open to scrutiny and informed by insights and concerns of relevant constituencies. Adopting such principles can impose ethical guardrails in the context of the asymmetrical power relations between companies that collect and use personal data and those affected by their decisions. Thus, in a comprehensive and cohesive ethical framework, substantive principles must be paired with procedural principles.

Three procedural principles are relevant in this context. *Transparency* does not require that companies disclose trade secrets but does require that they disclose relevant information about their data practices, including the purposes for which they collect and use personal data, the parties with which they engage in data transactions, the reasonably foreseeable risks associated with their data practices, and the efforts they take to mitigate those risks. Transparency can guard against harmful or discriminatory data practices by exposing them to scrutiny and informing advocacy for policy change at the industry or government levels. Transparency also demonstrates respect for individual autonomy by providing individuals with relevant information that can inform their choices to interact with, avoid, criticize, or support companies for their data practices.

Transparency is necessary but not sufficient for achieving *accountability*. Accountability requires standards for behavior and the creation of mechanisms by which organizations can be sanctioned for failing to meet those standards either through reputational sanctions imposed by the public or fines and other legal penalties imposed by regulators. The goal of accountability mechanisms is to ensure that companies collecting and using personal data are answerable to the communities that their practices affect.

Finally, the principle of *inclusion* expresses the idea that members of affected groups should have meaningful opportunities to shape the policies of companies that collect and use their personal data and the industries of which they are a part. Like transparency, inclusion demonstrates respect for individuals by taking their concerns seriously. Structured appropriately, opportunities for inclusion can also increase the benefits and guard against the potential harms of data practices by helping to ensure that those practices incorporate the perspectives and insights of affected people.

PRACTICAL IMPERATIVES

The foundational principles described above provide the ethical basis for evaluating the practices of companies that collect and use personal data. In this section, we describe several ways in which these principles can be translated into practical imperatives for companies (Table 2). These imperatives can be used together as a checklist for decision-makers within companies. By asking how well their policies and practices achieve these imperatives, decision-makers can better understand how well their policies conform to underlying ethical principles. Other groups, including the European Commission's High-Level Expert Group on

Table 2. Practical imperatives for organizations that collect and use personal data.

Practical imperative	Principles emphasized	Actions
Minimize collection and retention of personal data	• Fairly distributing benefits and burdens • Respecting individual autonomy • Minimizing harm	• Limit data collection, use, and retention to what is necessary to provide a requested product or service or what is reasonably anticipated in the context of a company's relationship with an individual.
Offer fewer but more meaningful choices about data	• Fairly distributing benefits and burdens • Respecting individual autonomy	• Eliminate the sharing or sale of personal information for low value uses rather than forcing users to opt out. • When sharing data for socially valuable purposes such as medical research, take steps to minimize burdens on individuals, including offering succinct statements of the risks and benefits of data sharing and using opt-in consent models that allow individuals to limit the sharing of their data by default.
Provide meaningful disclosure	• Fairly distributing benefits and burdens • Respecting individual autonomy • Transparency • Accountability	• Develop succinct, consumer-facing data use statements and context-specific notifications that can be readily understood by consumers. • Publish comprehensive privacy policies that can be assessed by regulators, watchdogs, and other third parties.
Assess the social impact of data practices	• Minimizing harm • Fairly distributing benefits and burdens	• Conduct periodic social impact assessments of data practices with a focus on identifying and correcting practices that burden, exclude, exploit, or discriminate against members of disadvantaged groups. • Use a checklist process prior to product launch to identify and mitigate risks associated with new products.
Ensure meaningful stakeholder engagement	• Minimizing harm • Accountability • Inclusion	• Involve members of the public and external experts in social impact assessments. • Create standing bodies, such as data access committees or representative assemblies, to allow for ongoing input from users and affected groups.

Artificial Intelligence, have defended similar checklist approaches to promoting ethical data practices (European Comission High-Level Expert Group on AI 2019).

Checklists of this sort can provide a useful roadmap for companies that are motivated to act ethically. Media, watchdog organizations, and the broader public can help to supply this motivation by applying pressure on companies to adhere to ethical principles. However, ensuring consistent adherence to ethical data practices will require external oversight. To that end, policymakers can use these imperatives as standards for evaluating corporate practices and policies, and ultimately developing legislation and enforcement mechanisms. While we focus on cases involving companies that process unambiguously health-related data, the imperatives apply to any organization that collects or uses personal information and could form the basis for a regulatory regime that crosses industry sectors.

The first three practical imperatives are closely related in that each aims to minimize burdens on individuals. This promotes fairness by lowering the costs to individuals of exercising control over their data, which benefits those with fewer time, financial, and informational resources. Minimizing burdens on individuals also promotes individual autonomy by letting people focus their attention on the decisions that matter most.

(1) *Minimize collection and retention of personal data*. This imperative combines ideas of purpose limitation—the idea that there should be limitations on the purposes for which companies can collect and share data—and data minimization—the idea that data should be limited to what is necessary for the purposes for which they are collected and used. While existing and proposed privacy legislation enshrine the ideas of purpose limitation and data minimization, they have been interpreted in different ways, not all of which are consistent with the goal of reducing individual burdens and thereby advancing the underlying goals of fairness and autonomy. In particular, purpose limitations are sometimes construed to allow data collection for a wide range of purposes so long as they are specified in companies' privacy policies (Kerry et al. 2020). This approach is too permissive to meaningfully reduce burdens on individuals. Instead, companies should limit data collection and use to that which is relevant to providing a product or service that is specifically requested or may reasonably be anticipated by an individual who enters into a relationship with a company (Kerry et al. 2020).

The imperative to restrict collection and retention of data to the minimum necessary is a stringent constraint that would impose significant limitations on common practices. Many health and fitness-related smartphone apps, for example, capture more user information than they need to function properly. When users download an app to map their runs or bike rides, they reasonably anticipate that the app will record their geolocation data while they are using it. However, such apps have been found to continue tracking users' movements even when they are not using the apps (Brandtzaeg, Pultier, and Moen 2019). Such round-the-clock GPS tracking, which can reveal the location and identity of individuals' workplaces, doctor's offices, and close relations, goes far beyond users' reasonable expectations (Valentino-DeVries et al. 2018).

Meeting this imperative would also require modifying common data retention practices. In a recent study of direct-to-consumer genetic testing (DTCGT) companies' privacy policies, 45% indicated that companies would retain genetic data indefinitely or until a customer requested its deletion, while 42% failed to address data retention explicitly but implied that customers' genetic data could be held indefinitely (Hazel and Slobogin 2018). By forcing consumers first to sift through dense privacy policies to understand how long their information will be retained and to take active steps to delete it, such practices disproportionately and unfairly burden individuals. The question of how long consumer data may be reasonably retained depends on the nature of the services being offered or of the research being conducted. While long-term data retention may benefit companies and facilitate socially valuable undertakings such as biomedical research, there are likely few scenarios in which a company would be justified in retaining personal information indefinitely unless consumers actively request that it be deleted.

(2) *Offer fewer but more meaningful choices about data.* Individuals currently face an overwhelming number of choices about how their data are collected, used, and transmitted to third parties. Reducing decisional burden lets people focus their attention on consequential choices.

To realize this imperative, companies should eliminate the sale or sharing of personal information for uses with low social value like behavioral advertising, which consumers find undesirable and invasive and which may be only marginally or no more profitable for ad publishers than contextual advertising (Boerman, Kruikemeier, and Zuiderveen Borgesius 2017; Davies 2019; Marotta, Abhishek, and Acquisti 2019). Indeed, in a survey of digital ad publishers, the majority of respondents reported that behavioral advertising had "not produced any notable benefit" for their businesses, while 23% reported that behavioral advertising had caused advertising revenue to decline (Weiss 2019). Although privacy laws like the California Consumer Privacy Act (CCPA) already give consumers a right to opt out of the sale of their personal information, exercising this right requires individuals to make an ongoing series of active choices to limit the sale of their data across a vast number of platforms (Waddell 2020). Simply eliminating the sale of personal information for low value uses spares individuals the burden of trying to navigate such "slow, confusing, [and] frustrating" processes to prevent unwanted data sharing (Waddell 2020). Better still would be ending the use of personal information for behavioral advertising altogether, which would also prevent organizations from using personal information collected on their own platforms for targeted ads.

Although organizations should be permitted to share data for socially valuable purposes such as medical research, they should still take steps to minimize burdens on individuals, such as publishing succinct statements on the purposes, risks, and benefits of sharing data. Another option is to use opt-in models of consent. Unlike opt-out models of consent, which require individuals to take action to limit data sharing, opt-in consent removes the onus from individuals by limiting the sharing of their data by default. Though opt-in consent may not be feasible or appropriate in all contexts, the experience of DTCGT companies suggests that opt-in models can be viable for organizations that wish to share user data for research purposes (Laestadius, Rich, and Auer 2017). The personal genomics company 23andMe, for example, reported that 80% of its users opted in to its research programs (Hart 2019).

Ultimately, questions of what constitutes a valuable data practice must be answered on a case-by-case basis. Organizations should develop processes for making such determinations that abide by the procedural principles of transparency, accountability, and inclusion.

(3) *Provide meaningful disclosure.* Respecting autonomy requires that companies provide information about their data practices that individuals realistically can use to make informed decisions about interacting with that company. Disclosure also helps stakeholders understand where to focus their efforts in advocating for policy reforms to better align data use practices with the public interest. Recent laws like the CCPA include a requirement that companies create and publish plain-language privacy policies. However, even CCPA-compliant privacy policies may still be long and difficult for average users to parse. As Kerry and colleagues propose, companies should develop succinct, consumer-facing data use statements and context-specific notifications that consumers can more readily use, while also publishing comprehensive privacy policies that regulators, watchdogs, and other third parties can assess to foster accountability (Kerry et al. 2020). Ideally, companies would also make public an annual accounting of the nature and extent of

their transmission of user data to other companies and to academic researchers.

(4) *Assess the social impact of data practices*. Data practices must minimize risks to both individuals and groups. Furthermore, to realize the value of fairly distributing benefits and burdens, organizations must also ensure that their practices do not disproportionately burden, exclude, exploit, or discriminate against members of disadvantaged groups. One mechanism for promoting consistency with these principles is conducting regular social impact assessments of data practices.

Social impact assessments expand upon the established model of privacy impact assessments (Edwards, McAuley, and Diver 2016; Mantelero 2018). While privacy impact assessments have traditionally focused narrowly on individual privacy, a more holistic social impact approach would consider impacts on "groups or categories identified by characteristics that include ethnicity, race, religion, national origin, political affiliation and sexual orientation" as well as the political effects of data practices (Raab and Wright 2012).

The proposed Consumer Online Privacy Rights Act (COPRA) provides a partial model in its requirement that covered entities conduct annual algorithmic decision-making impact assessments if they use algorithmic techniques for eligibility or advertising decisions related to housing, education, employment, or credit opportunities (Cantwell 2019). These assessments are intended to focus on "whether the algorithmic decision-making system produces discriminatory results on the basis of an individual's or class of individuals' actual or perceived" attributes (Section 108 (b)(1)(B) of the bill in Cantwell 2019). While guarding against algorithmic discrimination is critical, it should not be the sole focus of social impact assessments, which should address the broader risks and benefits of data practices. Impact assessment also offers an opportunity to consider issues of digital accessibility and ensure that products are designed in such a way that they can be used by those with disabilities (Lazar, Goldstein, and Taylor 2015).

Companies can also use a checklist process to assess potential risks of new technologies prior to launch (Loukides, Mason, and Patil 2018). Requiring developers to answer questions such as "have we studied and understood possible sources of bias in our data?" or "have we identified and addressed potential access barriers for people with perceptual disabilities?" can help companies work proactively to prevent harms and assure fair access to their products. Finally, both periodic and prospective assessments provide opportunities for companies—which often lack in-house ethics expertise and diversity across multiple dimensions among their workforces—to seek input from external experts and stakeholders, thereby furthering the principle of inclusion.

(5) *Ensure meaningful stakeholder engagement*. A requirement to conduct transparent social impact assessments can give companies incentives to maintain socially acceptable data practices. However, these incentives are strongest when individuals can sanction companies by cutting ties with them (Hirschman 1972). In reality, individuals often lack this ability because companies' platforms and services are integral to daily life. Additionally, once collected by a company (and possibly transmitted to others), personal information may be difficult for a user to withdraw. Similar dynamics can arise when health systems share patient data with technology companies for the purpose of developing improved electronic health record systems or better predictive algorithms (Cohen and Mello 2019; McCoy, Joffe, and Emanuel 2020b).

Particularly when individuals are compelled to interact with entities that collect, use, and share their personal data, the principle of inclusion requires entities to provide meaningful opportunities for affected persons to inform their policies (Hirschman 1972). Involving members of the public in social impact assessments is one way of ensuring meaningful stakeholder engagement. Involving members of disability communities in such assessments can, for example, be one way to vet accessibility of new technologies. Organizations should also consider developing independent standing bodies to allow for ongoing engagement. In the context of data sharing arrangements between health systems and for-profit companies, Cohen and Mello have argued for the creation of data access committees to review requests for patient data. Under their proposal, at least half the committee's membership would comprise patients in the health system whose data are being sought (Cohen and Mello 2019). Similarly, Post recently proposed the development of a "Facebook representative assembly" selected by platform users (Post 2021).

IMPLEMENTATION

The Ethical Data Practices framework can serve both to guide self-regulation at the firm- or industry-level and to inform the development of laws and regulations.

On its own, self-regulation is an inadequate response to the misuse of personal data. Yet in the absence of comprehensive data protection legislation enforced by well-resourced regulatory agencies, enhanced self-regulation offers one of the only pathways to incrementally improving data practices. Self-regulation also has long term value as an adjunct to government regulation because even empowered regulatory agencies have finite oversight and enforcement capacities. Regulators could, for example, demand evidence that companies have conducted social impact assessments, but they are not well-positioned to ensure that companies are working proactively and deliberately at all phases of their decision-making to prevent harms and assure the fairness of new products.

One might ask why companies, whose primary goal is to generate financial returns, would adopt ethical standards that may interfere with their ability to monetize the data they collect. However, companies and industries have some incentive to self-regulate when they recognize that long-term reputational costs of engaging in unpopular data practices outweigh the benefits of pursuing short-term profit maximization, or that failure to self-regulate may spur more onerous public regulation (Cusumano, Gawer, and Yoffie 2021). Additionally, if companies anticipate the passage of more stringent data protection legislation, they may also find that adopting ethical data practices in the short term may lower the downstream compliances costs (Weinberg 2019; Floridi 2021).

At a time when many consumers are concerned about minimizing their digital footprints, individual companies may also see market opportunities in transparently committing to ethical data practices (Weinberg 2019). For example, churn in public use of different social media platforms may present an opportunity for new entrants to market their platforms based on privacy protections. Similarly, systematically assessing the accessibility of new technologies while helping to ensure that their benefits are fairly distributed may appeal to companies as a way to enlarge the market for their products (Lazar, Goldstein, and Taylor 2015).

At the industry-level, companies can be motivated by the prospect of a tragedy of the commons—whereby companies undermine public trust in the broader industry in pursuit of short-term individual gains—to take collective action to self-regulate (Cusumano, Gawer, and Yoffie 2021; Listokin 2017). Indeed, such motivations have driven recent attempts at self-regulatory reform within the tech industry (CDT and eHI Release Proposed Consumer Privacy Framework for Unprotected Health Data 2022; Robbins 2019).

To the extent that companies, individually or collectively, are incentivized by these factors to voluntarily adopt ethical standards, the framework we offer here can be adapted and implemented at multiple levels. Within organizations, decision-makers can use the imperatives to structure ethics review processes. Large organizations should appoint chief ethics officers or ethics review boards to oversee these processes, but even smaller organizations without in-house ethics expertise can use the imperatives to guide ethical review of their practices. Box 1 describes how a hypothetical developer of a mental health app designed to help users track changes in their moods might align its practices with the imperatives.

Box 1. How developers of a mental health app could meet the framework's practical imperatives.

Ensure meaningful stakeholder engagement

Beginning in the product development phase, decision-makers in the company convene external stakeholders including patients, mental health professionals, and, eventually, users of the app to shape the app's data privacy settings and controls and review any potential data sharing arrangements.

Minimize collection and retention of personal data

With stakeholder input, the company identifies and continually reassesses what types of personal information are necessary to the functioning of the app and ensures that there is a reasonable justification for the data it collects and that it is not collecting superfluous information or unnecessarily requiring users to link their account to other online profiles.

Offer fewer but more meaningful choices about data

By default, users' personal information is not shared with researchers. However, users may be given the opportunity to opt in to sharing their personal data with stakeholder-vetted research partners. In these cases, the company presents users with clear statements on the risks and benefits of research as well as clear instructions for how to stop sharing their data.

Provide meaningful transparency

The company engages patients and users of the app to develop a succinct, accessible statement that addresses the types of data collected by the app, how the data will be used, and the terms, if any, under which the data may be shared.

Assess the social impact of data practices

Prior to product launch, the company uses a checklist—informed by input from patients and mental health professionals—to assess and mitigate risks associated with the app and to ensure that it is meeting the needs of its target users, including those with disabilities. After product launch, the company regularly convenes stakeholders to review and correct practices that may stigmatize patients, lead to discrimination, or otherwise harm individuals or groups.

The framework can also form a core component of ethics training modules for data scientists. Sophie Zhang, a Facebook employee who blew the whistle on "blatant attempts by foreign national governments to abuse our platform on vast scales to mislead their own citizenry," showed that data scientists can provide an important check on unethical practices (Wong 2021). By training data scientists to better understand ethical issues raised by their work and empowering them to voice concerns and, if necessary, halt product development processes, companies can promote compliance with ethical standards (Loukides, Mason, and Patil 2018).

To foster accountability, companies could publicly communicate how they are meeting the framework's imperatives through the publication of ethical data practices reports that third-party organizations can digest and critique in a manner accessible to the lay public. Members of the public, too, can use the framework as a basis for evaluating the ethical implications of decisions by companies and regulators and advocating for improved policies and practices.

Third-party watchdog and consumer advocacy organizations can also adopt the framework as a rubric for evaluating or auditing companies' data practices. There are already platforms in place to support these processes, such as the Ranking Digital Rights Corporate Accountability Index (2020 Ranking Digital Rights Corporate Accountability Index 2021). Benchmarking platforms from other industries, such as the Good Pharma Scorecard, can also provide useful models (Miller et al. 2017). Similarly, trade associations could work to develop standards based on the framework and require compliance among their members.

Large companies like Apple, Google, and Amazon, which play a critical gatekeeping role by regulating the conduct of companies that use their platforms, could also require companies to comply with the terms of the framework in order to sell products or services on those platforms (Dempsey et al. 2021). Apple's recent decision to require app developers to seek users' permission before tracking them is a good example of how a large technology company can make it easier for users to exercise meaningful control over their data (Chen 2021). Similarly, health care organizations like hospitals can require compliance with the framework by developers that wish to use their data.

Although improved self-regulation may offer some benefit in the near team, the shortcomings of previous privacy self-regulation efforts show that ensuring organizations' consistent and reliable compliance with ethical data practices will ultimately require government action (Gellman and Dixon 2011; Hoofnagle 2005). To this end, lawmakers working to develop comprehensive data protection and privacy legislation could use the framework to map key principles to practical imperatives that ought to be enshrined in law. Similarly, a federal agency like the FTC or a newly-created federal data agency like the one proposed in the Data Protection Act of 2020, with the authority and resources to create and enforce data protection rules, could use the framework to develop rules for implementing and enforcing new legislation (Gillibrand 2020). Such an agency could also play a vital role in institutionalizing the principle of inclusion by ensuring that the public has an ongoing and meaningful say in how personal data are collected and used. While companies can and should explore mechanisms for engaging individuals and groups affected by their decision-making, federal agencies are better positioned to engage representative publics. There are precedents for this type of engagement in agency decision-making. For example, the FDA's Patient-Focused Drug Development initiative systematically engages patients to provide insight into how to navigate "tradeoffs between treatment benefit and risk outcomes" in drug development (CDER Patient-Focused Drug Development 2020). A data protection agency could similarly engage citizens in shaping and implementing federal data regulation standards.

CONCLUSION

While the use of personal data by private companies holds the potential to contribute to socially valuable research and innovation, it also involves recognized risks and burdens for individuals and communities. Making progress toward responsible data practices in private companies requires an ethical framework that is sufficiently parsimonious and concrete to be feasible to apply. The broad principles and specific imperatives proposed here can help companies take immediate actions to promote ethical data practices and, in the long run, can support the development of a robust accountability and enforcement regime.

ACKNOWLEDGEMENTS

The authors thank Amaya Diana and Cathy Zhang for research support and Marcy Wilder for her intellectual contributions to the project and feedback on an early version of the manuscript.

DISCLOSURE STATEMENT

Dr. Joffe is a paid member of a data and safety monitoring board for CSL Behring. Dr. Mello was an advisor to Verily Life Sciences LLC from 2020-22 on product designed to facilitate safe return to work and school during COVID-19.

Dr. Ossorio does bioethics consulting for Eli Lilly and Roche-Genentech.

Dr. Patil is an investor at GreatPoint Ventures and an limited partner in several other funds that focus on starting companies in the data and associated technology space. He is also a board member at Devoted Health, as well as a member of the Defense Science Board.

Dr. Emanuel serves on the following: Board of Advisors, Cellares; Occasional Advisor, Clarify Health; Unpaid External Advisory Board Member, Village MD; Occasional Advisor, Notable; Advisory Board Member, JSL Health Fund; Advisory Board Member, Peterson Center on Healthcare; Special Advisor, to Director General WHO; Expert Advisory Member, WHO COVID-19 Ethics & Governance Working Group; Advisory Board Member, Biden's Transition COVID-19 Committee; Advisory Board Member, HIEx Health Innovation Exchange Partnership sponsored by the UN Geneva. He also reports the following engagements: uncompensated speaker UNESCO Global Conference, uncompensated speaker School of Pharmaceutical & Biotech Business, uncompensated speaker Blue Cross Blue Shield, uncompensated speaker National Health Equity Summit, uncompensated speaker The Galien Foundation, uncompensated speaker Temple Shalom Chicago speaker series, uncompensated speaker AIFA Italian Medical Agency, uncompensated speaker Rainbow Push Coalition/CEF, uncompensated speaker IDSA, personal fees from Rise Health, uncompensated speaker Vin Future, personal fees from Well Sky, personal fees from Rightway, uncompensated speaker Brown University, personal fees from Signature Healthcare Foundation, uncompensated speaker Organisation for Economic Cooperation & Development, personal fees from Healthcare Leaders of New York, uncompensated speaker 21st Population Health Colloquium, personal fees from Medimpact, uncompensated speaker American Academy of Arts & Sciences, uncompensated speaker Village MD, uncompensated speaker The Galien Foundation, uncompensated speaker University of Sydney Australia, personal fees from Massachusetts Association of Health Plans, uncompensated speaker Virtahealth, uncompensated speaker Health Action Alliance, personal fees from Yale University Grand Rounds, personal fees from Hartford Medical Society, uncompensated speaker UCSF, uncompensated speaker Ichan School of Medicine, uncompensated speaker University of Minnesota, uncompensated speaker IPHS Addis Conference, personal fees from AAHC Global Innovation Forum, personal fees from HMSA & Queens Health System, uncompensated speaker Faith Health Alliance Project, Travel fees from Macalester College, uncompensated speaker CDC Learning event, Travel fees Oak CEO Summit, uncompensated speaker Primary Care Transformation Summit, uncompensated speaker 16th World Congress of Bioethics, uncompensated speaker Blue Cross Blue Shield Research Health Alliance, personal fees from Advocate Aurora Health Summit, Travel fees DPharm Conference, personal fees from UPMC Shadyside Medical Center, uncompensated speaker ASCO Quality Care Symposium, Travel fees UCSF Department of Urology Grand Rounds, personal fees from Advocate Aurora Health, personal fees from Cain Brothers Conference, personal fees from Bowdoin College, Travel fees The Galien Foundation Jerusalem Ethics Forum, Travel fees HLTH 2022 Las Vegas, uncompensated speaker National Academies Forum on Microbial Threats, uncompensated speaker National University of Singapore, uncompensated speaker Williams College, uncompensated discussant NIH Grand Rounds, Travel fees HMSA, uncompensated speaker Stanford Graduate School of Business, uncompensated panelist World Bank, Travel fees Tel Aviv University, Travel fees The Suntory Foundation, personal fees Ontario Hospital Association.

FUNDING

Work on this project was supported by a gift from Deloitte to the University of Pennsylvania.

REFERENCES

2020 Ranking Digital Rights Corporate Accountability Index. 2021. Ranking *digital rights.* Accessed March 14, 2021. https://rankingdigitalrights.org/index2020/

Acquisti, A., C. R. Taylor, and L. Wagman. 2016. The economics of privacy. *Journal of Economic Literature* 52 (2): 442–492. doi:10.2139/ssrn.2580411.

Allen, A. L. 2021. HIPAA at 25—a work in progress. *The New England Journal of Medicine* 384 (23):2169–71. doi:10.1056/NEJMp2100900.

Altman, A. 2020. Discrimination. In *The Stanford encyclopedia of philosophy*, ed. E. N. Zalta. Stanford, CA: Metaphysics Research Lab, Stanford University.

Angwin, J., J. Larson, S. Mattu, and L. Kirchner. 2016. Machine bias. *ProPublica*. Accessed May 25, 2023. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Auxier, B., L. Rainie, M. Anderson, A. Perrin, M. Kumar, and E. Turner. 2019. Americans and privacy: Concerned, confused and feeling lack of control over their personal information. Pew Research Center: Internet, Science & Tech.

Binns, R. 2018. Algorithmic accountability and public reason. *Philosophy & Technology* 31 (4):543–56. doi:10.1007/s13347-017-0263-5.

Boerman, S. C., S. Kruikemeier, and F. J. Zuiderveen Borgesius. 2017. Online behavioral advertising: A literature review and research agenda. *Journal of Advertising* 46 (3):363–76. doi:10.1080/00913367.2017.1339368.

Brandtzaeg, P. B., A. Pultier, and G. M. Moen. 2019. Losing control to data-hungry apps: A mixed-methods approach to mobile app privacy. *Social Science Computer Review* 37 (4):466–88. doi:10.1177/0894439318777706.

California Consumer Privacy Act (CCPA). 2018. California Department of Justice, Office of the Attorney General. October 15.

Cantwell, M. 2019. Text - S.2968 - 116th Congress (2019–2020): Consumer Online Privacy Rights Act. Webpage. 2019/2020. December 3.

Carmi, E. 2021. A feminist critique to digital consent. *Seminar.net* 17 (2):1–21. doi:10.7577/seminar.4291.

CDER Patient-Focused Drug Development. 2020. FDA. December 22. Accessed May 25, 2023. https://www.fda.gov/drugs/development-approval-process-drugs/cder-patient-focused-drug-development.

CDT and eHI Release Proposed Consumer Privacy Framework for Unprotected Health Data. 2022. Center for Democracy and Technology. Accessed October 21, 2022. https://cdt.org/press/cdt-and-ehi-release-proposed-consumer-privacy-framework-for-unprotected-health-data/

Chen, B. X. 2021. To be tracked or not? Apple is now giving us the choice. *The New York Times*, April 26, sec. Technology.

Cohen, I. G., and M. M. Mello. 2019. Big data, big tech, and protecting patient privacy. *JAMA* 322 (12):1141–2. doi:10.1001/jama.2019.11365.

Cusumano, M. A., A. Gawer, and D. B. Yoffie. 2021. Social media companies should self-regulate. Now. *Harvard Business Review*, January 15.

Davies, J. 2019. After GDPR, The New York Times cut off ad exchanges in Europe - and kept growing ad revenue. *Digiday*. https://digiday.com/media/gumgumtest-new-york-times-gdpr-cut-off-ad-exchanges-europe-ad-revenue/.

Dempsey, J., C. J. Hoofnagle, I. Rubinstein, and K. J. Strandburg. 2021. A broader look at privacy remedies. *Lawfare Blog*. Accessed May 25, 2023. https://www.lawfareblog.com/broader-look-privacy-remedies.

Edwards, L., D. McAuley, and L. Diver. 2016. From privacy impact assessment to social impact assessment. In *2016 IEEE Security and Privacy Workshops (SPW)*. doi:10.1109/SPW.2016.19.

European Comission High-Level Expert Group on AI. 2019. *Ethics guidelines for tustworthy AI: Shaping Europe's digital future*. Brussels: European Commission.

Fairfield, J. A. T., and C. Engel. 2015. Privacy as a public good. *Duke Law Journal* 65(3): 385–422.

Feathers, T., S. Fondrie-Teitler, A. Waller, and S. Mattu. 2022. Facebook is receiving sensitive medical information from hospital websites. *Stat News*. Accessed May 25, 2023. https://www.statnews.com/2022/06/16/facebook-meta-pixel-hospitals-data/

Floridi, L. 2021. The end of an era: From self-regulation to hard law for the digital industry. *Philosophy & Technology* 34 (4):619–22. doi:10.1007/s13347-021-00493-0.

Fowler, G. A. 2020. Nobody reads privacy policies. This senator wants lawmakers to stop pretending we do. *Washington Post*. Accessed November 11, 2020. https://www.washingtonpost.com/technology/2020/06/18/data-privacy-law-sherrod-brown/

Gellman, R., and P. Dixon. 2011. Many failures: A brief history of privacy self-regulation in the United States. *World Privacy Forum*.

Gillibrand, K. E. 2020. *S.3300 - 116th Congress (2019–2020): Data Protection Act of 2020*. Webpage. 2019/2020. February 13.

Grande, D., X. Luna Marti, R. Feuerstein-Simon, R. M. Merchant, D. A. Asch, A. Lewson, and C. C. Cannuscio. 2020. Health policy and privacy challenges associated with digital technology. *JAMA Network Open* 3 (7): e208285. doi:10.1001/jamanetworkopen.2020.8285.

Hart, K. 2019. A new data scandal: How ancestry DNA firms share your most intimate secrets. *Axios*, February 25.

Hazel, J. W., and C. Slobogin. 2018. Who knows what, and when? A survey of the privacy policies proffered by U.S. direct-to-consumer genetic testing companies. *Cornell Journal of Law and Public* Policy (28):35–66.

Hirschman, A. O. 1972. *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*. Cambridge, MA: Harvard University Press.

Hoofnagle, C. J. 2005. *Privacy self regulation: A decade of disappointment*. Washington, DC: Electronic Privacy Research Center.

Jobin, A., M. Ienca, and E. Vayena. 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 1 (9):389–399. doi:10.1038/s42256-019-0088-2.

Kern, R. 2022. Bipartisan draft bill breaks stalemate on federal data privacy negotiations. *POLITICO*.

Kerry, C. F., Morris, J. B., Chin, C. T., and N. E. Turner. 2020. Bridging the gaps: A path forward to federal privacy legislation. *Brookings Institute*.

Khoury, M. J., M. Engelgau, D. A. Chambers, and G. A. Mensah. 2018. Beyond public health genomics: Can big data and predictive analytics deliver precision public health? *Public Health Genomics* 21 (5–6):244–50. doi:10.1159/000501465.

Laestadius, L. I., J. R. Rich, and P. L. Auer. 2017. All your data (effectively) belong to us: Data practices among direct-to-consumer genetic testing firms. *Genetics in Medicine* 19 (5):513–20. doi:10.1038/gim.2016.136.

Lazar, J., D. Goldstein, and A. Taylor. 2015. *Ensuring digital accessibility through process and policy*. Waltham, MA: Morgan Kaufmann.

Libert, T. 2015. Privacy implications of health information seeking on the web. *Communications of the ACM* 58 (3): 68–77. doi:10.1145/2658983.

Listokin, S. 2017. Does industry self-regulation of consumer data privacy work? *IEEE Security & Privacy* 15 (2):92–5. doi:10.1109/MSP.2017.45.

Loukides, M., H. Mason, and D. Patil. 2018. *Ethics and data science*. Sebastopol, CA: O'Reilly Media.

Mantelero, A. 2018. AI and big data: A blueprint for a human rights, social and ethical impact assessment. *Computer Law & Security Review* 34 (4):754–72. doi:10.1016/j.clsr.2018.05.017.

Marotta, V., V. Abhishek, and A. Acquisti. 2019. Online tracking and publishers revenues: An Empirical Analysis presented at the FTC PrivacyCon 2019, June 27.

McCoy, M. S., S. Joffe, and E. J. Emanuel. 2020b. Sharing patient data without exploiting patients. *JAMA* 323 (6): 505–6. doi:10.1001/jama.2019.22354.

McCoy, M. S., T. Libert, D. Buckler, D. T. Grande, and A. B. Friedman. 2020a. Prevalence of third-party tracking on COVID-19–related web pages. *JAMA* 324 (14):1462–4. doi:10.1001/jama.2020.16178.

McGraw, D., and K. D. Mandl. 2021. Privacy protections to encourage use of health-relevant digital data in a learning health system. *NPJ Digital Medicine* 4 (1):2. doi:10.1038/s41746-020-00362-8.

Miller, J. E., Wilenzick, M., Ritcey, M., Ross, J. S., and M. M. Mello. 2017. Measuring clinical trial transparency: An empirical analysis of newly approved drugs and large pharmaceutical companies. *BMJ Open* 7 (12):e017917. doi:10.1136/bmjopen-2017-017917.

Mooney, S. J., and V. Pejaver. 2018. Big data in public health: Terminology, machine learning, and privacy. *Annual Review of Public Health* 39:95–112. doi:10.1146/annurev-publhealth-040617-014208.

Newman, N. 2014. How big data enables economic harm to consumers, especially to low-income and other vulnerable sectors of the population. *Journal of Internet Law* 18 (6): 11–23.

Ornstein, C., and K. Thomas. 2018. Sloan Kettering's Cozy deal with start-up ignites a new uproar. *The New York Times*, September 20, sec. Health.

Pike, E. R. 2020. Defending data: Toward ethical protections and comprehensive data governance. *Emory Law Journal* 69(4): 687.

Pillar, P. R. 2013. Big data, public and private. *Brookings Institution*. Accessed May 25, 2023. https://www.brookings.edu/opinions/big-data-public-and-private/

Post, D. 2021. The Facebook "Oversight" Board. February 17. *The Volokh Conspiracy*. Accessed May 25, 2023. https://reason.com/volokh/2021/02/17/the-facebook-oversight-board/

Price, W. N., and I. G. Cohen. 2019. Privacy in the age of medical big data. *Nature Medicine* 25 (1):37–43. doi:10.1038/s41591-018-0272-7.

Raab, C., and D. Wright. 2012. Surveillance: Extending the limits of privacy impact assessment. In *Privacy impact assessment*, ed. D. Wright and P. De Hert, 363–383. Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-2543-0_17.

Redden, J., J. Brand, and V. Terzieva. August 2020. *Data Harm Record*. Data Justice Lab. Accessed May 25, 2023. https://datajusticelab.org/data-harm-record/

Robbins, R. 2019. Health tech companies introduce guidelines to protect consumers' data privacy. *Stat News*. Accessed May 25, 2023. https://www.statnews.com/2019/09/12/health-tech-companies-introduce-guidelines-protect-consumers-data-privacy/

Sætra, H. S. 2020. Privacy as an aggregate public good. *Technology in Society* 63:101422. doi:10.1016/j.techsoc.2020.101422.

Solove, D. J. 2013. Privacy self-management and the consent dilemma. *Harvard Law Review* 126 (7).

Susser, D. 2019. Notice after notice-and-consent: Why privacy disclosures are valuable even if consent frameworks aren't. *Journal of Information Policy* 9:37–62. doi:10.5325/jinfopoli.9.2019.0037.

Susser, D., B. Roessler, and H. Nissenbaum. 2019. Technology, autonomy, and manipulation. *Internet Policy Review* 8 (2):1–22. doi:10.14763/2019.2.1410.

Valentino-DeVries, J., N. Singer, M. H. Keller, and A. Krolik. 2018. Your apps know where you were last night, and they're not keeping it secret. *The New York Times*, December 10, sec. Business.

Viljoen, S. 2021. A relational theory of data governance. *Yale Law Journal* 131:573–654.

Wachter, S., and B. Mittelstadt. 2019. A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. *Columbia Business Law Review* 2019 (2): 494–620.

Waddell, K. 2020. California's new privacy rights are tough to use, consumer reports study finds. *Consumer Reports*.

Weinberg, G. 2019. What if we all just sold non-creepy advertising? *The New York Times*, June 19, sec. Opinion.

Weiss, M. 2019. Digiday research: Most publishers don't benefit from behavioral ad targeting. *Digiday*, June 5.

Wong, J. C. 2021. How Facebook let fake engagement distort global politics: A whistleblower's account. *The Guardian*, April 12, sec. Technology.

World Health Organization. 2021. *Ethics and governance of artificial intelligence for health*. Geneva: World Health Organization.

Xafis, V., G. O. Schaefer, M. K. Labude, I. Brassington, A. Ballantyne, H. Y. Lim, W. Lipworth, T. Lysaght, C. Stewart, S. Sun, et al. 2019. An ethics framework for big data in health and research. *Asian Bioethics Review* 11 (3):227–54. doi:10.1007/s41649-019-00099-x.