THE OHIO STATE UNIVERSITY

Business Data Ethics: Emerging Trends in the Governance of Advanced Analytics and AI

Final Report
October 15, 2021

Research Team:

Dennis Hirsch, The Ohio State University, Moritz College of Law (Principal Investigator)

Timothy Bartley, Washington University – St. Louis, Department of Sociology

Aravind Chandrasekaran, The Ohio State University, Fisher College of Business

Davon Norris, The Ohio State University, Department of Sociology

Srinivasan Parthasarathy, The Ohio State University, Department of Computer Science

Piers Norris Turner, The Ohio State University, Department of Philosophy

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

A few years ago Amazon, finding itself flooded with employment applications, developed an artificial intelligence ("AI") tool to help it sort through the resumes. (Dastin 2018). It trained the tool on the resumes of its largely male workforce. As a result, the AI tool learned to penalize resumes that used the word "women's," as in "women's tennis team." According to media reports, Amazon's recruiters looked at the tool's recommendations when evaluating new hires, although they never relied entirely on the tool's suggestions. Amazon subsequently spotted the gender bias problem and, unable to fix the AI tool, abandoned the project.

This brief story shows both the promise of advanced analytics and AI (here, sorting resumes quicker than a human could) and the hazards (here, perpetuating gender bias).[1] It also shows the importance of governance mechanisms able to catch defects early and prevent them from hurting people.

In today's algorithmic economy (Schneider 2018), it is vital that companies exercise such governance and so make their use of advanced analytics and AI fairer, more just, and more accountable. Government's limited monitoring and enforcement resources make it unable to perform this vital task on its own. Effective corporate governance is an essential part of the solution.

The question is: how to motivate and achieve such governance? Here, the literature takes two main paths. Many authors focus on what it means for advanced analytics and AI to be "ethical." Scholars (Floridi and Cowls 2019), think tanks and others have generated dozens of sets of ethical principles, and have encouraged businesses and other users of advanced analytics to align their practices with them. (Fjeld, Achten, Hilligoss, Nagy and Srikumar 2020). A second stream of writing proposes new forms of regulation that would require companies to bring their advanced analytics and AI into line with ethical or fairness standards (Balkin 2016; Calo 2014; Citron and Pasquale 2014; Hirsch 2020; Richards and Hartzog 2015; Barocas and Nissembaum 2014;

[1] In recent years, a number of terms have been used to refer to the computational processes that derive insights from massive quantities of data. These include "big data," "machine learning," "data analytics," "advanced analytics," and "artificial intelligence," among others. Each of these terms has a distinct meaning, although the categories overlap considerably. We have decided to use the phrase "advanced analytics and AI" to refer to this suite of technologies. In its information technology glossary,, Gartner defines "advanced analytics" as "the autonomous or semi-autonomous examination of data or content using sophisticated techniques and tools, typically beyond those of traditional business intelligence (BI), to discover deeper insights, make predictions, or generate recommendations. Advanced analytic techniques include those such as data/text mining, machine learning, pattern matching, forecasting, visualization, semantic analysis, sentiment analysis, network and cluster analysis, multivariate statistics, graph analysis, simulation, complex event processing, neural networks." Gartner defines "artificial intelligence" as technology that "applies advanced analysis and logic-based techniques, including machine learning, to interpret events, support and automate decisions, and take actions." https://www.gartner.com/en/information-technology/glossary

Wachter and Mittelstadt 2019). While these two lines of inquiry are important, the literature is missing an essential, third dimension: an understanding of how companies today actually govern their advanced analytics and AI projects. Are these governance practices sufficient? Do they need to be improved and, if so, how? What existing practices can be leveraged and expanded? What gaps need to be filled? Good regulatory design requires an understanding of how companies implement privacy protections "on the ground." (Bamberger and Mulligan 2015; Waldman 2018a). Such knowledge can also help to identify productive practices and disseminate them to others. Yet there has been little empirical study of whether, and how, companies actually go about spotting, and preventing, the harms that their advanced analytics and AI projects can create.

This Report employs a more inductive approach in order to begin to answer these questions. The research team started with testimony and empirical data from practitioners of business data ethics themselves and, from this, sought to construct a picture of what motivates them, how they conceive of ethical decision-making that goes beyond legal compliance (and the limits of such thinking), and the management processes and technological tools that they have utilized to integrate data ethics into their business practices. In doing so, the researchers observed significant tensions between self-reporting and actual practice that informed our critical reconstruction of the companies' data ethics practices and their limits. Despite these tensions, the partial convergence of strategic and ethical considerations in the field of business data ethics is real, and this research started with an attempt to understand how a range of professionals tasked with navigating that convergence have articulated both (1) what constitutes responsible decision-making in the uncertain, beyond compliance domain and (2) the steps they have taken to establish responsible data practices within their companies. The interviewees typically refer to this governance area as "data ethics" or "AI ethics." As they see it, the law lags the rapid emergence of advanced analytics and AI and so, to address the risks that their use of these technologies poses, companies need to go beyond legal requirements and into the realm of "ethics."

The research team included scholars from the fields of law, computer science, business, philosophy and sociology, as each of these disciplines is needed to understand this emerging area. The study sought to answer three, fundamental questions about business data ethics management: (1) How do leading companies conceptualize the threats that their use of advanced analytics and AI pose for individuals, groups and the broader society? (2) If it is true that the law does not yet require companies to reduce these risks, then why are companies pursuing data ethics? (3) How are companies pursuing data ethics? What substantive benchmarks, management processes and technological solutions do they use to achieve this end? The answers to these questions, contained in this report, should give legislators and policymakers a more solid foundation for their work in this area, and provide companies with ideas on how to improve their own data ethics management.

The Report provides only starting place, however. There is disagreement about what one can learn from empirical research, such as that contained in this report, that relies primarily on interviews with chief privacy officers and similarly situated privacy managers. Bamberger and Mulligan's work suggests that the commitments and practices of these privacy managers result in better privacy performance "on the ground." (Bamberger and Mulligan 2015). By contrast, Waldman see a chasm between what the chief privacy officer says or does, and what the engineers and technologists do with respect to privacy (Waldman 2018a). As described in the pages below, the privacy managers interviewed and surveyed for this report said that they and their companies were doing quite a bit to identify and reduce AI risk. A useful area of future research will be to assess the extent to which these corporate data ethics programs affect the behaviors of data scientists and other technologists, and so whether they truly enable companies to spot and reduce the harms their use of advanced analytics and AI can generate.

This Report examines:

- The threats that corporate use of advanced analytics creates for individuals and the broader society (Part III);
- What "data ethics" means to the companies that practice it (Part IV);
- Why companies pursue data ethics when the law does not require them to do so (Part V);
- The substantive principles that companies use to draw the line between ethical and unethical uses of advanced analytics (Part VI);
- The management processes (Part VII) and technologies (Part VIII) that companies use to achieve these substantive goals; and
- Corporate projects that use advanced analytics for the social good (Part IX).

The Report's key findings are:

ADVANCED ANALYTICS AND AI POSE THREATS

Along with its many benefits, corporate use of advanced analytics poses important threats to individuals and the broader society. These include

- *Invasion of privacy*: Companies can use advanced analytics to take seemingly innocuous surface data about people and infer highly sensitive information from it with high levels of accuracy.

- *Manipulation of vulnerabilities*: Data scientists can employ advanced analytics to infer people's vulnerabilities. This can allow bad actors to manipulate, or even exploit, these individuals.

- *Bias against protected classes:* Algorithms and models can disfavor protected classes where protected class status (race, gender, religion, etc.) is expressly included in the data

set, or where facially neutral training data, shaped by past bias, produces algorithms or models that have negative disparate impacts on protected classes.

- *Increased power imbalances*: Businesses can use advanced analytics to achieve highly accurate insights into their customers and so build their advantage over them.

- *Error*: Inaccurate data or faulty algorithms can produce erroneous predictions.

- *Opacity and procedural unfairness*: Most people lack an understanding of and opportunities to challenge the corporate algorithmic determinations that can shape their life opportunities.

- *Displacement of labor*: Advanced analytics facilitates increased automation which, in turn, can displace human labor.

- *Pressure to conform*: Individuals may feel pressure to conform to behaviors that they think will please the algorithmic decision-maker.

- *Intentional, harmful use*: Companies can produce analytic tools that their customers may utilize for morally problematic ends.

The respondents said that their companies were far more cognizant of some of these risks than of others. As Figure 3 (below) shows, 80 percent of survey respondents said their company paid a "great deal of attention" to privacy risks; 50 percent said it paid a great deal of attention to discrimination and unfairness risks; 45 percent said this for transparency and error risks; and only 20 percent said that their company paid a great deal of attention to manipulation, and 10 percent said so for worker displacement risks. The interviews reflected a variation. Thus, companies appear to be actively addressing some threats far more than others.

COMPANIES SEE DATA ETHICS AS BEYOND COMPLIANCE RISK MITIGATION

- Corporate "data ethics" management, as companies themselves describe it, is a form of beyond compliance behavior that seeks to mitigate the risks that a company's use of advanced analytics and AI can create for individuals and the broader society, and so for the company itself.

- The law, including privacy law, lags the rapid development of advanced analytics and AI. As a result, compliance with the law is not sufficient to protect individuals or society from the threats that corporate use of these technologies can create. To protect people against these risks, companies need to do more than the law requires.

- Successful risk mitigation requires attentiveness both to specific ethical risks and to standards of responsible decision-making in the use of advanced analytics and AI. Reputation and public trust are understood to depend on responsible, beyond compliance ethical decision-making.

- Previous literature has described beyond compliance behavior with respect to corporate environmental performance (Gunningham, Kagan, Thornton 2006; Prakash 2011; Bartley 2018). This Report documents beyond compliance behavior with respect to corporate use of advanced analytics and AI.

COMPANIES PURSUE DATA ETHICS TO FURTHER THEIR INTERESTS AND VALUES

If data ethics requires going beyond what the law requires, why are companies pursuing it? Why don't they just focus on legal compliance? The respondents tell a story similar to the one found in the literature on "beyond compliance" behavior in the environmental area (Gunningham, Kagan, Thornton 2006; Esty & Winston 2006). They describe five main drivers:

- *Protect* reputation: Companies worry that if they use advanced analytics and AI in ways that harm people or the broader society, this will damage their reputation with their customers, business partners and regulators. They invest in data ethics to protect reputation and build trust.

- *Prepare for law*: Companies believe that regulation of advanced analytics and AI is coming. They undertake beyond compliance activities to pre-empt, shape and/or prepare for such regulation.

- *Recruit and retain* employees: Employees who perceive the company's data practices as harmful are more likely to leave, or not to accept a job offer in the first place. Beyond compliance behavior can enable companies to recruit and retain these employees.

- *Make better decisions*: Some companies have trouble making decisions about data analytics projects due to the uncertainties surrounding their risks and benefits. Companies that develop standards and processes for assessing the social acceptability of their projects are can make quicker and more effective decisions as to whether to proceed with such projects.

- *Fulfill corporate values*: Some respondents reported that it was their company's or CEO's deeply held values that motivated and informed its data ethics efforts. They see data

ethics management as an extension of a broader commitment to corporate social responsibility.

- *Individual commitment*: Interviewees commonly were those tasked with defining data ethics practices at their companies. They expressed a strong commitment to promoting ethical business practices.

A CORPORATE DATA ETHICS PROGRAM HAS THREE MAIN COMPONENTS

The interviews suggest that, to succeed in their journey towards data ethics, companies need, at a minimum, to do three things: (1) draw the substantive line between ethical, and unethical, analytics and AI practices; (2) manage their operations to ensure that the company stays on the "ethical" side of this line; and (3) develop and implement technologies that facilitate the ethical use of advanced analytics. The Report addresses each of these areas in turn.

SUBSTANTIVE BENCHMARKS DETERMINE WHAT IS ETHICAL

- *Formal principles and frameworks*. Over the past few years, governments, multi-stakeholder groups, companies and others have published dozens of data ethics frameworks that companies can use to distinguish ethical from unethical advanced analytics projects and so accomplish the first of these tasks (Fjeld, et al. 2020). Part IV describes a number of these frameworks and the general principles that they set forth.

- *Informal and intuitive benchmarks.* Surprisingly, most of the companies that we spoke with evaluated data analytics projects according to much more informal and intuitive benchmarks, such as: "Would my mother think this is okay? Would I want this to happen to my kid? Do I feel good about this personally?" (Interviewee #6). Or, how would this look if it appeared on the front page of the newspaper (one interviewee referred to this as the "newspaper test") (Interviewee #21). Some companies have publicly adopted sets of AI objectives or principles. But even these companies appear to revert to more informal standards for actual decision-making. Perhaps the difficulty of articulating a method that can yield a determinate moral judgment or decision, after taking into account various ethical threats and general principles, may push companies toward informal benchmarks. The above-stated finding that that companies engage in data ethics to protect reputation and trusted relationships may also explain this commitment to informal, expectation-based standards for judging what is or is not "ethical."

- *Policies.* Some companies are beginning to capture their data ethics decisions and formulate them into evolving policies that stand between broad principles and informal benchmarks and that the companies can apply prospectively.

- *Risk management.* A small number of companies say they use a risk-management approach to data ethics decision-making in which they identify the a project's benefits and risks to a broad array of stakeholders, consider how the company might mitigate the risks, and then weigh the benefits against the mitigated risks to decide whether to go forward.

- Difficulties in applying general principles or in determining what counts as "acceptable" risk highlight the need for standards of responsible decision-making in the face of moral uncertainty. Beyond compliance data ethics is not only about being sensitive to particular ethical threats and relevant general principles. It also intersects with traditional business ethics debates about corporate responsibility to society and how to establish processes that reflect that responsibility. The fact that companies themselves are often best able to understand the impacts of their use of advanced analytics and AI serves to heighten this responsibility.

Further experimentation and research will be required to determine whether one of these, a combination of them, or some other decision-making approach not mentioned here, is preferable.

MANAGEMENT PROCESSES ACHIEVE SUBSTANTIVE GOALS

It is not enough to draw substantive lines. A company must also manage its operations to ensure that it abides by the lines that it has drawn. Data ethics management innovations break down into three, main areas: organizational structure; issue spotting; and issue resolution.

- *Organizational structure*. The data ethics function goes well beyond privacy to include management of bias, manipulation, opacity, labor displacement and other risks listed above. Some companies house this function in their privacy unit since it has traditionally handled externalities associated with use of personal data. Others locate it in the legal or strategy departments. Companies generally do not see data ethics management as a compliance function.

    - *Data Ethics Officer.* Corporate use of advanced analytics poses threats that go well beyond privacy. This appears to have led to a new type of corporate executive—the Data Ethics Officer – who manages potential harms from all data about humans (not just personally identifiable information) and has expertise in bias, manipulation, opacity and other risk areas that go beyond privacy. Data ethics officers tend to focus on beyond compliance solutions since law does not yet

address sufficiently the threats that advanced analytics can pose. Companies increasingly see this role as a strategic one, focused on building goodwill and trust, rather than as a strictly legal or compliance one.

- *Spotting ethical issues.* Companies used a variety of processes and tools for spotting the ethical issues that their use of advanced analytics and AI could be creating. These include

    - *"Hub and spokes" model.* This places a junior privacy or ethics professional in each business unit. These professionals are trained to spot ethical issues and, where they are significant and difficult to resolve, to elevate them to the central ethics team for further evaluation and resolution.

    - *External advisory group.* Companies have created external groups to review their data analytics projects and identify ethical issues. Such committees, which generally consist privacy advocates, academics, former regulators, and others, have an advisory function but not a governance one. Companies use them to increase their sensitivity to potential ethical issues and gauge the expectations of important stakeholders.

    - *Checklists*. Interviewees reported using AI ethics checklists to spot issues. These tools are at an early stage of development and we did not collect any. A model checklist developed by a team of Microsoft Research and Carnegie Mellon researchers (Madaio, et al., 2020) is currently the best resource for companies or policymakers interested in understanding such checklists. This model consists of questions to consider, actions to take and items to document at six distinct stages in the product development process : (1) Envision (envisioning or greenlighting meetings); (2) Define (spec or design reviews); (3) Prototype (go/no-go discussions and code reviews); (4) Build (ship reviews); (5) Launch; and (6) Evolve (product reviews). The document, which runs almost six pages in length, contains several core themes that are repeated throughout the various stages:
        - Identify demographic groups and others whom the AI system might impact.
        - Examine the fairness-related harms that the AI system might impose and compare them to the system's benefits.
        - Scrutinize and clarify definitions – of system architecture, datasets, and fairness-related harms, criteria and metrics – and revise as necessary.
        - Solicit input from a diverse group of reviewers and stakeholders
        - Where feasible, test the product with these diverse reviewers
        - Monitor product implementation for unanticipated fairness-related harms.
        - Revise the vision, definitions, datasets, fairness criteria, etc.
        - Where harms cannot be mitigated, explore and document why this is the case, future mitigation or contingency plans, and whether it makes sense to proceed.
        - Revise the system at regular intervals to improve its fairness performance

Companies interested in developing their own checklist should consult the Microsoft Research article.

- *Discuss ethical issues regularly.* Regular reflection on and discussion of AI ethics issues can help to build a culture in which employees are more likely to spot such issues. One data ethics manager found it useful to circulate articles and other reports about AI ethics incidents, concepts and solutions, so as to get her associates to think more about them. For a model of how to go about this, consider The Omidyar Network's recently-released a toolkit for sparking such data ethics discussions: The Ethical Explorer Pack.[2]

- *Engage in peer-to-peer discussions*. Interviewees reported the emergence of informal, peer-to-peer, conversations to talk about ethical risks and how to address them.

- *Reaching a decision.* Once a company has spotted an ethical issue, the next step is to make a sound decision about it. Often this involves weighing the project's potential harm against its benefits, or against a different potential harm. Decisions of this type often require judgment.

  - *Data Ethics Committee.* A growing number of companies locate the data ethics decision-making function in a committee that typically includes representatives from legal, privacy, security, communications, data analytics and engineering The committee approach allows the company to see the issue from a variety of perspectives and so to gauge public and regulator expectations better. The data ethics committee often operates by consensus, with all members required to agree before an ethically challenging project can move forward. Some committees have the power to cancel projects or contracts. A recent Accenture and Northeastern University report spells out some of the important design choices that companies should consider when creating a data ethics committee (Sandler and Basl 2019).

TECHNOLOGIES REDUCE POTENTIAL HARMS

Our interviews and survey focused on those who occupy Chief Privacy Officer and similar positions. These managers focus more on substantive and management solutions to data ethics issues than on technological ones. That said, we did learn about some technological solutions to the ethical issues that corporate use of advanced analytics and AI raises.

[2] Available at https://ethicalexplorer.org/ .

- *Data Privacy and anonymization technologies*. Modern technology-based efforts to protect privacy include research efforts on k-anonymity, l-diversity and epsilon-differential privacy (or differential privacy for short). The interviews make clear that the last of these, differential privacy, is the current de-facto standard for privacy preserving data analysis. Several interviewees also discussed simpler rule-based aggregation strategies for anonymization and de-identification aligned with the classical notions of k-anonymity. Others note that privacy and ethics are overlapping, but not synonymous. They point out that organizations that protect privacy can still utilize data in ways that are ethically problematic.

- *Algorithmic fairness technologies.* Several interviewees brought up the importance of technological solutions to facilitate fair artificial intelligence (AI) and Machine Learning (ML). For example, companies need to make an effort to identify and use datasets that are both inclusive of marginalized groups (so that the resulting AI does not treat them less accurately or less well) and, at the same time, are not themselves shaped by harmful societal bias.

- *Making algorithms more explainable.* Several interviewees emphasized the importance of explainability to facilitate procedural fairness and trust and to ensure that regulatory policies are being met. Others pointed out that companies may have to go beyond simple explainability and develop explanations as to why models fail, and the risks associated with such failure.

- *Auditing of algorithms.* Algorithms are complex multi-step processes that can produce several sources of risk at each step. To mitigate such risks, companies need to be able to audit the algorithms to ensure that they are consistent with stated data ethics policies and regulatory requirements. Organizations are beginning to work on this type of auditing of algorithms.

- *Systems Technologies.* Several interviewees discussed the development of systems for enabling or enhancing ethical governance. Some use insulated data warehouses and data lakes, or virtual data, for this purpose.

COMPANIES CAN USE AI FOR THE SOCIAL GOOD

As companies grapple with the need to go "beyond compliance" in their data ethics practices, some of them have welcomed opportunities to promote the social good, either by directly trying to improve individual lives (Facebook's suicide prevention initiative is a controversial example) or by producing information that municipalities can use to improve evacuation planning during natural disasters, track infectious diseases, or relieve traffic congestion or for other purposes. These projects do not have a direct impact on the company's bottom line.

As companies enter the world of data ethics, they will encounter many opportunities to use their powerful data analytics tools to benefit their communities. Whether most companies will integrate moral values and the broader interests of the public into their decision-making for their own sake, and not because of the coincidence of morality and interest, remains an open question

Acknowledgments

The research team, each of whom is a Fellow of The Risk Institute at The Ohio State University Fisher College of Business, would like to acknowledge financial support from The Risk Institute and from Facebook, as well as administrative support from The Ohio State University Moritz College of Law and Translational Data Analytics Institute. Special thanks to Christina Drummond and to Gillian Thomson for their assistance with this project, and to the Privacy Law Scholar's Conference for giving us an opportunity to workshop of an early draft of this report.

I. BACKGROUND

A growing proportion of human activity is being captured as data, monetized, or otherwise used by for-profit firms, as well as by governmental and non-profit organizations. As analytic capacities have skyrocketed, companies have moved into new frontiers of forecasting, complex decision-making, and tracing of geographical and virtual movements. According to some lines of research, companies that have capitalized on what IBM has called the four Vs of big data—volume, velocity, variety, and veracity (Herschel and Miori 2017)—have seen benefits in operational efficiency, market valuation, profitability, consumer satisfaction, and decision-making (Faroukhi et al. 2020; Fosso Wamba et al. 2015; Lavalle et al. 2011; McAfee and Brynjolfsson 2012). Corporate use of big data analytics may also have broader social benefits related to public health, education, transit, and policy-making (Glaeser, Kim, and Luca 2017; Khoury and Ioannidis 2014; Marx 2013; Nuaimi et al. 2015).

At the same time, it is increasingly clear to companies, policymakers, and the public that big data analytics carries significant socio-political risks and ethical conundrums. In addition to longstanding debates about privacy, the use of complex predictive analytics and algorithms raises serious questions about bias and discrimination (through racially biased facial recognition or "digital redlining," for instance), manipulation (through micro-targeting to exploit vulnerable individuals), and procedural unfairness (due to opaque, non-transparent algorithms). By some accounts, companies are forging ahead with a style of "surveillance capitalism" that extracts an ever-growing amount of data from human experience in order to track, predict, and ultimately reshape behavior for commercial purposes (Zuboff 2019). Even classic questions about privacy are being reconsidered as complex data supply chains (Ebeling 2016) and data analytics capacities that exceed human understanding challenge "notice and consent" systems, and data that is so

fine-grained as to be re-identifiable stymies anonymization (Ohm 2010; Tene and Polonetsky 2012). The lack of a sufficiently protective framework allows advanced analytics to have privacy-threatening consequences, as with Target's mailing of personalized coupons to customers predicted to be pregnant, a marketing campaign that revealed a 15-year-old girl's pregnancy to her family. In addition, beyond concerns about discrimination, there are worries that big data analytics will indirectly exacerbate inequality, whether through the "off label" use of credit scores and zip codes or through the displacement of workers through automation (Barocas and Selbst 2014; Fourcade and Healy 2017; O'Neil 2016; Rona-tas 2017).

Scholars from a variety of fields are increasingly grappling with the benefits and harms of big data analytics. These writings have most often focused on either the legal and regulatory terrain, such as the applicability of existing law or the design of new regulatory approaches, or on the normative principles that might guide ethical practice in this arena, from industry codes to philosophical frameworks. Our research takes a different approach, seeking to examine the policies and procedures that companies have begun to put into place to manage the ethical dilemmas of big data analytics. This overlaps to some degree with inquiries into ethical norms, law, and regulation, but it provides a distinct point of entry, focused on how companies in the U.S. are interpreting and acting upon the evolving legal and regulatory terrain, as well as the scrutiny they face from media, advocacy organizations, employees, and customers.

We focus our attention on firms' policies and procedures for two primary reasons. First, the ethics of big data analytics represents the next frontier of corporate responsibility, self-regulation, and industry benchmarking. Even if governments develop robust legal and regulatory regimes, there is little doubt that firms will play a central role in managing and mitigating risks, whether to ensure their legal/regulatory compliance or to go beyond what is legally required in order to build their reputation or avoid disruption to their operations and markets. In today's

data-intensive environment, many companies are no longer deciding what they can or cannot do with data analytics but instead grappling with what they *should* and *should not* do. While our study cannot specify what firms should be doing, it can provide insight into what some companies are doing, why they are doing it, and what kinds of challenges they have faced. While the last two decades of corporate responsibility have revolved around globalization, environmental sustainability, working conditions, and human rights, data ethics are likely to be at the center of the conversation in the coming decade.

Second, while some research has examined the practices and positions that firms put into place to address privacy (Hirsch 2011; Mulligan and Bamberger 2015; Waldman 2018a), this mostly predated the rise of risks associated with prediction, automated decision-making, and artificial intelligence, which are not limited to privacy. As we will see, our research suggests that policies and positions put in place to address privacy—such as the establishment of Chief Privacy Officers—continue to structure many firms' approaches to big data analytics more broadly. This may suggest that firms have a solid foundation to build on, or that many are still catching up with the data privacy concerns of the past two decades, while simultaneously navigating a new set of ethical dilemmas. The extent to which companies are structured to deal with the risks of big data analytics likely also extends beyond simply having positions or policies in place, but also about how well the concerns of the risks of big data are integrated into the activities of employees on the ground (Waldman 2018a).

Overall, big data analytics is at an important crossroads. Corporations, governments, and society writ large want to obtain the promise of big data while minimizing the risks. But they do not yet know how to do this. In the sections that follow, we first briefly review research on the legal/regulatory terrain and emerging normative principles that often serve as ways to balance

the benefits and risks of big data before turning to our own research on corporate policies and practices.

A. LEGAL CONCEPTS AND REGULATION

As mentioned, big data analytic capabilities have a propensity to infringe on due process norms, undermine trust, and antiquate cornerstone legal concepts and regulation (Cate and Mayer-Schonberger 2013; Tene and Polonetsky 2012). To address these shortcomings, a line of research has emerged to revise foundational legal terms for a digital age. Some advocate for a "technological due process" or "procedural data due process" to modernize procedural standards with increased transparency and pragmatic improvements like training judges on how to assess the reliability of expert testimony involving big data (Citron 2016; Citron and Pasquale 2014; Crawford and Schultz 2014). Others argue for extending fiduciary responsibilities to corporations handling data to align their interests with consumers and ultimately foster trust (Balkin 2016; Richards and Hartzog 2015; Waldman 2018b), or for using commercial unfairness law to set parameters for the fair use of predictive analytics (Hirsch 2020). Likewise, scholars articulate similar revisions for a variety of other terms bringing concepts like autonomy, manipulation, fairness, accountability, and transparency up to date in the era of big data (Calo 2014; Felzmann et al. 2019; Hellman 2020; Kroll et al. 2017; Selbst and Barocas 2018; Wachter, Mittelstadt, and Russell 2018).

Updating legal concepts helps identify ways current regulatory authority may be deployed to regulate big data and serves as a foundation for new omnibus legislation. The Federal Trade Commission, in particular, has received attention as some argue its authority to regulate unfair trade practices may be expanded to encompass issues related to big data analytics (Hartzog and Solove 2015; Hirsch 2015; Hirsch 2020). With respect to new legislation, the European Union's General Data Protection Regulation (GDPR), implemented in May 2018, represents a significant

overhaul in European privacy law replacing previous privacy guidelines from the 1990s (Mayer-Schonberger and Padova 2016). In the US, GDPR was quickly followed by the California Consumer Privacy Act (CCPA) which became effective January 2020. Additionally, two bills have been introduced in the U.S. Congress, the Algorithmic Accountability Act (H.R.2231) and Consumer Online Privacy Rights Act (S.2968). These bills seek to require corporations to study and fix biased or unfair algorithms and to comprehensively alter federal online privacy regulation. Importantly, the efficacy of these legislative efforts has yet to be seen as the GDPR and the CCPA are relatively new, and the Algorithmic Accountability Act and Consumer Online Privacy Rights Act are only in their infancy as bills. In fact, the rapid development and expansion of capabilities have led some to become pessimistic about the ability of any legislation to fully address the concerns of big data (Mayer-Schonberger and Padova 2016; Zarsky 2017). As a result, a second stream of scholarship moves beyond regulatory proposals to understand what ethical principles big data analytics should reflect.

B. NORMATIVE PRINCIPLES

Scholars representing a range of fields from philosophers to computer scientists in addition to corporations, professional associations, and standard-setting bodies have begun to delineate guiding normative principles they argue should be reflected in corporate uses of big data (Drosou et al. 2017; Herschel and Miori 2017; Jobin, Ienca, and Vayena 2019; Mcdermott 2017; Mittelstadt et al. 2016; Richards and King 2014; Yang et al. 2018; Zwitter 2014). In their review of the global landscape, Jobin et al. (2019) identified approximately 80 different frameworks. Likewise, Fjeld and colleagues (2020) survey over 30 frameworks put forth by a diverse set of institutions. Though each framework outlines several principles, all valuable in their own right, there has been a convergence on a familiar core set of ideas. Floridi and Cowls (2019) exemplify this convergence as they condense the multitude of considerations down to five elements: beneficence (promoting

well-being and preserving dignity), non-maleficence (ensuring privacy and security), autonomy (avoiding manipulation), justice (preventing unfairness), and explicability (enabling transparency and accountability).

While principles provide a roadmap for what firms should consider, providing a roadmap and having users use that roadmap are distinct things (Mittelstadt 2019). Likewise, the abstract nature of principles often precludes concrete implementation (Whittlestone et al. 2019). For example, how should a corporation "preserve dignity" in its operations? This ambiguity and disconnection from implementation leave a gap and so serve as our entrée in the study of big data.

C. WHAT IS MISSING?

If corporations lack clear legal foundations to understand how they *should* self-regulate their big data capabilities because law is quickly outdated, and normative principles lack enough specificity to be implemented, then how do corporations address the risks of big data? Prior studies have occluded the question of how corporations negotiate a patchwork of legislation and complex landscape of social norms/expectations in their governance of emerging advanced analytics techniques. Our goal with this study is to begin to fill this gap by focusing on several descriptive questions. What benefits and risks are top of mind for corporations with respect to their use of big data, advanced analytics and AI? How well do corporations believe legislation addresses these risks? What policies and processes do corporations have in place to address these risks? How well integrated and efficacious do managers believe their policies and processes are? Gaining insight into these questions provides a novel extension of academic literature that is experiencing a paucity of empirical investigations (Flyverbom, Deibert, and Matten 2019) and contributes to broader conversations among scholars, policymakers and practitioners about how to balance the possibilities and the pitfalls of big data.

II. METHODOLOGY

We employ a mixed-methods research design, including in-depth interviews and a survey of managers within corporations. The interview component served as an open-ended way to map the terrain of the contestation around big data ethics and inform the construction of a meaningful survey instrument. The survey component attempts to then synthesize insights from the interviews to understand how to systematically assess corporate uses of big data, the risks, and specific policies and processes. We treat the research components as complementary, collectively contributing unique dimensions to the first empirical investigation (to our knowledge) of corporate practices for addressing the risks of big data analytics. Targeting higher-level executives for our interviews and survey gives us the view of corporate practices from the top, but precludes us from assessing the coupling between high-level policies and the actual daily work of engineers and employees on the ground (Waldman 2018). Likewise, our survey sampling methodology discussed more below lends itself to selection bias as many of our respondents are a part of trade associations that are specifically engaged with questions of data and privacy.

A. INTERVIEWS

We conducted in-depth interviews targeted at Chief Privacy Officers and relevant individuals within U.S.-based companies, individuals from professional services firms, and representatives of consumer organizations. We used a purposive sampling method in which we first identified individuals prominently engaged in writing about big data ethics by publicly leveraging social networks of the research team (Singleton and Straits 2010). We then snowball sampled by asking interviewees to identify individuals involved in public discussions/forums on big data ethics or were actively grappling with big data ethics in their position. The snowball

method facilitated access to new interviewees which was important as reaching high-level managers is particularly challenging (Biernacki and Waldorf 1981; Cycyota and Harrison 2002).

The interview protocol was developed to broadly probe respondents about corporate big data ethics. The protocol had four major sections: (1) risks of big data, (2) motivations and goals of mitigating the risks of big data, (3) management processes including substantive frameworks and technological solutions, and (4) the broader regulatory and legal environment. The protocol was adjusted to account for differences between companies, law firms, advocacy organizations, and consulting firms. When interviewing representatives of corporations that used big data analytics, our interviews probed respondents on the structure, design, and perceived efficacy of internal processes. Interviews were conducted primarily over the phone with one interview conducted in person. We conducted interviews from September 2017 to March 2019 and interviews ranged from 60 mins to 160 mins with an average of 75 minutes. We transcribed interviews to facilitate coding and analysis which involved descriptive coding according to the sections of interviews followed by close readings to identify prevalent themes across interviews.

Overall, we interviewed 23 respondents. The industries represented in the interview sample range from telecommunications, information technology, social media, pharmaceuticals, and insurance. Both publicly traded and private companies are represented in the sample. The interviewee's titles included, at various levels of seniority: Privacy Officer, Data Ethics Officer, Counsel, Public Policy Executive, Compliance Executive, and Partner.

B. SURVEY

We paired the interview study with a survey using the Qualtrics online platform. As with the interview component, we targeted higher-level management. Contacting this specific population is notoriously difficult using a large random survey (Cycyota and Harrison 2002).

Survey research of corporate management has indicated that an important way to increase response rates is to have the survey delivered through legitimated or trusted organizations. As a result, we opted for a convenience sampling approach that leveraged the social networks of corporations through membership in industry trade associations and industry-funded think tanks. Specifically, we contacted five trade organizations and think tanks engaged with issues of data and privacy and asked them to send our survey as a proxy to their member companies. In this way, we conceptualize our survey results as representing a more optimistic view of current corporate practices as our targeted sample have selected into membership in organizations engaged with privacy and data accountability.

We provided trade organizations with email language/script and survey links for their members ensuring that companies that overlapped in membership with multiple organizations only received one survey link. The think tanks and trade organizations agreed to send reminder emails one-week after the initial survey was sent. Data was collected from November 2019 to January 2020.

In total, our survey was sent to 246 companies. We received a total of 51 responses with 24 fully completed yielding a response rate of approximately 20% for all surveys and approximately 10% for fully completed surveys. This response rate is fairly consistent with the range of other surveys of corporate managers (Cycyota and Harrison 2006). Given our targeted sampling strategy and exploratory nature of the study, we are unable to make strong claims; however, we can identify cleavages of variation and associations which will serve as an important entry point for future research. In particular, our findings from this targeted survey provide evidence that a much larger sampling of corporate big data ethics is necessary and would likely yield valuable insights for both scholars, policymakers and companies.

Results we discuss from the survey focus on what refer to as our "core sample" of 31 respondents that answered question 9 of our survey which relates to the policies a respondent's company has in place to address the risks of big data analytics. Figures 1 and 2 display the size variation of the sample by number of employees and revenue, respectively. We expected that our sample would be comprised of larger companies on average given the membership of the organizations we sampled through, and this is the case. The largest proportion of the sample, approximately 30%, are the largest companies with more than 40,000 employees or more than $15 billion in revenue. While skewed towards large companies, almost 50% of the sample has less than 10,000 employees. As it relates to industries, Table 1 shows that most of our respondents were information technology companies. The remaining companies were evenly split amongst financial services firms, communications, and industrials with healthcare composing the smallest proportion of the core sample.

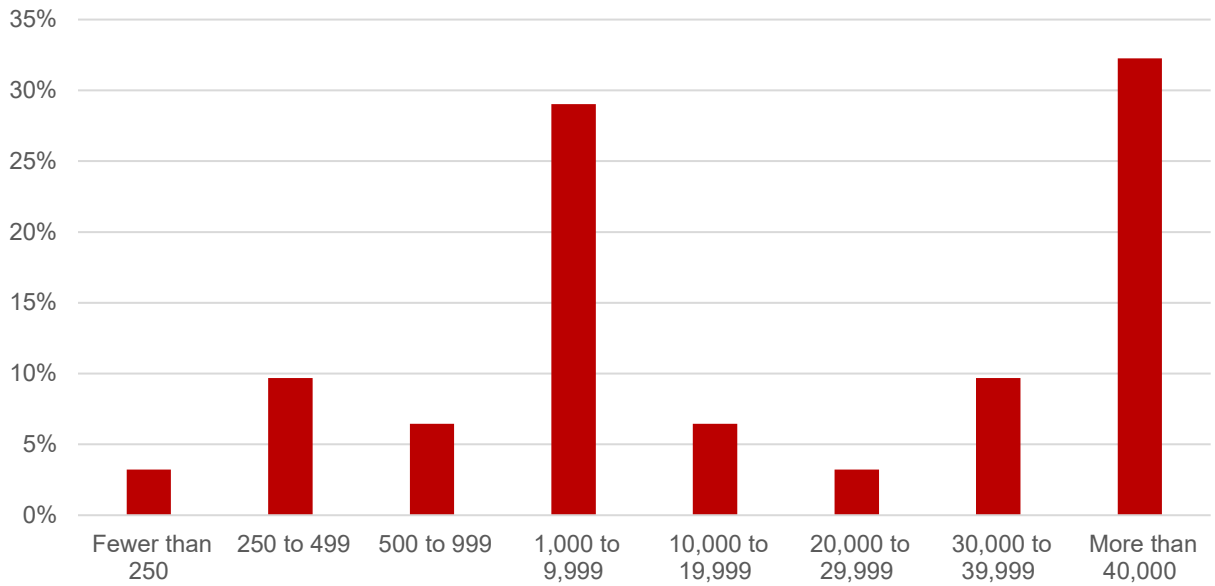

Figure 1. Distribution of sample company size according to total number of employees

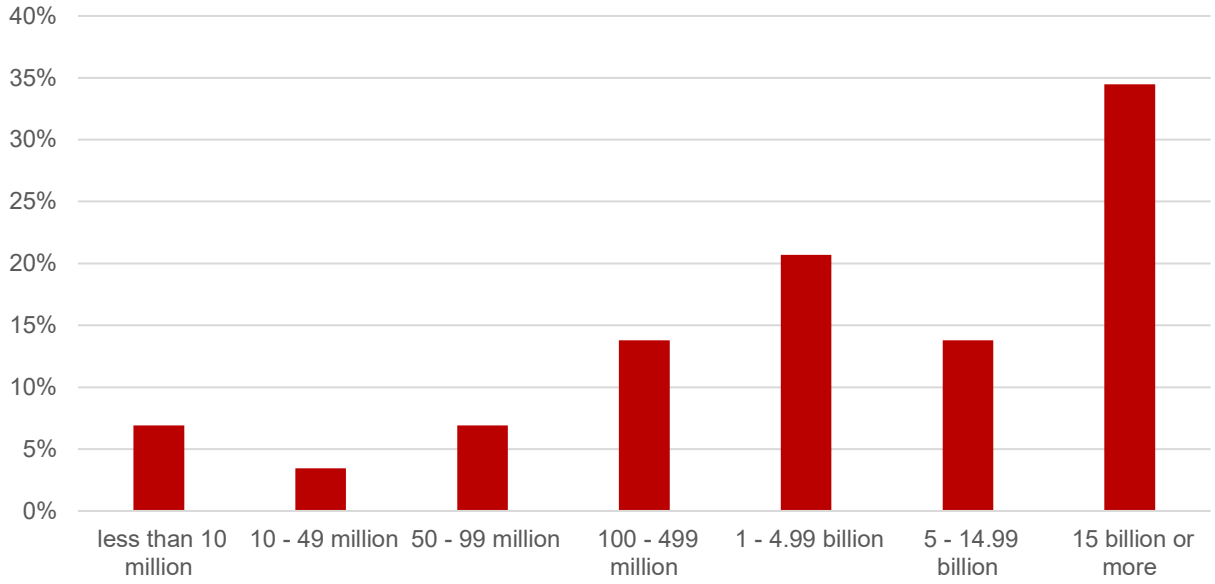

Figure 2. Distribution of sample company size according to 2018 revenue

Table 1: Survey respondent industry

	Percent
Communications services (including Telecommunication Services and Media & Entertainment, including Advertising)	16.1
Information Technology (including Software & Services, Technology Hardware & Equipment, and Semiconductors & Semiconductor Equipment)	41.9
Financials (including Banks, Diversified Financials, and Insurance)	16.1
Health Care (including Health Care Equipment & Services and Pharmaceuticals, Biotechnology & Life Sciences)	9.7
Industrials (including Commercial & Professional Services, Human Resource & Employment Services, Office Services, Capital Goods, and Transportation)	16.1

III. RISKS FROM ADVANCED ANALYTICS

The great majority of the survey respondents, and each of the corporate interviewees, said that their company's use of advanced analytics poses risks to individuals and/or the broader society, and that their company is taking steps to address these risks. The survey and interview components of the research project each utilized small samples shaped by selection bias. They accordingly tell us only what the particular respondents say about their specific companies. Still, their consistency on the question of risks is striking. Virtually all of the respondents – both in the survey sample, and in the interview sample – confirmed that their companies' use of advanced analytics does pose risks, and that the companies themselves are aware of these potential harms. Most companies that talk publicly about their use of advanced analytics focus on the valuable insights that it produces. The survey and interviews suggest that this rosy view is only part of the picture and that some of the more sophisticated companies, at least, also see and are paying attention to the very real risks that these activities create.

All survey respondents stated that their company recognized its use of advanced analytics created important risks. As illustrated in Figure 3, the companies paid more attention to some risks than to others. Eighty percent of respondents said that their company pays a great deal of attention to the privacy risks that advanced analytics creates; fifty percent said that it pays significant attention to both discrimination and unfairness risks; and almost half that their company pays attention to transparency risks and error risks. A smaller percentage of respondents indicated that their company paid significant attention to manipulation and worker displacement risks.

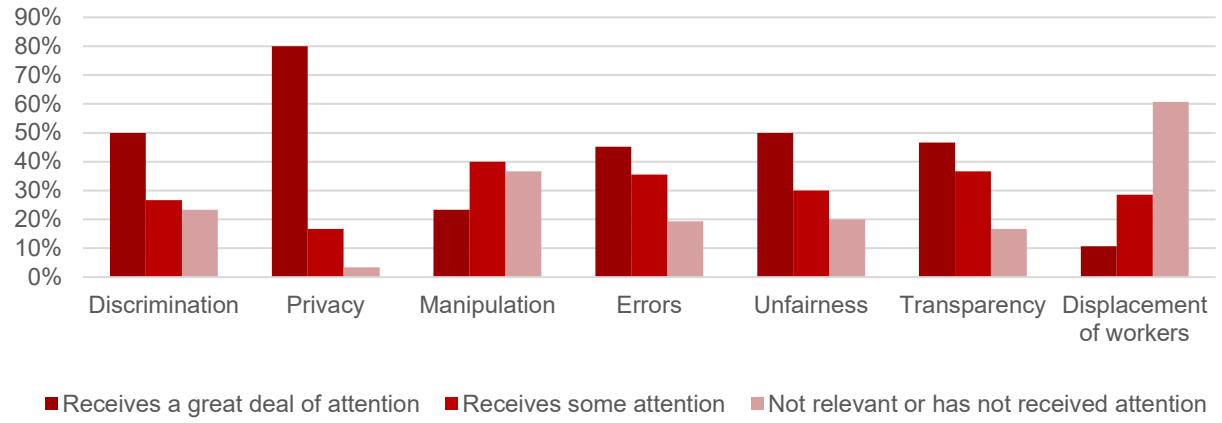

Figure 3. Corporate attention to risks from advanced analytics

The interviewees identified an even broader array of risks than the survey respondents. The remainder of this section sets out the risks as the interviewees described them.

A. PRIVACY VIOLATIONS

Companies can use advanced analytics to take seemingly innocuous surface data about people and infer highly sensitive information from it with high levels of accuracy. For example, researchers at Cambridge University were able to take a person's Facebook likes and infer their gender, sexuality, age, race, and political affiliation "with remarkable accuracy" based solely on these surface data ((Rosen 2013). Predictive analytics thus poses a profound threat to personal privacy ((Rubinstein 2013). The interviewees expressed keen awareness of this threat to privacy. As one remarked: "You're learning my weaknesses or learning my pregnancy status, you're learning whether I'm gay, you're learning intimate information about what I do in my home. Is it ethical for you to be doing that even though your policy said you do research and we collect that information for product improvement?" (Interviewee #19).

The interviewees distinguished between different types or levels of privacy invasion. Some predictive insights feel "creepy," such as when Facebook inferred which of its users were

Jewish and sent Rosh Hashanah (Jewish New Year) greetings to them (Interviewee # 23). Other insights are more invasive and can cause severe embarrassment, distress or even danger. For example, some gay teens have been outed to their parents as a result of their receiving gay-themed advertising from companies that inferred their sexual orientation (Interviewee #23). Finally, companies may deny people opportunities for jobs, loans or other important life opportunities based on predictive insights about their physical or mental health status, sexual orientation or other highly personal attributes. One interviewee gave the example of a producer of smart toothbrushes that faced economic pressure to sell household tooth brushing data to insurance companies who could infer risk of heart disease from it. "This might go to future insurability of the kids or payment for pre-existing condition of the adults. My point is, the world is changing and measurement or observation of us, which is happening, this is the way it all works, is very, very important. We've got to decide what the rules are now. Right?" (Interviewee #6).

B. MANIPULATION

Data scientists can employ advanced analytics to infer people's vulnerabilities. This can allow bad actors to manipulate, or even exploit, these individuals. For example, a business might predict that an individual is likely to experience early stage dementia and target the person with predatory loans intended to take advantage of her diminished, but undiagnosed, mental state. Or, as actually happened, a company such as Cambridge Analytica might take people's Facebook "likes," use them to infer their personality types, and then target them with political advertisements that appeal to their unconscious in ways that they find hard to resist (Rosenberg 2018).[3] One interviewee saw such manipulation as a growing issue:

[3] An interviewee explained how this issue can arise on the smart retail environment where sensors, including those placed in the mirrors at beauty counters, can infer from shoppers' facial expressions their likelihood

> *[P]eople are becoming more sensitive to some of the risks that I might put into the category of being unfairly manipulative, or kind of unfair in some way. That might be using predictive analytics to sell people things they don't need, or can't really afford. Or, targeting people based on vulnerabilities, whether it's age, or cognitive abilities, or some other disability…. When you see them happening . . . people recoil as slimy and nobody wants to be that… When do those lines get crossed? So that's not always obvious. There's certainly a sensitivity around that (*Interviewee #12).

The difficult questions lie in identifying the point at which marketing becomes unacceptable manipulation or even exploitation. An interviewee from the retail industry talked about how they approach this issue:

> *How much imputation can you do before you're actually manipulating and defining the behavior and causing the behavior, rather than responding to it? . . . I've said this a lot to our marketing teams, I was like "So long as you are persuading." In your gut, [if you] know that you are persuading and providing an offer, and something of value -- you're good. The moment you feel that you are manipulating, you've gone too far and we need to have a conversation (*Interviewee #17).

This gut-level, know-it-when-you-see-it approach to drawing the line between marketing and unacceptable manipulation leaves a great deal of room for interpretation.

C. BIAS AGAINST PROTECTED CLASSES

The law distinguishes between disparate *treatment* and disparate *impact* discrimination. Disparate treatment occurs when one intentionally disadvantages another based on a protected characteristic (e.g. race, or gender). Disparate impact occurs when a facially-neutral standard unintentionally, but meaningfully, disadvantages those who possess a protected characteristic. Predictive analytics can produce disparate treatment. For example, a company could infer

---

of buying particular good. The store can then market that good to the individual in real-time in the store. "I'm not saying that we're there, I'm not saying that anybody is necessarily there. But I think that's where we're going." Interviewee #16.

someone's protected characteristic (e.g. pregnancy), and intentionally discriminate against the person on this basis.[4]

It can also produce disparate impact discrimination when past bias has shaped the training data. For example, Amazon tried to develop an AI tool that could separate viable from non-viable resumes (Dastin 2018). It trained the tool on the resumes of existing Amazon employees, most of whom, due to pre-existing bias in the technology field, were male. The tool accordingly learned to reject applicants whose resumes identified them as female (e.g. by listing an all-women's college). Amazon discovered this problem early on and ultimately abandoned the project. But bias in the training data can be subtle and many companies may miss it.

Harmful bias seemed to be one of the top, if not the top, concern of the interviewees:

> *Algorithmic discrimination is a top tier issue for me and my group, and I've made it a priority. What I mean by that is to work, and help, and focus, our engineering teams on evaluating outcomes as we build out especially our machine learning portfolio. You're never going to be able to be 100 percent positive, in a testing environment, that your algorithm isn't creating some disparate impact. That's very difficult to do . . . How do you get data that doesn't have a lot of bias in it? That's also tricky, but there's some data sets that we all know to have tremendous bias in it, so maybe steering away from those insofar as you're training the models might be helpful, right?* (Interviewee #18).

Increasingly, companies seek to address the problem of algorithmic bias by identifying, and either modifying or not using, biased data sets. This is an important strategy. The ethical question that the interviewees posed was how far to go with this. Specifically, do companies have an obligation to "fix" long-standing social inequalities that are accurately reflected in the training data? For example, should a facial recognition tool that learned it could identify gender by whether the person was standing in a kitchen (women were more likely to be in the kitchen)

[4] Such a practice would likely violate employment discrimination laws, *see e.g.,* The Pregnancy Discrimination Act of 1978, amending Title VII of the Civil Rights Act of 1964, 42 U.S.C. §§ 2000e (prohibiting an employer from discriminating against an employee or applicant based on pregnancy status), but would be very hard to detect.

deliberately ignore this finding? (Interviewee #19). If women, through their online behavior, express less interest in certain high-paid jobs than men, should the company nonetheless advertise the jobs equally to both women and men? (Interviewee #19). Should the company ignore or alter the training data in these cases, or modify the conclusions that emerged from it? The interviewees talked with their data scientists about this question. As one explained, "[t]he concern is now you've taught this thing, this code, to be biased. On the other hand, do they have some obligation to have the algorithm be less accurate . . . do you want me to pull [those data that are the product of bias] out? So these sorts of questions are being asked of us by the AI folks: 'We'll figure out how to do or not do . . . but tell us when and where prediction is discriminatory in a way that is to be deterred.'" (Interviewee #19).

Another grey area was when, if ever, it is acceptable to use a protected characteristic in algorithmic decision-making. For example, when data shows that different racial or ethnic groups have different preferences, is it appropriate to take this into account in marketing to the members of these groups? An interviewee from the retail industry provided an example:

> *so we know that there's different body sizes, or different body types perhaps, for different ethnicities. You might need wider-thighed jeans. We're conscious of that. And again, this is just matching the customer with what they need. So in that case, if we have a special on jeans and we want to make sure they're the right jeans, that ethnic code might actually be important (*Interviewee #17).

These interviews raise deep and interesting questions about what a society that values equality and justice should look like, and about what companies should do to try to achieve that vision. They suggest that at least some corporate privacy professionals and data scientists are discussing these issues but are doing so without the benefit of well-developed tools, resources or ethical frameworks that could help them navigate the grey areas.

D. INCREASED POWER IMBALANCES

Businesses that employ advanced analytics to achieve highly accurate insights into their customers can use this to build an advantage over them. For example, a company could infer the highest price that each customer would be willing to pay for a given good or service, and then charge the individual that price. This would allow the company to capture all the gains from trade. On another front, corporate use of advanced analytics to determine eligibility for loans, jobs or other important opportunities can entrench existing inequalities. If more privileged groups tend to possess more frequently the attributes (proxies) that predict success in these areas, the algorithm will more likely select them for these opportunities. This can reproduce existing hierarchies and further lock the poor into their poverty. Additionally, analytics can also enable companies to segment groups into much finer categories than was previously possible. This can have social and distributional effects. For example, it can undermine the pooling of risk that has long been one of the social functions of insurance.

E. ERROR

Inaccurate data or faulty algorithms can produce erroneous predictions. In the marketing area, such errors can result in annoyed or dissatisfied customers ((Interviewee #12). In the government context, the stakes can be much higher. As one interviewee recounted: "Our number one risk is if someone is killed because of our analytics. We're working with the military, we're working with intelligence and law enforcement, and I've impressed this on the engineers a number of times, you're pointing a loaded gun at someone basically. Are we 100% confident in the analysis that we're supporting here, and if we're not, then the consequences are that level of seriousness." (Interviewee #10).

F. OPACITY AND PROCEDURAL UNFAIRNESS

What really distinguishes algorithmic decision-making from its human counterpart is not its capacity for error (humans make them too), but rather its opacity and imperviousness to challenge. For example, where a company determines through advanced analytics that an employee would not succeed in a higher position and denies the person a promotion, the employee would have no way to know what data or algorithm had resulted in this determination, and no way to challenge them ((Rubinstein 2013). Such algorithmic determinations are a "black box" as far as the individual is concerned. (Pasquale 2016). In some advanced machine learning, even the company or other decision-maker may not understand how the technology arrived at its determination. The risk to the individual, then, is that machine-driven decisions deny people the core procedural rights—transparency and the right to be heard—to which they are entitled when others are making important decisions about their lives. One interviewee articulated this risk:

> *In this case, if a harm occurs, there is no mechanism to even understand why suddenly am I on the No Fly List. . . .How did I get on the No Fly List? There is no mechanism to ask. You will be told, "[it's] none of your business, you simply can't fly anymore." . . . What if [the list placement] was because in the third generation of processing, where they were not using data about me but data inferred about me, something got in there that was a horrible inaccuracy or trigger and now it is perpetuated because suddenly, it's no longer about the data about me, it's about data that has been inferred about me. Some risk score. And there is no mechanism to actually understand why [it happened] or to have [the data] corrected (*Interviewee #21).

G. DISPLACEMENT OF LABOR

Advanced analytics facilitates increased automation which, in turn, can displace existing labor forces. As one interviewee explained:

> *The thing that worries me enormously in this way is driverless cars. . . . You're going to put people out of work: trucking, cab drivers, low skill workers, people who aren't going to be able to get other jobs and I don't think the industry thinks*

*it has to care about that. The speed at which it's developing these things, if it builds a driverless car that works really well and starts replacing everybody before society is able to figure out what are we going to do with all these people that it's displaced... that's hugely irresponsible. That's the kind of thing that topples governments, leads to the French Revolution, you know? This is significant, and I don't think industry really takes responsibility for that . . . And what's the legal solution? Ban driverless cars? Maybe, but that's a hard call. What's the rationale for that? I think these are the huge challenges that engineers have to own; but, I'm not sure they know they should (*Interviewee #10).

H. PRESSURE TO CONFORM

One interviewee expressed a deep fear that constant data collection about people, combined with analytic use of that data to allocate goods and opportunities, would create a profound pressure on individuals to conform to behaviors that they think will please the algorithmic decision-maker.

*[M]y biggest fear, which is almost Orwellian, is that . . . [a]t some point, we as individuals will begin to realize that we are being observed. And everything about our behaviors and our patterns of behaviors are being understood, compiled, inferences are being created. And there will be a point in time in the near future . . . where we're going to internalize that. And you know what's going to happen ... we are going to be the person we think people want us to be all the time. And what impact is that going to have on creativity? What impact is that going to have in ultimately funneling us all down into behavior that we believe or, worse case, know that we must conform to? . . . What impact is that going to have on society? On culture? On us as individuals? It scares the hell out of me. And it's happening right now (*Interviewee #21).

I. INTENTIONAL, HARMFUL USE OF ANALYTICS

Some companies worried that customers or others would use their analytic tools for morally problematic ends. For example, one company had an internal debate about whether to sell its technology to customers that may have ties to the Chinese government, since the Chinese government might use the technology to create facial recognition tools capable of distinguishing

members of the Uighur minority (Interviewee #2).[5] In a well-publicized 2018 incident, thousands of Google employees signed a letter protesting the company's work on a Pentagon pilot program, Project Maven, which used machine learning to interpret drone imagery and, potentially, to better target drone strikes against suspected terrorists or other individuals (Wakabayashi and Shane 2018). The letter expressed the employees' view that "Google should not be in the business of war." A few months later, Google announced that it would cease its involvement with the controversial Pentagon program (Harwell 2018).

The difficult question is where to draw the line. One interviewee described an employee complaint about the company's analytic work for a cosmetics manufacturer. "[S]omebody sent an email to me and they said 'What good does it do the world to perpetuate working with companies whose primary mission is to make women feel bad about how they look?' I thought about the question, it's not really a civil liberties or privacy question, but we didn't feel like we should ignore it, so we started having a conversation . . . but this was interesting how do we evaluate?" (Interviewee #10). The lines are not clear. Even the question of whether to do advanced analytics work for the Pentagon has no obvious answer. Several months after Google's announcement on Project Maven, Microsoft and Amazon separately affirmed their willingness to contribute to the Department of Defense's AI efforts. (Gregg 2018).

[5] This interview, and the internal debate to which the interviewee referred, took place prior to the United States' decision to add these Chinese companies to the Entity List, and so to prohibit U.S. companies from exporting certain items to them without a license on the grounds that doing so could compromise U.S. national security. *See* U.S. Dept of Commerce, Bureau of Industry and Security, Addition of Certain Entities to the Entity List, 84 Fed. Reg. 54002 (Oct. 9, 2019). Thus, at the time of the internal debate referred to, it was still legal for U.S. companies to export to these Chinese companies.

IV. WHAT IS "CORPORATE DATA ETHICS"?

Corporate use of advanced analytics produces benefits. But it also creates the harms outlined in the previous section. How should companies balance these positive and negative impacts? How should they determine whether a given advanced analytics project is legitimate and socially acceptable?

Traditionally, companies have looked to privacy law, and the Fair Information Practices (FIPs) that underlie them, as a guide in such matters. Privacy law generally requires that companies notify individuals before they collect and use their personal information; afford them some choice as to whether to allow this; and use the data only for the purpose specified in the notice and to which the individuals have acquiesced. So long as a company adheres to these core principles – notice, choice and purpose limitation – and the individual in question consents to the data processing, the company feels relatively comfortable that its data practices are legitimate. The individual consented to them, after all.

A number of interviewees explained that, while this approach may work for simpler forms of data processing, advanced analytics puts great strain on it. To begin with, the above-described harms that advanced analytics can impose extend well beyond privacy to bias, increased inequality, and other such areas. Privacy law's individual consent model is not designed to perceive and address these social harms. Second, U.S. privacy law governs only certain sectors, leaving important ones (social media, data brokerage, search engines, etc.) lightly regulated. Third, U.S. privacy law generally applies only where companies process personally identifiable information (PII). Advanced analytics, however, can find correlations in, and develop predictive

algorithms from, de-identified information. Where companies de-identify data before analyzing it, they arguably take advanced analytics outside the scope of privacy law.[6]

Finally, the interviewees explained that, even if a company were to try to comply with the spirit of privacy law and the FIPs, the nature of advanced analytics makes this difficult and places great strain on the notions of notice, choice and purpose limitation. Companies that engage in advanced analytics typically start by compiling or gaining access to a massive dataset drawn from multiple sources. Later, they look for correlations in the data and, based on these correlations, make inferences and actionable predictions. As a result, the company carrying out the advanced analytics may be several steps removed from the entity that first collected the data. This makes it difficult to go back and obtain consent from the individuals whose data make up the data set. "[Y]ou may be so many steps removed, you can't possibly have gotten consent from the individual in way that is reflective of what you want to do with that data. So, I think that in of itself is a problem." (Interviewee #22).

Interviewees further explained that they often do not figure out what they are going to try to learn from the data – the purpose of their processing – until after they have amassed the data set. Thus, even where a company is involved in the data collection and so in a position to seek consent, it often cannot specify the purpose at that moment and can only obtain the most general kind of consent. As one interviewee put it: "[T]he problem is that in order for it to be meaningful consent, the person or organization who was seeking that consent at the time would had to have thought through every possible way those data could be used and at least to have framed up, at least at a generalized level, a consent that's actually broad enough in scope. They don't have the ability to do that, unless you say, 'You're consenting to everything that we possibly ever might

[6] As one interviewee put it: "[A]ll privacy laws in the world are written with the caveat of personal information. But if you think about the number of potential technology-enabled decisions or impacts an individual could be subject to that have nothing to do with personal information, you start to say: 'Well, wait a sec, that just doesn't work anymore . . . .'" Interviewee #7.

want to do with this data.' It's problematic." (Interviewee #22).[7] Another concurred: "in the area of big data, where data is used well beyond the purpose of primary collection, it's almost impossible to get consent, informed consent, and consent that is useful." (Interviewee #23). For all of the above-described reasons, the interviewees believed that privacy law does not adequately address advanced analytics' potential harms, and that compliance with such laws is not a sufficient way to protect people from these harms.

The survey data shows something similar. In Figure 4, we see that a clear majority of respondents either do not address, or do not adequately address, the harms that advanced analytics can impose on individuals and the broader society.

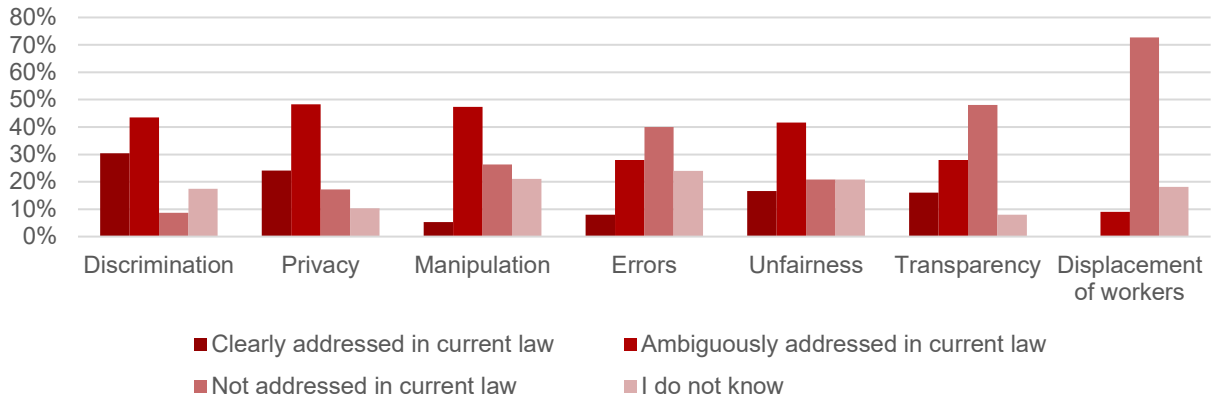

Figure 4. How well does current law address the risks from advanced analytics

[7] Another interviewee made a similar point:

> Let's put that in the context of big data. We have laws today that really are built on foundations of providing notice, providing a purpose specification, and gaining consent for whatever purpose you're specifying. That's simply inconsistent and does not take into account the reality of this observational world that we live in, or advanced algorithms where. . . the whole purpose of discovery is to discover causation or correlations that we can't anticipate. Otherwise it's really just research, or analytics, which we've been doing for decades. The fact that we don't know what we're going to discover is simply inconsistent with specifying purpose. (Interviewee #21)

Unable to rely on privacy law and individual consent as a source of legitimacy, some companies have begun to assess for themselves the social acceptability of particular advanced analytics projects.[8] They see themselves as venturing beyond privacy law and into the realm of substantive value choices, of ethics. For these companies, "data ethics" means assessing the legitimacy and acceptability of advanced analytics projects so that the company can act—or at least appear to be acting—in socially responsible way. "[T]he laws and regulations haven't caught up yet to this new, innovative use of data. Therefore, we will have to make our best guess at how to be ethical and responsible." (Interviewee #21).[9]

Framed in this way, data ethics shares much with other forms of corporate social responsibility, such as reducing carbon emissions, where companies go beyond compliance with the law. Data ethics is, in a sense, a form of corporate social responsibility for the algorithmic economy.

> *I almost think that it comes down to just this perfect storm of the company's history and philosophy around social responsibility. Because I actually think that everything we're actually talking about here, ultimately, is social responsibility. Not unlike the labor issues; not unlike the environment. I see patterns that are similar to the waves that we've seen of other social responsibility. And, I know we've never thought about this, or data protection, as a social responsibility function. But, I think ultimately it will be and those always align to the ethics department and tend to get pulled away from the legal department (*Interviewee #21).

We hypothesize that, in the next five to ten years, growing numbers of U.S. companies will include data ethics as part of their general corporate social responsibility initiatives and reporting.

[8]As one interviewee put it: where "the ability to go and get consent, really meaningful consent, just doesn't exist, you need another basis on which to do what you're doing." (Interviewee #22)

[9] One interviewee directly tied the rise of data ethics to the reluctance to rely on consent: "[the] reason that the ethics conversation is important and interesting . . . is: when do ethics let me use data without consent? . . . I mean, could Google use all the searches and go play the stock market? Who knows what machine learning will enable them to predict? And what can Amazon do with the data it learns about my home?" Interviewee #19.

This shift from a consent-based model to substantive assessments of impacts and harms is reminiscent of the longstanding debate in the privacy law literature between individual control-based, and use- or harm-based, approaches to privacy regulation. What seems to be happening is that, when it comes to advanced analytics, some companies are starting to move on their own from a consent-based model, to a harm-based one. One interviewee spoke to this directly: "the FIPs were not created with this world in mind. Transparency and choice in a world that is opaque and complex are not going to solve all of these problems. . . . I think the conclusion of the White House report about looking at use-based rules in the complexities of this world was right. The difficulty of that is how do you do it?" (Interviewee #3).

V. MOTIVATIONS – WHY DO COMPANIES PURSUE DATA ETHICS?

This account of data ethics as a form of corporate social responsibility leaves an important question unanswered: Why do companies make this effort? If existing U.S. privacy law does not require companies to be more responsible in their use of advanced analytics, why are they investing resources in doing so?

A. BUILD REPUTATION AND SUSTAIN TRUST

For one thing, companies want to build and maintain their reputation and the trust that others have in them. Negative incidents involving advanced analytics, such as the Facebook-Cambridge Analytica episode, which was particularly salient among survey respondents, can erode this trust, damage reputation, and so hurt the business. The interviewees cited the need to maintain reputation and trust as among the most important reasons that they invested resources in identifying and seeking to reduce the threats that advanced analytics can pose.

> *Well, there's a lot of different people like me at other companies that are trying to ensure that the trust in their brand is maintained and extended and the trust in the marketplace because trust is the fundamental of all human relationships and that's why if you act ethically and ensure the data use is ethical and you are fully accountable for that, then your brand is trustworthy and I think that is the most important. That's what we're all trying to achieve, so there's many, many companies get it and are trying to stand up programs or extend programs that really get at this fundamental of trust and operating ethically (*Interviewee #6).

While legal compliance is necessary for building a strong reputation and trusted relationships, it is far from sufficient. Negative incidents that do not violate the law can still affect the public's perception of a company. For example, one interviewee pointed to the recent controversy in which ProPublica found that it was able to market ads to Facebook users who included the term Jew Hater in their profile.

> *Well obviously it's a big reputational issue for Facebook. They don't want to be viewed as a company that's serving up ads based on antisemitism, and likewise Google doesn't want to do that as well. So it doesn't matter what the law requires. It's really a question of what's good for their business's reputation. And we run into that a lot with companies (*Interviewee #23).[10]

Companies that engage directly with consumers have the strongest incentive to avoid negative incidents and protect reputation and trust. For example, an interviewee from the retail industry explained that their company had developed a management approach to vetting advanced analytics that centered on whether the data practice would be seen as benefiting the consumer.

> *After ascertaining that a given project complies with the law, the company then asks: "does it put the customer first, is it something we ought to do, is it brand right. . . . If we have a reputational hit and we have customers that either decrease spend or are not spending with us at all for whatever reasons, that's really something that we want to avoid. That's the harm, and that's the basis. But at the end of the day, it's really about the customer . . . does this put the customer first?"* (Interviewee #17).

Other companies, particularly those that do not transact directly with consumers, focused on their standing with the general public. For example, one interviewee attributed their company's data ethics initiative to its decades-long reputation as "a really ethical company in the eyes of the public," and to the company's desire to maintain this reputation.

Other companies worried more about their reputation with regulators. As they saw it, legal compliance was not sufficient to maintain a good relationship with these public officials. They also needed to show that they were good actors. A high-profile incident involving the unethical use of advanced analytics could damage regulators' image of them, even if it did not involve a violation of law. The Facebook-Cambridge Analytica episode, which may or may not have involved a legal

[10] An attorney put it to their clients this way: "I can tell you this thing you're doing, that you're proposing, it's perfectly legal. But, there's a really good chance there's going to be a really crappy New York Times article about you on this. I don't think you want that. So, let's brainstorm about ways that we can avoid that and achieve the business objective you're trying to achieve in a different way." (Interviewee #12).

violation but certainly caused regulators to scrutinize Facebook more carefully, illustrates this. As one interviewee explained:

> *"I think some of the companies are very motivated by wanting to . . . have good, trustworthy relationships with regulators. So they're seeking to balance a number of different factors: how regulators perceive them, how their customers perceive them as well as how actively they're able to use and move data around the world."* (Interviewee #16).

One interviewee focused, not on customers or regulators, but on their company's reputation among its business partners. As this interviewee saw it, individual consumers lack the resources and expertise to assess meaningfully a company's data practices. Business partners, particularly those who share their data with a company, are much more likely to scrutinize these practices carefully, especially where a negative incident could reflect back on them. A company needs to handle data ethically in order to earn the trust of these business partners.

> *As we all know, [individual] people don't read privacy policies…. When you're [a company that is] signing up for a CRM [customer relationship management] contract, for years and millions of dollars, you're reading every word of every agreement, right? . . . . Companies that sell these types of products need to make sure what they say in these contracts is true. But, more importantly that they're following their own controls and public statements about privacy and security protocols, and living up to the values that they leverage when they sell these products (*Interviewee #9).

The survey data, too, however, suggested that reputation among business partners is may not be as strong of a an important driver of data ethics management relative to consumers. In Figure 5, Companies in our data set that primarily sell to businesses have 4.5 times higher lower odds of having a policy in place to manage the risks from big data than do companies that primarily sell to consumers though the difference is not statistically significant ($p=0.21$).

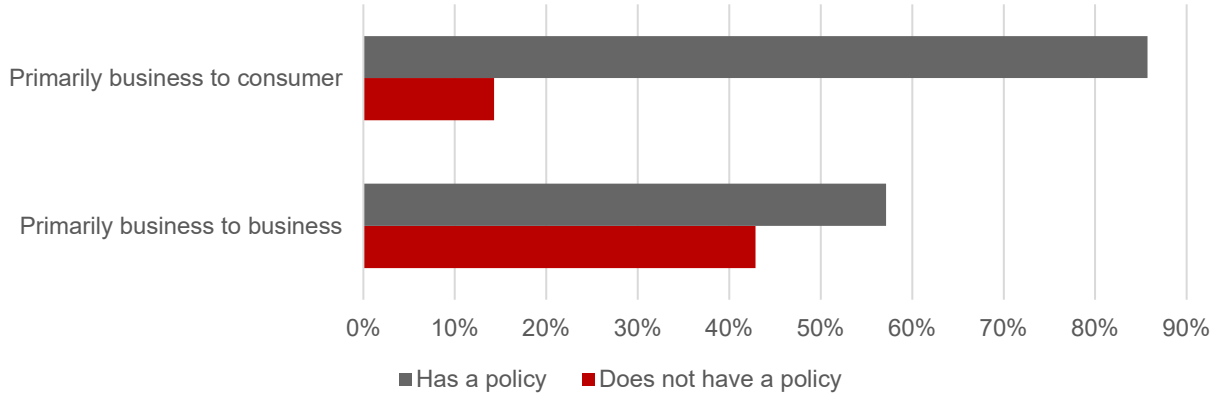

Figure 5. Incidence of company policy for managing risks of big data by whether a respondent's company is primarily business to consumer or business to business

In sum, the interviews suggest that—whether to protect their standing with customers, regulators, business partners, or all three—companies pursue data ethics in part to protect their reputation as responsible stewards of people's data, even where the law does not require them to do so.[11]

> *Preying on vulnerable populations, treating people unfairly, manipulating people in ways that could harm them . . . . There's some of that stuff that's perfectly legal, but it still may not be a good business decision. I'll throw out the word ethics. It's not the ethical thing to do. Some companies that I work with, they take that stuff very, very seriously. They don't want to do things that feel, or could be perceived as, unethical (*Interviewee #12).

The survey data supports the idea that trust and reputation are important drivers of corporate data ethics management. For example, we asked survey respondents whether their company had adopted a policy for managing the risks of big data. We further asked respondents whether the media (Figure 6), advocacy groups (Figure 7) or

[11] As one interviewee explained that a company derives great value from it reputation and that this justifies an investment in data ethics. "A company kind of built their reputation over a 25-year period. And it's worth billions of dollars as an asset. And so they're very protective of that. And so . . . they are willing to devote substantial resources in making sure that they avoid [things that detract from their reputation." Interviewee #15.

employees or investors (Figure 8) had brought pressure on companies in their industry to achieve better data ethics performance. As conveyed in the following figures, companies were more likely to adopt a policy for managing big data's risks when their industry had experienced such pressures. While these findings are not statistically significant, the direction of associations is informative. Collectively, Figures 6-8 suggest a link between external pressures and having a policy in place to manage the risks from big data.

While the current survey was designed as exploratory and does not seek to assess causality, we did look further into the timing of media pressures and the creation of policies. Approximately 50% of respondents indicated that the media pressures on their industry began in 2015-2019, and almost 75% of these companies stablished their policies over the 2016-2019 period. The overlap of periods provides useful direction for future work to investigate external pressures. being meaningful for the timing of corporate policies.

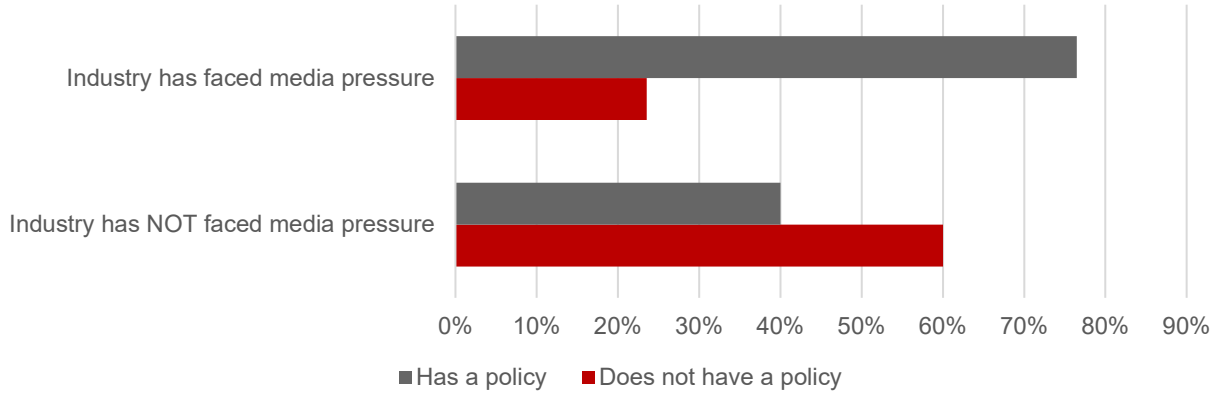

Figure 6. Incidence of company policy for managing risks of big data by whether respondent's industry has experienced media pressure

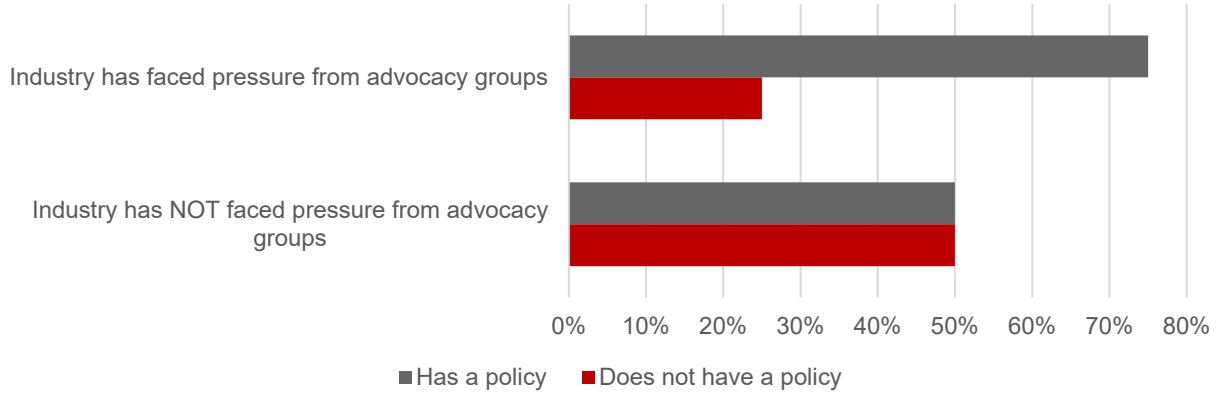

Figure 7. Incidence of company policy for managing risks of big data by whether respondent's industry has experienced pressure from advocacy groups.

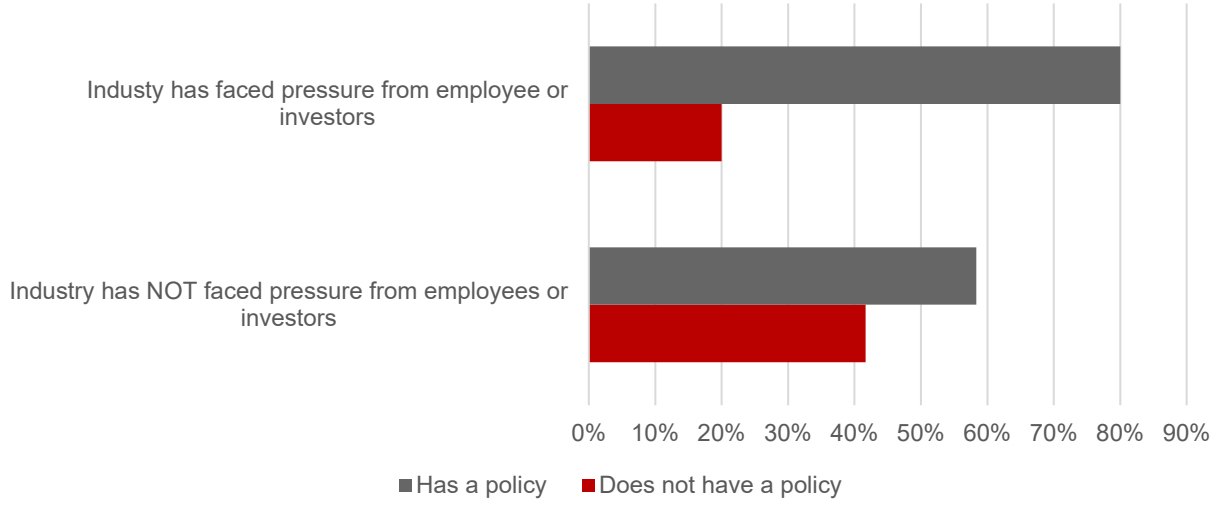

Figure 8. Incidence of company policy for managing risks of big data by whether respondent's industry has experienced pressure from employees or investors.

B. OPERATING IN THE SHADOW OF THE LAW

The fact that the U.S. law does not currently impose many limits on the corporate use of advanced analytics does not mean that it will not do so in the future. Indeed, a number of factors are pushing the law in this direction. To begin with, the Facebook-Cambridge Analytica episode, which centered on Cambridge Analytica's use of advanced analytics to manipulate voters,[12] ignited a firestorm of Congressional and public criticism of Facebook and the tech sector more generally. Public reaction to this incident contributed to the passage of the California Consumer Privacy Act, arguably the most significant piece of state privacy legislation in the United States in a generation.

Much is happening on the federal level as well. Congress is currently considering a number of privacy bills, some of them with sponsors from both sides of the aisle. Several of these bills go beyond the notice-and-consent approach to privacy regulation and impose use-based restrictions that could more directly curtail abusive data analytics practices (Kerry 2019). On the regulatory level, the Federal Trade Commission is starting to focus its attention on business use of analytics. In a recent report, the Commission suggested that it might use its Section 5 unfairness authority to reign in algorithms that had disparate, negative impacts on racial minorities or other protected classes.[13]

Many of the interviewees believe that governments will soon regulate advanced analytics in order to reduce the risks that it poses. The survey respondents also believed that future regulation was likely, although they were a bit less certain on this point than the interviewees.

[12] Cambridge Analytica obtained the data of 87 million Facebook users. Using advanced analytics, it inferred the personality types of these individuals. It then sent them political ads, at the behest of the Trump campaign, that appealed to each voter's particular personality type and so influenced them in ways that they could not consciously detect. Cambridge Analytica's use of advanced analytics, and its potential impact on the Presidential election, is one of the reasons that this incident outraged so many people.

[13] Federal Trade Commission, *Big Data: A Tool for Inclusion or Exclusion?* 23 (January 2016) (stating that "Section 5 may also apply . . . if products are sold to customers that use the products for discriminatory purposes.135 The inquiry will be fact-specific, and in every case, the test will be whether the company is offering or using big data analytics in a deceptive or unfair way.")

We asked the survey whether they agreed—from 1 (strongly disagree) to 5 (strongly agree)—that there will be new US Federal or State government regulation of big data analytics in the coming years. As conveyed below in Figure 9, almost 70% of the sample agreed that some state regulation is likely, while almost 50% agreed that federal regulation is likely. This different perceptions of the likelihood of state versus federal regulation may reflect perceived gridlock at the federal level.

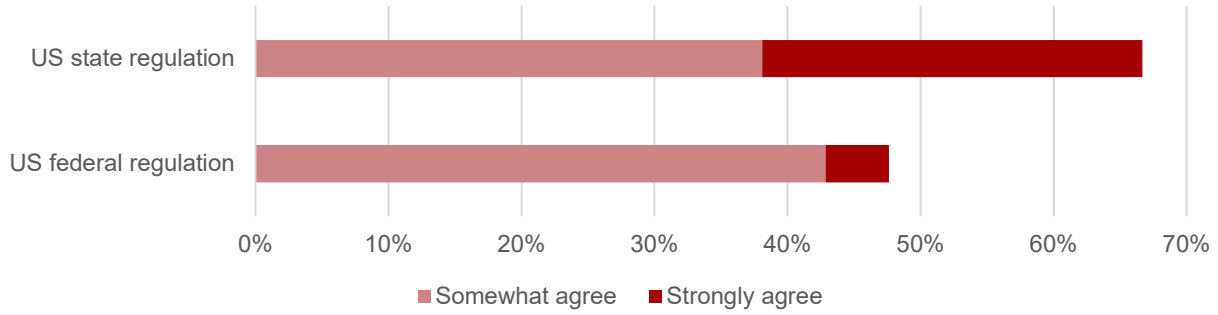

Figure 9. Do you agree with the statement that there will be new regulation (federal or state) of big data analytics in the next 5 years?

The interviews suggest that the prospect of future U.S. regulation drives companies to focus more on data ethics. Some companies believe that if industry can demonstrate that it understands and is able to reign in advanced analytics' harmful aspects, it may be able to pre-empt and prevent the passage of stringent government regulation. As one interviewee who works with a variety of businesses on their data ethics initiatives explained:

> *a business with smarts, they would be advocating for self-regulatory or even co-regulatory types of models that held them accountable to a different standard of accountability or stewardship as a means to stave off what will invariably be badly written [regulation that will] have negative consequences from a data perspective to businesses (*Interviewee #7).

A second group of companies focus not so much on pre-empting future regulation as on shaping it. They worry that ill-informed government regulation could be incompatible with their business models and operations. They believe that, if they take the initiative to develop and implement strategies for reducing advanced analytics' harms, policymakers may draw on these models when drafting legislation and regulations. This could make future regulation more effective and more feasible from a business perspective.

> *But the smart ones are going to say wait a sec, this is the inevitable future and I want to stay at least one step ahead of it. And they're going to start to both work to influence the development of those regulatory guidance frameworks and . . . to implement their own ethical or fair data processing standards as a means to achieving sort of trusted data optimization (*Interviewee #7).

Finally, there are companies that want to get ahead of future regulation so that, when it does come, they can adapt to it more easily, and at lesser cost, than their competitors. These "smart organizations are seeing the tea leaves and saying, 'I really want to make sure that I stay at least one step ahead of that.'" (Interviewee #7). These three motivations – to pre-empt, shape and prepare for future regulation – are important drivers of corporate investment in responsible advanced analytics and AI. Still, the move in this direction remains more the exception than the rule. "[F]or many organizations, they just don't want to invest in something unless they have to." (Interviewee #7). In our survey, we see a positive association between having a policy in place to address the risks of advanced analytics and the extent to which respondents see US state regulation in the near future (Figure 10).

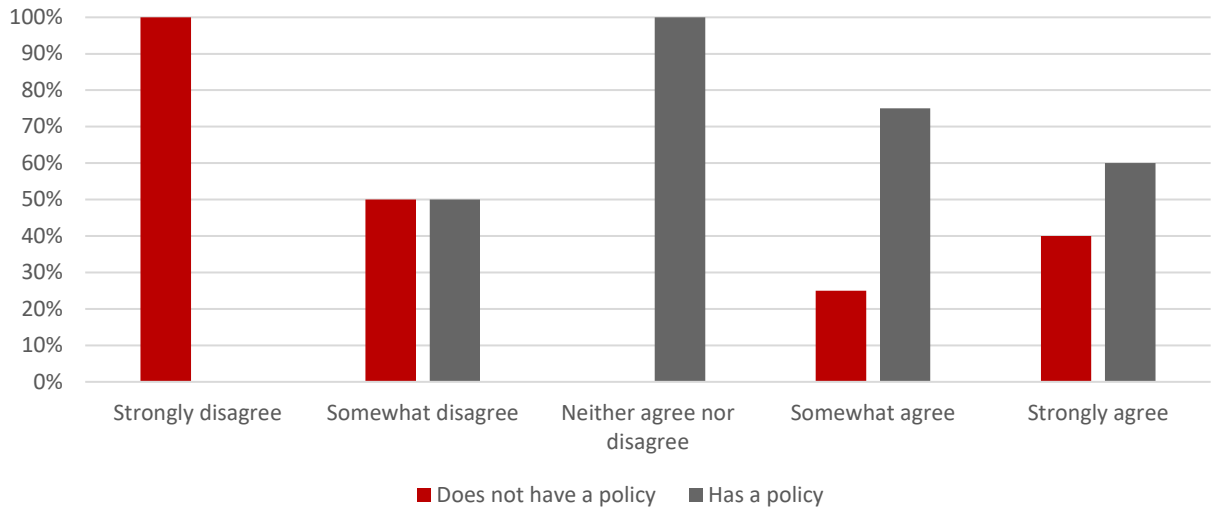

Figure 10. "Do you agree with the statement that there will be new state regulation of big data analytics in the next 5 years," by whether a company has a policy in place.

C. EUROPEAN DATA PROTECTION LAW

Some of the interviewees worked for companies subject to the European Union's General Data Protection Regulation. They cited the GDPR as a separate, additional motivation for pursing data ethics with respect to U.S. citizens' data. Article 22 of the GDPR focuses expressly on "automated decision-making," a term that encompasses much of what this report refers to as advanced analytics. Somewhat surprisingly, however, the interviewees did not cite Article 22 as the reason for their company's investment in data ethics with respect to U.S. citizens' data. Instead, they talked about the effect that the Article's "legitimate interests" balancing test had on their companies.

Article 22 of the GDPR, titled "Automated decision-making, including profiling," expressly governs machine-based data processing such as advanced analytics. Article 22 provides that "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly

significantly affects him or her." Article 22 assumes that the automated decision-making in question utilizes the data subject's "personal data," since the GDPR governs only the processing of personal data. In essence, Article 22 prohibits companies from taking an individual's personal data, using a machine to analyze the data and, without any human input ("a decision based *solely* on automated processing"), making a decision that has legal effects on the person.

One might expect this data analytics-oriented provision to push companies towards data ethics but, according to the interviewees, it does not do so in any significant way. One reason is that companies are able to avoid Article 22's requirements either by de-identifying the data so that it no longer qualifies as "personal data" and is no longer subject to the GDPR; bringing a human into the decision-making loop so that the company is not making decisions "based *solely* on automated processing"; or using advanced analytics only for decisions that do not have a legal effect on individuals.[14] Secondly, even where Article 22 does apply, it does not draw substantive lines that separate ethical from unethical data analytics but rather *prohibits* the decision-making operation entirely (unless one of the exceptions applies). Thus, even those companies that seek to comply with GDPR standards in their processing of U.S. citizens' personal data, as some do,[15] do not credit Article 22 with pushing them towards data ethics.

It is, instead, Article 6, "Lawfulness of Processing," that has spurred companies towards data ethics. European data protection law differs from U.S. privacy law in a fundamental way. In

[14] "With the GDPR coming into effect and . . . quite clear rules around profiling and automated decision making. . . . They are trying to think about ways to mitigate that risk. Either avoiding using automated decision making for the things that presumably would trigger GDPR, . . . things that have legal effect or very significant effect . . . Or building in some kind of human review." (Interviewee #12).

[15] One interviewee reported that their company does this so as to facilitate the global data flows on which the company depends. "So one of the things . . . we were trying to do is to have a common internal standard that went above the law, that provided the flexibility that if you work in Spain or you work in India or you work in Chile or the U.S., that you follow the same standards everywhere, and that you protect the data the same way everywhere in which you are. That should be a basis for which data can move across the business and systems that support people on a global scale . . . [and so to] move data around the world . . ." Interviewee #22.

the US, companies may process personal data so long as no law forbids it. In the EU, the default is the opposite: companies (and other data controllers and processors) cannot process personal data unless a law authorizes it. The GDPR codifies this requirement in Article 6 which requires companies to have a lawful basis for any processing of personal data. Consent of the data subject is one such lawful basis. But where a company is not able to obtain consent, as is often the case when it comes to advanced analytics, then an entity may lawfully process the data if the "processing is necessary for the purposes of [its] legitimate interests," and is not "overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data."

This lawful basis, known as the "legitimate interests balancing test," shares much with data ethics as the interviewees and their companies practice it.[16] Even where companies are in compliance with Article 22 (by, for example, bringing a human into the loop), Article 6 requires them to consider how their data analytics practices could injure the "fundamental rights and freedoms" of data subjects, and to balance those negative impacts against the company's own interest in carrying out the processing. One interviewee explained how this substantive balancing analysis differed fundamentally from traditional privacy law's focus on notice and consent and, in so doing, promoted a new "data ethics" mindset.

> *It's GDPR . . . that's helping organizations think about these things a little bit differently . . . [T]his concept of legitimate interest. . . . [i]t was the first time that there was a beginning, if you will, of starting to frame out ways of thinking about risks and benefits, impacts and interests. . . . Big data analytics, oftentimes you don't as an organization have the ability to get consent. . . . So, organizations are looking at other ways of trying to ensure that they are lawfully using data that are subject to GDPR, and it's forced them to start thinking about. . . who might be affected by this data if I do what I want to do? Who might be harmed by what I do with the data? Who might benefit from it? . . . .[H]ow do I set up the way in*

[16] As one interviewee put it, "there is not a lot of difference between the type of analysis you do, to make sure that the score card is balanced for what legitimate interest, and what you would do to make sure that benefits and risks are balanced in big data." (Interviewee #1).

> *which I safeguard or put protections in place, such as encryption or greater transparency . . . . That's the kind of analysis that organizations who are starting to think about this are being forced to do, which is just a very different way than they've ever had to think about it before (*Interviewee #22).

The GDPR's most important contribution to corporate data ethics may stem, not from its automated decision-making provisions, but rather from its legitimate interests balancing test and the way that it gets companies to articulate and balance the benefits and the risks that their data analytics operations can generate.

D. RECRUIT AND RETAIN EMPLOYEES

Several interviewees linked data ethics to employee retention. This was particularly evident among firms that employed data scientists. These individuals are in high demand. They are more likely to move to another employer when the company for which they work uses data in ways that offend their values. A pro-active approach to data ethics thus becomes important to employee retention which, itself, is critical to the company's success.

> *I will say one thing about engineering companies in the Valley . . . the engineers themselves –these 24 year old kids – are powerful. . . . They are valuable and they are the strength of the company. So, . . . there's two markets you compete in for a tech company: one is to sell your products, the other is to attract talent – the talent market. If you don't get the best talent, you don't have the best product. . . . [T]hese are engineers coming out of generally elite, progressive, generally liberal institutions. They're going to come in with sort of a mindset that is very pro-privacy and civil liberties. They want to do the right thing (*Interview #10).

E. MAKING RISK-BASED DECISIONS

It may sound counter-intuitive, but the driving force driving for some companies to limit their use of advanced analytics is their desire to use advanced analytics more fully. This stems from companies' uncertainty over where the line between socially acceptable, and unacceptable, uses of this technology lies. Faced with this uncertainty, many companies find it difficult to make

risk-based decisions about particular advanced analytics use cases. Risk-averse companies hold off on such uses, and so lose out on the value that advanced analytics, as applied to their data, could have generated.[17] Companies that put resources into data ethics management and so develop a way to spot and navigate ethical issues connected to their use of advanced analytics not only protect people better; they also improve their ability to make quick and effective decisions about whether to proceed with particular analytics projects. This frees them up to use advanced analytics, and their own data, more fully. One consultant who works with many companies, referred to this as overcoming "reticence risk."

> *[Q]uite frankly, . . . the biggest driver of data value creation loss, and the increasing problem organizations face, is what I call self-inflicted reticence risk. It is their inability to make internal decisions about whether they should or shouldn't do something related to the use of data. . . . In the absence of a decision-making process inside the company, the risk voices always win. The result is these organizations end up leaving value on the table. . . . The decision-making process should address: 'I don't know whether I can use data;' and, 'even if I legally can do it, I don't know whether I should do it.' Absent a more formalized decision-making process, organizations find that many, if not every, stakeholder inside the company has an opinion on this and that these opinions cannot be reconciled. The result is that data activity grinds to a halt . . . and these organizations end up only using only 30%, 40% of the data or the value because they can't reconcile the risk. . . . Reticence risk is leaving value on the table because you just can't make a risk-based decision* (Interviewee #7).

F. ACHIEVE COMPETITIVE ADVANTAGE

A surprisingly large number of interviewees said that their companies pursued data ethics for competitive advantage. In a sentiment related to the above comments on reputation and trust, some focused on the market benefits of a good reputation.

> *[I]f you get people to believe . . . that you are handling things in a responsible manner, they're more likely to keep doing business with you or want to do business*

[17] One study estimated that the median Fortune 1000 company could increase its revenue by $2.01 billion a year just by marginally improving the usability of the data already at its disposal. Anitesh Barua, Deepa Mani, & Rajiv Mukherjee. "Measuring the Business Impacts of Effective Data," University of Texas McCombs School of Business, September 1, 2010, p. 3.

> *with you. . . . It can help you in the marketplace. . . [Anytime] we are able to talk about how we are handling or managing data in a responsible way, it does nothing but help our brand and our reputation (*Interviewee #9).

Others framed the advantage differently. As they saw it, an ethical product or project is one that benefits, rather than harms, customers. They believe that, by pro-actively working to prevent harms to consumers, they improve the customer experience and so make the company's products more attractive.

> *[I]t's also . . . what's the customer experience? It's just making them think through things that I wouldn't have had a problem with. They're not necessarily data ethics concerns. But suddenly they'll just realize this is actually to have a crappy customer experience. . . . So they're actually seeing this as a benefit from the business perspective (*Interviewee #20).

This is also how they explain their role to the business units that they work with. They find this message to be more effective than an explicitly ethics-based one.

> *If you go to a team and say, "Hey, I'm here to do ethics review." They immediately think, "What? Am I being unethical? Am I doing something wrong?" . . . It sends the wrong message. So I often frame ethics questions as product improvement questions. Like, "I just want to know how you're doing things, and let me see if I can help you figure out what the sensitivities might be, and how we can resolve them." And those questions then, are ethics questions. And that's really my job right now. . . So, that's what we're doing here, is we're making these projects better because we ask questions. . . . we have many, many examples where we are truly, proud of the work that we did as a team, because we know that we improved the project (*Interviewee #14).

A 2016 Price Waterhouse Coopers report suggested a number of levels on which data ethics could create competitive advantage, maintaining that "[t]hose that [have more developed ethical frameworks] could find themselves a magnet for employees, customers, and even investors who increasingly favor organizations that operate ethically and responsibly. In fact, several studies have confirmed that companies operating ethically outperform others in revenue and profitability. . . . [they] gain a strategic advantage by excelling in leveraging data's upside

while managing risk and reducing costs." (PwC 2016). One data ethics manager expressed a similar sentiment: "the time has come . . . where privacy is a differentiator, and data ethics is even something that's going to further differentiate . . . . I don't think there's any way you can escape it." (Interviewee #16).

Data ethics managers likely have a vested interest in believing that their work makes their company more competitive and successful. These same managers may not be the most reliable reporters on whether, in fact, data ethics management has this effect. Based only on these reports, we cannot conclude as to whether data ethics does, in fact, produce such an advantage. But it is interesting that some data ethics managers see themselves, and explain their role to their companies, in this way.

G. FULFILL CORPORATE VALUES

The above motivations behind data ethics management are rooted, in one way or another, in the company's self-interest. However, a number of the interviewees felt strongly that their companies pursued data ethics for more intrinsic reasons having to do with the company's core values. As one said:

> *[T]hat's not the only driver, following the rules, following the law. I think [my company] has other drivers. One, is company values. Those might extend beyond the exact letter of the law, and I think [company] is a values-based company, and I'm not just saying this as a marketing pitch, this is what I think. I think [company] has strong values about protecting its customers, protecting its own information, and its employees, and then making sure that it's a good steward of public information. [The company] puts a lot of energy, and puts forth a lot of resources to be a good steward in those regards (*Interviewee #18).

The above, anecdotal account of the motivations behind data ethics management does not allow us to say which motivations are most important. But it does suggest that there are many reasons why a business might put resources into data ethics management, even the law does not require it to do so.

VI. DRAWING SUBSTANTIVE LINES

Having looked at why companies are pursuing data ethics, we turn now to how they are doing so. The interviews suggest that, in order to establish a data ethics management program, companies need, at a minimum, to do three things: (1) draw a substantive line between ethical, and unethical, data analytics practices; (2) manage their data analytics operations to ensure that the company stays on the "ethical" side of this line; and (3) develop and implement technologies that facilitate the ethical use of data analytics. This part will examine how companies draw the substantive lines between ethical, and unethical, data analytics practices. Part VII will describe how companies manage their operations in order to achieve this substantive goal. Part VIII will discuss some of the technologies that companies are deploying to this end. Our interviewees and survey respondents mainly held chief privacy officer or other, similar positions. These professionals tend to focus on the substantive and managerial dimensions of data ethics management, rather than on the technological one. The following discussion thus focuses more on the substantive and managerial, than on the technological, dimensions.

When a company sets out to improve its data ethics, the first thing that it needs is a way to distinguish between advanced analytics practices[18] that are ethical, and those that are not. Without this, it does not know what it is trying to achieve.

[18]Some of this literature uses the term "artificial intelligence" rather than advanced analytics. These are distinct, but heavily overlapping categories. Big data analytics, machine learning, advanced analytics and artificial intelligence all refer to analytic operations that take advantage of the massive data sets and processing capabilities that have become available in the past decade or so, and that use them to find correlations and make predictions. This White Paper uses the term "advanced analytics." But this term overlaps significantly with the others just listed.

A. PUBLISHED DATA ETHICS PRINCIPLES

In the past few years, an extensive body of literature has grown up on how to draw substantive lines between ethical, and unethical, advanced analytics and artificial intelligence. Scholars, governmental bodies, multi-stakeholder groups, industry think tanks, and even individual companies have contributed to this literature.

The literature largely follows a similar pattern. The author first sets out an ethical framework grounded in human rights, a school of philosophy, bioethics, fiduciary duties or some other established set of principles. The author then sets out these principles as the basis for distinguishing between ethical and unethical advanced analytics and AI practices. Frequently, the author suggests that companies can use these principles as the basis for its data ethics decision-making.

In the scholarly arena, Professors Luciano Floridi and Josh Cowls (2019) illustrate this approach in their recently published *Unified Framework of Five Principles for AI in Society*. Floridi and Cowls maintain that data ethics shares much in common with bioethics.[19] They set out a unified framework for data ethics that adopts the key principles of bioethics – *beneficence, non-maleficence*, *autonomy*, and *justice*[20] – as well as one additional principle, *explicability*. They maintain that these "Five Principles for AI in Society" should guide specific sectors and industries as they decide which AI practices are ethical and which are not.

On the governmental front, the European Data Protection Supervisor's Ethics Advisory Group's (EAG) 2018 report, "Towards Digital Ethics,"[21] offers its own list of guiding principles. These

---

[19] *Id.*

[20] The first four of these principles emerge from the dominant approach to bioethics and medical ethics (Beauchamp and Childress 2013).

[21] Governmental bodies in the European Union have led the way in in articulating data ethics guidelines and principles. As in the realm of privacy regulation more generally, other countries will likely follow the Europeans' lead. It is therefore useful to consider examples of how EU governmental bodies contribute to the data ethics literature.

include Dignity, Freedom, Autonomy, Solidarity, Equality, Democracy, Justice and Trust (European Data Protection Supervisor 2018). The Ethics Advisory Group put forth this set of principles so that companies and others engaged in advanced analytics could "integrate [them] in both their designs and business planning reflection about the impact that new technologies will have on society."[22]

A year-long multi-stakeholder process involving policymakers, industry stakeholders, civil society organizations, and professional orders, among others, produced the Montreal Declaration. The Declaration *identifies* ten principles to guide the use of artificial intelligence: (1) Well-being, (2) Respect for autonomy, (3) Protection of privacy and intimacy, (4) Solidarity among people and generations; (5) Democratic participation, (6) Equity; (7) Diversity inclusion, both social and cultural; (8) Prudence in anticipating potential adverse consequences; (9) Human responsibility; and (10) Sustainable development.[23] It establishes these principles as a guide for private and public entities to use in developing and deploying AI in ways that "are compatible with the protection and fulfilment of fundamental human capacities and goals."

Industry-oriented think tanks and trade associations articulate similar sets of principles to guide corporate use of advanced analytics. For example, the Information Accountability Foundation, an influential industry-funded think tank based in the US, published a Unified Ethical Frame for Big Data Analysis. (Information Accountability Foundation 2015). This document recommends that, in "developing an assessment framework necessary to assure a balanced, ethical approach to big data," companies should seek to align their advanced analytics practices with five core values: "Beneficial, Progressive, Sustainable, Respectful and Fair."

[22] EDPS 2018 at 7. *See also, id.* at 15 (describing the principles "as a means to fill critical gaps in existing legal regulations and as a way of supporting those actors who work to adapt ethical principles to rapidly evolving issues, which often outpace the evolution of law.")
[23] https://nouvelles.umontreal.ca/en/article/2018/12/04/developing-ai-in-a-responsible-way/

Finally, a growing number of companies have begun to adopt and publish their own sets of data ethics or AI ethics principles. For example, Google's Objectives for AI Applications states that AI should: "1. Be socially beneficial; 2. Avoid creating or reinforcing unfair bias; 3. Be built and tested for safety; 4. Be accountable to people. 5. Incorporate privacy by design principles. 6. Uphold high standards of scientific excellence."[24] Microsoft's AI Principles are quite similar: (1) Fairness. All systems should treat people fairly (2) Reliability and Safety. All systems should perform reliably and safely (3) Privacy and Security. All systems should be secure and protect privacy (4) Inclusiveness. AI systems should empower everyone and engage people (5) Transparency. AI systems should be understandable (6) Accountable. People should be accountable for AI systems.[25]

These examples are just a slice of a much broader array of articles, reports and statements that set out abstract ethical principles to guide the deployment of advanced analytics and AI. In a 2020 report, Harvard University's Berkman Klein Center for Internet and Society identified and analyzed several dozen such frameworks from government, civil society, the private sector, multi-stakeholder groups and inter-governmental organizations (Fjeld, et al. 2020). The report identified eight core themes that many of them share: privacy, accountability, safety and security, transparency and explainability, fairness and non-discrimination, human control of technology, professional responsibility, and promotion of human values.

Rather than consolidating all sets of principles into a single framework, as the Berkman Klein Center did in its report, we find it helpful to distinguish between two categories of such frameworks which we call "moral" and "practical." On the one hand are frameworks that appear to be grounded in moral philosophy or human rights traditions. The EU Data Protection

[24] https://ai.google/principles/
[25] https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimaryr6.

Supervisor's Ethics Advisory Group's focus on "Dignity, Freedom, Autonomy, Solidarity, Equality, Democracy, Justice and Trust,"[26] and the Montreal Declaration,[27] with its emphasis on "well-being," "solidarity," "autonomy" and "equity," exemplify the "moral" category. They integrate moral and human rights ideals that are at once so universal and essential that they are almost beyond question, and so abstract that, unless they are further elaborated, would prove difficult for a company to operationalize. By contrast, Google's Objectives for AI Applications[28] emphasizes practices – such as accountability, privacy by design, avoiding unfair bias, building and testing for safety – that are grounded in traditions of privacy management and practice. They appear more practical and implementable, even as they leave out essential moral and human rights commitments that might drive a company towards something more worthy of the term "ethics" (note that Google refers to "objectives," not "ethics.")

Figure 11 shows that large percentages of our survey respondents noted that some of these types of substantive frameworks from academics and think tanks played a role in shaping internal policies for addressing the ethical risks they face with advanced analytics. Specifically, Figure 12 shows that many respondents also had seen specific documents produced by organizations like the Information Accountability Foundation and Future of Privacy Forum. While this would suggest that published external principles are important, it is not clear from the survey just how influential these types of ethical principles are. In fact, most interviewees stated that their companies resorted to informal benchmarks (discussed below) to make decisions rather than formal, ordered sets of ethical principles. One key issue is that, although the lists of principles may inform discussions within companies, in and of themselves they frequently do not lead to an all-things-considered judgment of *what to do.*

[26] *Id.* at 16-21.
[27] https://nouvelles.umontreal.ca/en/article/2018/12/04/developing-ai-in-a-responsible-way/
[28] https://ai.google/principles/

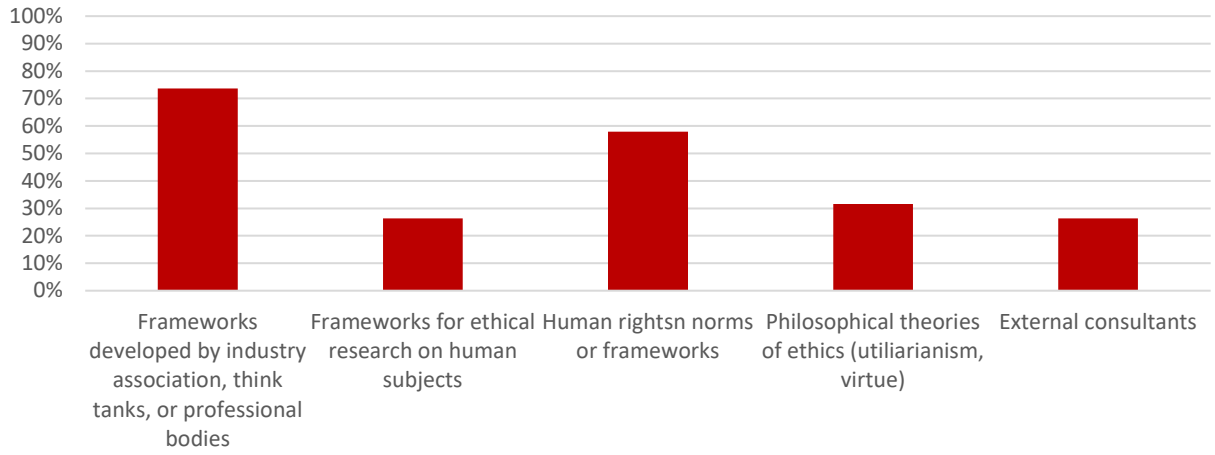

Figure 11. Did any of the following shape the content of your internal policy for dealing with the ethical risks of big data analytics?

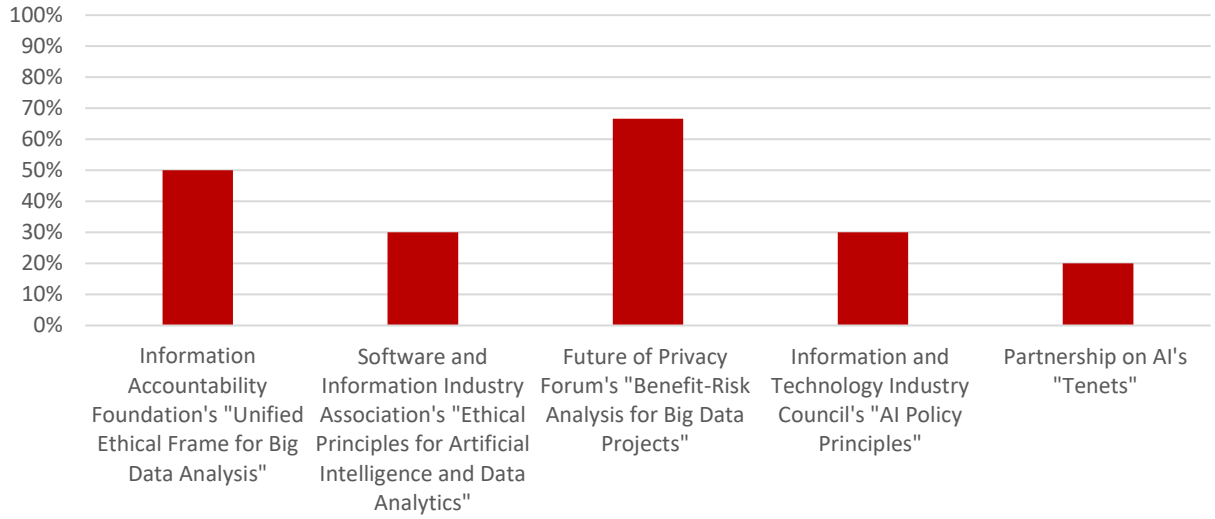

Figure 12. Has anyone within your company to your knowledge seen any of these documents?

B. INFORMAL STANDARDS

Given the abundance of relevant ethical principles, our research team expected the interviewees to describe how their companies were using such substantive frameworks in their

ethical decision-making. But that is not what we found. Fourteen of our interviewees were corporate employees (the other nine worked for law firms, think tanks, or were consultants that advised companies on data ethics matters). Of the fourteen who worked for companies, only three referred to formal principles when explaining how their companies made data ethics decisions. We describe their accounts below. The remaining eleven companies described their companies' heavy reliance on informal benchmarks for making these decisions. The ethics lead for a large tech company explained that their approach was "heavily leaning [towards] the informal [approach to data ethics decision-making]. We don't have any: 'Hey, based on this document that we wrote six months ago, this is now sensitive, or meets that qualification, or meets the definition of risky." (Interviewee #14).

The survey data on professional training of those who handle the data ethics function is consistent with the interviewees' reliance on informal benchmarks rather than formal ethical frameworks or sets of principles. The survey asked whether the respondents' "work has been influenced by any type of formal ethics training." Of the twenty-two survey recipients who responded to this question, only one indicated that they had a formal degree in ethics, and only six said that they had received ethics training of any type from a source outside the company. By contrast, twelve of the twenty-two said that they had a law degree. This suggests that those charged with making data ethics decisions are unlikely to have deep knowledge of ethical frameworks or philosophies of ethics. They are more likely to have received training in the kind of practical judgements, informed by laws and by broad human rights or ethical concepts, that lawyers tend to make.

The interviewees were very clear about applying informal standards. For example, a privacy professional at a leading technology company explained that, when presented with a sensitive or

highly innovative project, they first apply a cursory "ear test." Only if the project passes the ear test does it get sent on for full Review Board consideration. The ear test is highly informal.

> *The ear test simply means to me: does that sound right, does that sound like a bad idea? Do the words coming out of your mouth make sense from a legal, ethical, and business standpoint?". . . [W]e really think of those as kind of cursory, baseline ethics analysis. Our attorneys ask themselves: 'does that feel right what you're saying, what you're suggesting? You want to use this data for this purpose . .. Does that make sense? . . . does that just feel right?* (Interviewee #18).

Another highly experienced privacy officer at a major company described employing a "fairness check." The executive described this as: "Would my mother think this is okay? Would I want this to happen to my kid? Do I feel good about this personally? . . . We all know unfairness when we see it and I think that's an important construct and you'll hear it. It's a resonant term. Everybody in the policy circles is beginning to talk about, 'Is it fair to the individual?" (Interviewee #6). A third interviewee explained that "the standards we use are primarily two things: One, are you finding this creepy? Which is an undefined, but everybody knows it means standard – the creepy standard. Two, do you want to live in the world that this creates?" (Interviewee #10).

Creepiness. The "ear test." What would my mother think? Do I want to live in the world that this data practice creates? These are informal, intuitive, expectation-based judgements, not formal ethical principles. Most of the companies that we spoke with were using standards of this type to draw the substantive lines between ethical, and unethical, uses of advanced analytics. Two central ideas permeate this informal approach. One is desire to stay within the expectations of important stakeholders.[29] One privacy officer explained that they ask engineers: "do you really think grandma's expectation was that her data was going to be used in the way you're suggesting

[29] Where companies went beyond customer expectations, they tried to do so gradually and carefully. "You can't flip that overnight. . . . You've got to put in some work here to bring the customers, bring the consumers, bring the regulators, bring everybody that might be looking at this in concern along, so they can understand what's happening and why it's happening, and what consumers are getting back from this. So that a conversation can take place and we can develop a new norm." (Interviewee #12).

when she allowed for it to be collected?" (Interviewee #18). Another, talking about the informal test that their company applies, recounted that "[t]here's one person in the company that calls it the newspaper test. There's another person that has the grandmother test. There's all these metaphors that are used when these kinds of things are decided. If we're going to end up telling an individual, and sitting down with them for an hour to explain exactly what we're going to do, if there's any chance that that person would object to that, then the general rule is, then we shouldn't do it." (Interviewee #21).

The second theme is the Golden Rule – "Do unto others as you would have them do unto you." When privacy professionals pose the question: "do you want to live in the world that this creates," (Interviewee #10), or "[w]ould I want this to happen to my kid?" (Interviewee #6), they are, in a sense, asking their engineering teams and organizations to follow the Golden Rule. Abstract ethical frameworks may help one to think about these questions. But ultimately, as one experienced attorney told us, it comes down to "more of a gut feel, to be honest." (Interviewee #12). That is what we found companies to be doing. They are making ethical judgments based on whether the practice in question "feels right" after considering stakeholder expectations and the Golden Rule.

How to understand this? Given the abundance of available ethical principles, why are these leading companies instead going with what "feels right"? The interviews suggested a number of reasons. To begin with, abstract principles such as "justice," "autonomy," "freedom" and "solidarity"—those that one finds in the Montreal Declaration, EU Ethics Advisory Group report, and other frameworks that we have put in the "moral" category – are too general and subject to interpretation to serve as effective guides to decision-making. They are more likely to tangle decision-makers in debates than lead them to an efficient resolution. As one privacy leader put it: "I don't want to turn everybody into a pointy headed philosopher, and we wouldn't

get anywhere, right? That was a little bit of a concern when we first started talking about this internally, was that we would get into some kind of analysis paralysis, we'd never move things along, and things would always get stuck in data governance. . . . We want to keep things moving, right? Innovation doesn't mean we just sit here." (Interviewee #16).

The informal standards that companies use – public expectations, the Golden Rule – are themselves open to interpretation. But people can more readily apply these standards based on their own experience. "What would *I* expect?" "How would *I* want to be treated?" That is a way of framing the question that can produce a relatively quick and useful resolution, even if not a philosophically grounded one. Informal standards are thus more practical than abstract ethical principles.

They are also more accessible to corporate employees who frequently lack formal training in philosophy or ethics. "If you say, 'Hey, have some ethical thoughts,' they're not going to know what that means because they are not ethically trained. So that's when you say, 'Hey, just think about the 'what if' questions. Like, what if this project does this? And what if this project actually does not do this for that population? Is that fair?' You're putting ethics questions into their heads without telling them they're ethics. And that's the trick." (Interviewee #14).

Informal standards also align well with the purposes behind corporate data ethics initiatives, which may be the main reason that companies adopt them. As was discussed above, most companies view data ethics as a form of beyond compliance risk mitigation. They pursue it in order to be seen as trustworthy and responsible, and so to protect their reputations and reduce the threat of regulation. Conforming to people's and regulators' expectations, and living by the Golden Rule, are good ways to show that one is responsible and trustworthy. Informal, expectation-based standards thus align with, and serve the purpose behind, data ethics initiatives. Public-facing, broad statements of principle also connote trustworthiness and responsibility. That

may be why companies adopt them while, at the same time, relying on more informal standards for the actual decision-making.

The key challenge for such an approach is drawing the line between acceptable and unacceptable risk. Because responsible decision-making in the beyond compliance domain requires sensitivity to, and balancing and weighing of a wide range of ethical risks, even the risk management approach cannot avoid consulting substantive principles, public expectations, the Golden Rule, intuitive "feel", or some other standard in order to guide judgments.

C. RISK MANAGEMENT FRAMEWORKS

A few data ethics managers and consultants embrace the risk mitigation function more expressly. They frame data ethics as a form of risk management. "I think the path through this is, we'll call it ethical, call it responsible, call it fair, whatever word it is, it's being able to design and implement responsible data practices that include an impact assessment on individuals or, quite frankly, a risk assessment as to the individual, as the receiver of that risk." (Interviewee #7).

Generally, risk management is defined as the identification, evaluation, and prioritization of risks followed by an economical application of resources to minimize those risks (Hubbard 2009). The interviewees expanded on this basic concept in two ways. First, they emphasized the importance of considering the *benefits* of a given data analytics project, in addition to its risks, and then of balancing the two. As one explained, "the risk management tools that I implement with organizations do benefits, minus inherent risks, minus controlled ineffectiveness to get at a net benefit risk score." (Interviewee #7).This approach is reminiscent of the Future of Privacy Forum's (FPF) 2014 Report, *Benefit-Risk Analysis for Big Data Projects (*Polonetsky, Tene and Jerome 2014). In this report the FPF, a privacy think tank largely supported by contributions from its corporate members, emphasized the importance of considering a project's benefits along with

its risks. It suggested that companies first identify the benefits of a given data analytics project; then evaluate the project's risks; then consider how to mitigate these risks; and, finally, balance the benefits against the mitigated risks. If the mitigated risks outweigh the benefits, drop the project. If the benefits were greater than the mitigated risks, proceed.

Second, the interviewees stressed the importance of considering impacts, not only on the company and its customers, but on a much broader array of stakeholders. "I would say it starts with first thinking about the actual individuals that are affected by the decisions you make . . . . That is not necessarily part of the mindset when people are just thinking about compliance . . . . Whereas an ethical approach is much more centered on who's affected by this, what are the risks, and what are the harms, but what also are the benefits . . . ? So, it's a weighing of what I'll call risks and harms and benefits and the different stakeholders." Data ethics, particularly when framed as risk management, gets the company to think about impacts on stakeholders that it might not otherwise have considered.

Only a few of our interviewees expressly mentioned the risk management approach. Given the close alignment between risk management and the risk mitigation goal behind corporate data ethics initiatives, we expect to see more companies adopt this approach to substantive line drawing and begin to build the risks of engaging in advanced analytics into their broader risk management efforts.

D. FORMAL PRINCIPLES IN ACTION

As was mentioned above, three of the fourteen corporate interviewees said that their company had established a formal set of principles to guide their use of advanced analytics and AI. One privacy manager at a major health care company was quite explicit about the need to move beyond informal judgment calls to principle-based decisions.

> *[When] we look at things through an ethical lens, we really do try to apply a principled approach. . . . I'm in the stages right now of drafting our code of data ethics for the organization, because people do need to see ... they need to see some enumerated framework, right? When we go into our data governance meetings, what does that mean, right? [W]e provide, again, principles of ethics to consider, as opposed to just saying, "Does this feel right? Does it not feel right?" I think that's where ethics sometimes gets stuck, because folks don't know how to think ethically, and I don't mean that in a disparaging way. It's not to say we can't be moral thinkers, but what does that mean in terms of data?* (Interviewee #16).

When we dig deeper into this interviewee's approach, however, we see that even they combine these formal principles with more intuitive, user-friendly standards. The interviewee starts from "health care ethics . . . autonomy, beneficence, nonmaleficence, and justice" but recognizes that "[t]hat's not going to mean much to a data scientist." (Interviewee #16). The interviewee then translates these principles into "questions that you might want to ask yourself." (Interviewee #16).

> *[T[he principle of autonomy . . . [w]e've reshaped that a bit to say that when we look at data, we need to continually remind ourselves that there is a human being behind this data. . . . there is a respect for the person who is behind that data. . . . We use the principle of empathy, which is to say, "Let's put ourselves in the shoes of our customers." If you're looking at length-of-stay reports, for example . . . [i]t isn't enough to say they should not be there more than three [days]. We need to look at what are the consequences. . . . when we're looking at drawing inferences from data. So we use the principle of empathy . . . . gathering as much . . . data as you can about this person and apply principles of empathy to it. 'Is it right?' 'Do you feel right about what you're doing?' We've used that principle as well.* (Interviewee #16).

This interviewee begins by making the case for enumerated principles and explaining that informal standards such as "Does this feel right" are insufficient. But neither are the principles themselves sufficient. In operationalizing the principles of autonomy, beneficence, etc., the interviewee first translates them into "empathy" and ultimately invokes "Do you feel right about what you're doing?" Even where there is a desire for enumerated principles, the practical value of informal, intuitive standards asserts itself.

In another example, an interviewee reported that their Silicon Valley company had articulated a set of ethical principles to guide its data practices. "It's pretty basic. It talks about . . . privacy and civil liberties but other things as well, and it has a few basic things like we would never be involved in supporting work that might repress a democratic group, . . . or that represses speech." (Interviewee #10). As with the interviewee from the health care industry, this interviewee almost immediately transitioned into talking about how difficult it can be to operationalize these formal constructions. For example, the interviewee posed the question of whether working with law enforcement in Europe to investigate and prosecute hate speech would count as "working for a group that represses speech?" (Interviewee #10).

The interviewee went on to explain how the company had tried to translate its set of principles into a re-usable set of questions for ethical decision-making, but had to abandon the project after the still-growing list of questions reached thirty-four pages in length.

> *We tried to break it down into a reusable framework of questions and we worked with our advisors to do this, to figure out what questions do we need to ask, what framework do we need to use and we stopped at 34 pages of questions. Because we just realized trying to capture it all in advance wasn't working. Trying to create these redlines in advance, again incredibly difficult (*Interviewee #10).

This account of the difficulty that a company experienced in trying to turn broad principles into usable interrogatories supports the idea, stated above, that companies adopt informal benchmarks because formal ethical principles do not lend themselves to practical decision-making.

Some companies do use high-level rules in a way that seems to work. They identify a set of data-related actions that the company believes to be harmful, and then steer clear of these "no go" areas. For example, one retailer refused to accept customer ethnic codes from third parties (Interviewee #17). A number of companies that collect personal data for marketing purposes (customer data, web surfing data, search data) decided not to sell it to third parties

who might use it for other purposes (Interviewees #17, 19). Some companies decide that, while they will sell data to other commercial entities, they will not sell it to the government.[30] Others, who collect customer data for advertising purposes, decide that they will not use it for other, secondary purposes. These are bright line rules about specific situations, rather than the type of broad concepts (autonomy, equality, etc.) that one finds in the sets of data ethics principles discussed above. But this approach also suggests that principles can inform a company's sense of what not to do, even if they do not easily result in a judgment of what *to* do.

If broad data ethics principles do not lend themselves to practical decision-making, then why are companies adopting them? They may serve a hortatory purpose by setting aspirational goals that inspire employees to think more seriously about data ethics and that communicate to the public that the company takes its data ethics responsibilities seriously. They also play an important role in issue spotting. As one interviewee explained:

> *But I think that the big value [of data ethics principles] is to direct people's attention to issues. There's issue spotting. . . . Given people's backgrounds and interests and expertise, you may be tempted to think narrowly in what you're doing, just in terms of achieving the short-term business goals. And what these principles do, especially if they're made part of corporate culture, is to say I know your job is to come up with ideas that cause more engagement among our members . . . . but here's some other things that you should do at the same time. That's where these principles can do some good* (Interviewee #15).

[30] "A company I talked to a year ago had been approached by the intelligence community for its mobile ad date, it sells this data to its clients, I don't know why the intelligence community wanted it but they said, 'we think this is a really bad idea right?' and I said, 'yeah, it's a really bad idea, it is legal, and they may be able to go get it from your client, but you should not sell it to them, that's going to be viewed as unethical by your customers who don't believe that because they saw your ad or they saw a pixel that the government should have it.' So that's a place where people are drawing clear lines." Interviewee #19.

E. POLICY: THE MISSING MIDDLE LAYER

There is a third alternative that lies between broad, abstract principles and intuitive, expectation-based judgments: corporate policy. Policy can be prescribed from the top. But it can also emerge in a common law fashion when managers, confronted with a difficult question, take broad principles, interpret and apply them based on common sense and "what feels right," and so produce a decision. If captured and compiled, those decisions constitute a growing set of corporate policy in much the same way that judicial decisions create the common law, or administrative adjudications produce agency policy.

The interviews showed a glimmer of such policy development. An interviewee from the pharmaceutical industry and one from the health care industry each explained that their company captures and stores its data ethics decisions and then makes them available as a type of precedent for future decision-making. Over time, such a process should yield a corpus of policy guidance that is far more functional than broad, hortatory principles, and more consistent and unified than case-by-case judgments grounded in gut feeling and ever-changing public expectations.

The interviewee from the pharmaceutical industry explained that their company maintains a set of rules to govern data-related actions, including the use of advanced analytics. An employee who wants to initiate a new project must consult these rules and, where the rules are ambiguous or do not speak to the question, the employee must then consult with a member of the team who is trained to answer such grey area questions. The decision then gets recorded and becomes part of the set of rules that guide future decisions. "[O]nce guidance is provided, it automatically loops back and gets instantiated . . . . It's like case law." (Interviewee #21). The interviewee from the health care industry explained that, once the company has built up such a set of precedents, they speed up the review process. "[S]o there's more, what I will call precedents, to

go off of. If something looks like the one we just looked at in July, [then] you can [follow the precedent and] keep it moving." (Interviewee #16).

An interviewee from a Silicon Valley-based technology company provided a very different picture, describing "ad hoc" decision-making that does not draw on prior precedents:

> *And so that means every time you get this ad hoc decision-making it runs huge risks . . . . [A]re we building a common law here? I don't think we are because we don't necessarily record, . . . I'm not sure we record the nuanced decisions in a way that lets us say "okay, how did we do this in the past." We obviously have a lot of churn, it's a tech company, obviously everybody's young, people start their own business, stuff like that. The institutional knowledge – at [number less than 10] years I'm one of the more senior people at the company now – the institutional knowledge isn't necessarily there. It creates a ton of challenges, how do you actually do this in a meaningful way that you can repeat?"* (Interviewee #10).

This anecdotal evidence suggests that companies in highly regulated, long-standing industries such as pharmaceuticals or health care may have existing organizational structures for making, capturing, and compiling policy precedents that they are utilizing with respect to advanced analytics and data ethics. Newer, Silicon Valley-type companies, which lack these institutional structures and, perhaps, need to move more quickly, may struggle more with policy development in this area. Precedent-based policy, which is both practical and consistent, appears to bridge the gap between impractical aspirational principles and ad hoc intuitive judgments. We expect more companies to produce this middle layer of data ethics policy as the field matures.

Whatever the strategic motivations of the companies in this study, it seems clear to both the participants and the research team that there is no way to build a reputation for the responsible use of people's data without entering thoughtfully into the world of beyond compliance data ethics. Our examination of the interviews and survey results revealed an important distinction that shapes our analysis, namely, the distinction between (1) ethical standards or principles that define particular wrongs (or harms or risks) and (2) standards that define what constitutes responsible decision-making by a company. Any comprehensive, beyond compliance business

data ethics approach will need to offer companies not just an enumeration of substantive ethical principles and their associated harms or risks, but a separate standard or procedure that tells them how to weigh and apply those principles to reflect their moral or social responsibilities in uncertain terrain. Appreciating this distinction ties specific data-related ethical concerns to long-standing debates about corporate obligations in society, and draws attention to the need for effective processes within a company that will allow them to track and meet those obligations. We turn to those now.

VII. MANAGING FOR DATA ETHICS

It is not enough to draw substantive lines. A company must also manage its operations to ensure that it abides by the lines that it has drawn. The interviewees spent the bulk of their time describing the management practices that their companies use to try to achieve this. These management innovations break down into three, main areas: organizational infrastructure; issue spotting; and issue resolution.

A. ORGANIZATIONAL INFRASTRUCTURE

In setting up a data ethics management operation, companies need to decide who within the organization should "own" this area. Who should be responsible for data ethics?

1. Privacy office

Of the companies that we spoke to, the majority assigned this function to a Chief Privacy Officer or some other privacy manager. The interviews suggest the thinking behind this. For some time now, the main risks associated with personal information have been privacy harms. When companies that use data analytics confront new threats from their uses of personal data – bias, manipulation, etc. – they take them to the privacy office. As one interviewee explained about the privacy team that they lead: "We've become the de facto ethics team. We're the people that people come to with far more than just privacy questions, so we end up being a conduit for that. . . . they say 'alright well, these are the sorts of questions this team does, we'll take it to them.'" (Interviewee #10). Statements like this suggest that the companies allocate this role to the Chief Privacy Officer and privacy team more by default than by design.

2. Legal department

Another common choice is the legal or compliance office, units that may, or may not, encompass the privacy office.[31] One interviewee explained that the Legal Department is generally responsible for doing due diligence on uses of data throughout the company. This gives it representation throughout the company and so enables it to spot and process data ethics issues wherever they arise (Interviewee #19). A second interviewee drew a distinction between the Legal and Compliance Departments and explained that Legal was preferable for the data ethics function because it is accustomed to making risk-based judgments under conditions of uncertainty, whereas Compliance is more used to bright-line rules.

> *My area reports up through the law department, which is interesting, because when I originally assumed this role, it was part of compliance . . . It made sense to move under legal, we also wanted to get out of the checkbox kind of compliance thinking. When you think of compliance, you think, "I check the box and I take care of what I need to do." . . . [T]hat's really ... not the appropriate way we want our folks to think about it.* (Interviewee #16).

The survey data suggests that, in most companies, either the legal department or the privacy office (which may, in some companies, be part of Legal), has primary responsibility for managing the ethical issues that the company's use of advanced analytics may create. We asked respondents: "Who in your company has primary responsibility for managing ethical risks associated with big data analytics?" Table 2 displays these results and indicates that the Chief Privacy Officer or a Legal executive have primary ownership for the ethical risks. We asked a follow-up about this person's background (not shown), and over 50% of the specific individuals charged with managing ethical risks have a legal or compliance background.

[31] In the survey, 18.5 percent of respondents indicated that the Legal or Compliance Offices housed the data ethics function.

Table 2: Who in your company has primary responsibility for managing ethical risks associated with big data analytics?

	Percent
No one in particular	10.7
Privacy Officer or similar	32.1
Legal or Compliance executive or manager	32.1
Other high-level officer (e.g., Chief Data Officer)	3.6
Data Ethics Officer or similar	14.3
Other ("combination")	7.1

3. The Chief Data Ethics Officer

An interesting development is the emergence in the past few years of a new executive position related to advanced analytics and customer trust: Chief Data Ethics Officer[32] and, in some cases, the creation of an Office of Data Ethics. In some companies, this function is combined with and incorporates the privacy one. In others, it is distinct. The Data Ethics Officer role is still quite rare. Companies that had made a significant commitment to data ethics management made up our entire interview sample and, due to selection bias, were likely over-represented in our survey sample as well. Yet only 21 percent of the companies in the interview sample had recently created a data ethics officer or similar position, and only 17 percent of those in the survey sample had done so. By contrast, almost ninety percent of survey respondents indicated that their company had a Chief *Privacy* Officer.

Table 3: Does your company have a Chief Data Ethics Officer?

	Percent
No	82.8
Yes	17.2

[32]One company refers to it as the "AI Ethics," rather than "data ethics," function, and makes a group of people, rather than a single individual, responsible for it. Interviewee #2.

Table 4: Does your company have a Chief Privacy Officer?

	Percent
No	10.3
Yes	89.7

The Chief Data Ethics Officer role goes well beyond that of the typical Chief Privacy Officer. To begin with, the Data Ethics Officer is responsible for all data about humans that could harm people, not just personally identifiable information (PII). A Chief Privacy Officer, by contrast, generally focuses on PII. As one former chief privacy officer explained:

> *I've just changed the name of the global program and my title has officially changed. My official title is now [title that includes "Data Ethics"] and I've changed the name of the global program to [name that includes "Data Ethics."] And it is because the way that we've done it at [company] is full accountability of all the data that we process and that we steward. That's a very different thing than ensuring you of just privacy requirements like notice and choice. [The idea that the company] should be comprehensively accountable for the data collection, the data activation, the data transformation, the data distribution, is a very next-generation program. It's always been built on ethics. We've been talking about the program as ethical data use for about five years. Then I, as I say, a few weeks ago, I made the official change. That's our journey (*Interviewee #6).

The data ethics function also goes beyond privacy to encompass responsibility for other data analytics-related risks such as bias, manipulation, labor displacement and many of the other threats described above.

While privacy officers tend to focus on compliance with privacy laws, the data ethics function must focus on beyond compliance solutions since law does not yet address the threats that advanced analytics can pose. One such professional explained that at the beginning of their tenure the CEO said to her: "I want compliance out of your title. This is not about compliance.

This is about customer trust. Let's figure out a new title. So that's the birth of the title." (Interviewee #20). Another expressed a similar evolution:

> *we actually added data ethics last year, so my title and my department changed. . . . if we are to do what we need to do for our customers . . . [w]e need to get folks to think of what privacy means a little differently, that it isn't simply complying with the law or policies, it is looking at things through an ethical lens. Because much of what we're doing with data is . . . in a space that is not occupied by law. . . . [D]ata ethics is getting a primary spot. That's the name of our department now (*Interviewee #16).

4. Philosophers in the corporate ranks

Another personnel-related innovation is the hiring of PhD philosophers onto the privacy and data ethics team. One interviewee, explaining the role that the philosopher plays in their groups, discussed the debate that the company had as to whether to create encrypted communications that the government could not access:

> *[A]t the heart of that is the question, what is the consequences of that, and even that, why do we have government? What is the purpose of government and what happens if we change the fundamental way the world operates by creating this extra-governmental space, and is that good or bad . . . . And so being able to think through those questions and recognize those questions is a big part of what we do. Lawyers . . . our job is to look at the legal implications; engineers' tunnel vision is: "I want something that works fast and effectively," and so philosophers are helpful in dragging us out of those mindsets and thinking about, looking at the broader implications. It's incredibly valuable insight. And we're employing philosophers, which has got to be valuable (*Interviewee #10).

This comment suggests that the data ethics team's need to consider broadly the social implications of advanced information technologies has led to the integration of philosophers trained to think rigorously about such matters.

B. SPOTTING ETHICAL ISSUES

Once companies have created an organizational infrastructure around data ethics, they need to institute processes for spotting and resolving ethical issues. The interviewees described a number of issue spotting practices.

1. Touring the business units

Under the first, which we saw more in fast-paced, Silicon Valley companies, the team with primary responsibility for data ethics (e.g. data ethics office, or privacy office) largely assumes the issue spotting function. This team goes out into the business units to meet with developers, learn about their projects, and help them to spot potential ethical issues. This model gets the ethics team out of its office an into the business units, allowing it to problem-solve and address issues quickly. This may be why faster-paced, Silicon Valley-type companies preferred it. The disadvantage, however, is that it relies on a small group of individuals to spot ethical implications throughout the entire company and so can lead to important issues being missed. One ethics specialist explained just how challenging this can be:

> *[O]ur team is small, there are 12 of us trying to support 2,000 deployments all over the world. I am currently at this year: 250,000 miles on [airline]. We are stretched very thin trying to keep up with everything . . . So in terms of flagging issues it is very spotty, and ad hoc and one of our big worries is something is going to happen that we're missing. And you think about code and how many millions lines of code there is, how many complex, how many little decisions might actually have huge implications, it's difficult to figure out how to scale it in a way that would systematically catch everything."* (Interviewee #10).

2. Hub and spokes

The second approach is to place a junior privacy or ethics professional in each business unit. These professionals are trained to spot ethical issues and, where they are significant and

difficult to resolve, to refer them back to the central ethics team for further evaluation and resolution.[33] One interviewee referred to this as a "hub and spokes" model.

> *[P]rivacy reviews are initially conducted by a privacy manager, which is typically a non-lawyer, sitting in a privacy team within the business. So we have sort of a hub and spoke model, where we have distributed a set of privacy managers who are out there in the business. Close to the business decision makers, close to the engineers, doing the privacy reviews according to the processes and standards that have been developed at the hub, in the center, and distributed it out. They are supposed to flag those issues. And the high-risk issues will get escalated to a legal person, who may then further escalate them to one of the central subject matter experts. . . . So there's a process for initial review, sort of issue spotting escalation. And that often works. . . . [H]aving that process in place is invaluable in that we do get eyes on these things very early, at different levels (*Interviewee #16).

This decentralized, hub-and-spokes approach seems to scale better than the centralized one. It appears to be gaining popularity, particularly among larger, more established companies that have many business units in which such ethical issues might arise.[34]

3. External advisory group

Some companies used an external advisory group to spot issues. Such a group -- made up of privacy advocates, academics, industry people, former regulators, and others -- gave the ethics officers a sense of what others might find troubling and so increased their sensitivity to potential ethical concerns. One referred to this as "pressure test[ing]" the company's future data

---

[33] The "structure, from a management system's perspective, tends to be a privacy function with point people out into the business to make sure that there's good oversight and monitoring and that it ties back into an organizational-wide view." Interviewee #22.

[34] An interviewee described this model as a "growing paradigm, and that is appointing people within each of the business units that are not only liaisons into a centralized privacy office or privacy function, but also they have responsibility for being the first point of review and oversight for whether or not that particular business unit is following the standards that have been established by the organization." Interviewee #22. In one company, the ethics team supplements its own capacity by partnering with the audit group which is already out in the business units. As the lead ethics officer explained, "I work very closely with internal audit. They'll be out doing what they normally do, and they'll see something and say, 'I heard this area's doing X,' and then we can go out, and take a look at it, and bring it into governance." Interviewee #16.

practices from an external perspective. (Interviewee #9). Consulting with the external advisory group also gave the ethics team a way to gauge public expectations and so, consistent with the risk mitigation approach to data ethics, align the company's data practices with these expectations.

In some instances, the external advisory also provided the ethics team with additional leverage for advocating its views within the company. As one privacy and ethics leader put it, "we needed backup. We needed a credible group of people who could provide the really solid [feedback], who we could point to and say 'look, they agree with this analysis' . . . . So that's what it was initially formed as. . . . that's the sort of network we built up to do that . . . primarily academics, but we also wanted to get advocates in there." (Interviewee #10).

Some companies set up sitting, external stakeholder committees. For example, one set up an external advisory board that included leading privacy advocates and academics. This board met regularly during the year and corresponded on a more ad hoc basis through emails. The interviewee explained that "they're under a NDA, [so] we can bounce ideas off them, we can show them deployments, we show them technologies, and get their feedback, so that catches things we might have missed, or gives us a perspective from outside the company which is very helpful." (Interviewee #10).

Other companies used a more ad hoc approach, convening groups of stakeholder experts to address particular issues when they arise. "We have the ability to contact consultants and people on the outside . . . and say, 'We're tackling with this issue, can you help us review this?' When do we do it? . . . [We do it] when we feel like the project is about something that we do not have in-house expertise in. And literally, if we feel like we're probably not the right people to review this, then we can go external." (Interviewee #14). This additional input can be helpful. For example, one interviewee recounted a time "when [a company that ran an Internet search

engine] wanted to know if it was a good idea to give people the option of sharing all their searches on Facebook. And so they convened a consumer panel. They said it would be purely voluntary, but should we even allow it as an option? And the panel unanimously said no – you shouldn't allow people to trap themselves, because while they think there isn't any harm in that, you can come up with a parade of horribles from sharing your searches on Facebook." (Interviewee #23).

The use of external data ethics advisory committees remains relatively uncommon. Even among the companies represented in the survey sample, only eleven percent utilized an external advisory committee for this purpose (Table 5). Certainly, the absence of a formal external committee does not necessarily mean companies are *not* seeking external insight. Indeed, our survey sample consists of companies explicitly involved in external organizations focused on data accountability or other trade associations. Methods of seeking external advice, then, may be much more informal or ad hoc.

Table 5: Does your company use an external advisory committee?

	Percent
No	81.5
Yes	11.1
I do not know	7.4

Companies do need to be thoughtful about who they appoint to such external bodies. Google's appointment of a polarizing figure to such a group provoked such an adverse reaction that the company had to disband the group a week after creating it (Waters 2019).

4. Checklists

In his book The Checklist Manifesto: How to Get Things Right , Dr. Atul Gawande (2009) popularized the idea that checklists can be a useful way for organizations to get their people to operationalize broad concepts and apply them consistently. Many industries and professions,

including medicine, aviation and structural engineering, use them for this purpose. The interviewees indicated that some organizations are beginning to use AI ethics checklists in order to get employees to operationalize and apply AI ethics principles (Interviewee #19). One interviewee described their company's instrument as a "set of interrogatories that we're developing right now to get in front of the analytics teams that are going to be asking for data. It's based on some of [our AI ethics] principles, but they're very simple questions, and they're more reflective. They get folks to think [about AI ethics issues] before they take the deep dive into the data." (Interviewee #16).

The companies in our sample are still at an early stage in their development of AI ethics checklists public and were not able to make them available to us. At this point, a 2020 Microsoft Research article constitutes the best resource for companies or policymakers interested in seeing what such a checklist might look like (Madaio, et al., 2020). These researchers, which included a Carnegie Mellon PhD candidate, conducted semi-structured interviews with fourteen data analytics practitioners to get a general sense of what these data scientists would look for in an AI ethics checklist. They then engaged in an iterative process with 48 practitioners working on a variety of AI systems to co-design a model AI Fairness checklist.

The Microsoft Research team's interviews resonated in some ways with our interview findings. Practitioners explained to the Microsoft Research team that they found abstract AI ethics principles to be hard to put into practice. They viewed checklists as a way to operationalize, and make more concrete, abstract concepts such as AI fairness. The practitioners also highlighted a potential downside to using checklists. They can breed a compliance-oriented mentality in which employees check the required boxes without engaging with the nuanced and context-based questions that AI ethics issues often raise. In their view, checklists were best used to initiate reflection and conversation about issues such as fairness, bias, manipulation or transparency,

rather than to provide discrete technical actions that engineers must follow. This fits with our finding, described above, that companies are coming to see data ethics as a strategic activity focused on improving the customer experience and building trust, and not as a compliance function.

The model AI Fairness checklist that the Microsoft Research team and practitioners co-designed, and which is included at the end of their article, consists of questions to consider, actions to take and items to document at six distinct stages in the product development process : (1) Envision (envisioning or greenlighting meetings); (2) Define (spec or design reviews); (3) Prototype (go/no-go discussions and code reviews); (4) Build (ship reviews); (5) Launch; and (6) Evolve (product reviews). Consisting of six sections, and running almost six pages, the checklist is quite long. But it becomes easier to comprehend when one realizes that it contains several core themes that are repeated throughout the various stages. These are:

- Identify those whom the AI system in question might impact, including particular demographic groups;
- Examine the types of fairness-related harms that the AI system might impose on such stakeholders (e.g. allocation, quality of service, stereotyping, denigration, over- or underrepresentation), how these compare to the system's benefits, and whether there are trade-offs between particular fairness criteria.
- Scrutinize and clarify definitions – of system architecture, datasets, potential fairness-related harms, fairness criteria and metrics – and revise them as necessary to mitigate any fairness-related harms.
- Solicit input from a diverse group of reviewers and stakeholders regarding vision, potential harms, definitions, fairness criteria, datasets, etc.
- Where feasible, test the product with these diverse reviewers so that they can better understand and provide feedback on them.
- Monitor product implementation for deviation from expectations and for anticipated or unanticipated fairness-related harms.
- Revise the vision, definitions, datasets, fairness criteria, prototype, etc. in order to mitigate potential harms.
- If it is not possible to mitigate the potential harms, explore and document why this is the case, future mitigation or contingency plans, and whether it makes sense to proceed with the project at all.
- Revise the system at regular intervals to improve its fairness performance and take account of changing social expectations or norms.

The Microsoft Research model checklist goes into far more detail than we are able to provide here. Companies interested in developing their own checklist should consult the Microsoft Research article (Madaio, et al., 2020).

5. Sparking discussion about data ethics issues

Interviewees explained that regular reflection on and discussion of AI ethics issues can help to build a culture in which people throughout the organization are more likely to spot and raise such issues. The idea is that developers and others need to become sensitized to these issues in order to be able to identify them, and that group discussion is an effective way to build this awareness.

Companies go about building this sensitivity and data ethics culture in various ways. One data ethics manager described their practice of circulating articles and other reports about AI ethics incidents, concepts and solutions. "I'm really big on any article I get on data ethics, distributing it broadly, . . . These are what typically would be a garden variety way of communicating with people. But we're customizing it for data ethics. That's part of my ask from our leadership when they said, "How are we operationalizing this?" Communications is one of my performance goals, actually, so I'm working on it." (Interviewee #16).

The companies that we spoke with have not yet developed their techniques for initiating data ethics discussions in their organizations. For a model of how to go about this, it is interesting to consider that the Omidyar Network recently-released a toolkit for sparking such data ethics discussions: the Ethical Explorer Pack.[35] This toolkit goes well beyond the ethical issues that corporate use of advanced analytics and AI can raise (the focus on this report) and considers a

[35] https://ethicalexplorer.org/ .

much wider range of data ethics risk areas. But companies could adopt its approach for the ethical risks that their use of advanced analytics produces.

The Ethical Explorer pack consists of two components. The Tech Risk Zones resource provides a "card" for each risk areas associated with advanced technologies: surveillance, disinformation, exclusion, algorithmic bias, addiction to technology, loss of control, bad actors and outsized power. The card describes the particular risk areas. The Ethical Explorer Field Pack provides discussion questions and exercises that utilize the Risk Zone cards and that companies can use to spark discussion about the various risk areas and how to avoid them in their products and services. For example, one exercise instructs individual employees to select two or three of the risk zones and then find and read current news articles related to these areas. (One might add that they should circulate these articles to their manager and/or team). A second encourages a team to select a technology, product, or feature they are working on, identify which Risk Zones this project might implicate, and then discuss "anticipating risk questions" such as: "What don't we know about this risk? Who needs to be involved to better understand this risk? What are we going to do about it? Who are the decision-makers that will take action?" Companies interested in provoking discussion about data ethics risks within their organization could draw inspiration from the Ethical Explorer Pack.

6. Peer-to-peer conversations

In a sign of just how important companies find ethical issue spotting to be, interviewees reported the emergence of informal, peer-to-peer, conversations to talk about ethical risks and how to address them. One interviewee who works in the Bay Area described off-the-record meetings of twenty or so privacy professionals to discuss the risks associated with advanced analytics and how best to deal with them. "The whole point is to really have a very genuine conversation about the topic, and a lot of people have started to convene them. . . . there's a lot

of interest and activity around wanting to have these really genuine conversations." (Interviewee #2).

C. ISSUE RESOLUTION

Once a company has spotted an ethical issue, the next step is to make a sound decision about it.

1. Just in time data ethics

Where senior privacy or ethics executives go out into the business units to spot issues they can often decide even difficult issues right away. This is the fastest approach. However, it quickly runs into resource constraints. "[T]he challenge is obviously that's not a process that scales very well. The bigger we get, the more difficult it is to have that in any consistent and meaningful way. So it's an incredible challenge." (Interviewee #10).

2. Triage and escalation

The majority of companies that we spoke with employ the hub-and-spokes approach to issue spotting in which a junior person, located in the business unit, identifies issues and refers the hard ones back to the center. Such companies empower the junior person to make decisions about relatively straightforward ethics issues that arise in their unit, perhaps after a quick consultation with the legal department. However, they require the person to escalate more complex, grey area issues to the more senior and experienced decision-makers at the center.[36]

[36] One privacy and ethics leader explained that the "growing paradigm . . . is appointing people within each of the business units that are not only liaisons into a centralized privacy office or privacy function, but also they have responsibility for being the first point of review and oversight for whether or not that particular business unit is following the standards that have been established by the organization." The internal data ethics board resolves "higher-risk uses of data." Interviewee #22. Another described a two-level process where "we have a number of specialists that evaluate it, send it around to a security person and a legal person and an engineer and we get agreement that it conforms to the rules and we sign off. It goes fast. Some of them are really, really big and we get in a room and it might take two weeks and we storyboard it out on a whiteboard. It takes a bunch of stakeholders. It's a heavy lift because it's something new and

One ethics lead analogized this triage and escalation approach to Institutional Review Boards that declare projects that raise few ethical issues to be "exempt" after only cursory review, and that reserve full IRB review for the more ethically complicated proposals (Interviewee #14). A leading consultant described it as "a basic risk assessment process that has escalate-able decision points relative to the commensurate level of risk." (Interviewee #7). A third interviewee used a medical analogy:

> *If you think about the concept of assessments, it's like a triaged process in an emergency room of the hospital. Somebody comes in, they have cuts and scrapes, I can deal with the cuts and scrapes, I do not have to escalate to a doctor. I don't need to escalate that to the operating room. You have other people come in and they have broken bones that have to be set by a doctor so you move to a second level of assessment to determine what is the right treatment level. You have a third level, a fourth level, then you have a level where the issues require assessment by a full range of people who have multiple skills, who will then decide whether what's being done is legal, fair and just (*Interviewee #1).

3. Cross-functional data ethics committee

Once a complex ethical issue gets escalated, who decides it? Here, again, we see a distinction. Some companies, particularly Silicon Valley firms that emphasize speed and innovation, authorize a senior privacy or ethics official to make these calls. In at least one such company, this official can directly engage the CEO when necessary to reach a resolution (Interviewee #10). This yields a quick, streamlined process in which the senior data ethics officer, backed by the CEO, is empowered to make decisions on behalf of the company.

Most of the companies we spoke with, however, place a cross-functional data ethics committee, rather than an individual, at the center of the decision-making process. Where privacy or ethics managers in the business units confront difficult or novel issues that they cannot comfortably resolve themselves, they refer it to such a committee.

---

massive. Right? But we do about 800 a year. This is not something small. We do this at volume and scale." (Interviewee #6).

> *Many of the more sophisticated organizations . . . have started to set up these ethics review boards within the organization. So it's not just about compliance. It's about thinking through these broader sets of data uses and thinking about whether or not they are meeting the company's standards for appropriate data use, if you will. Those tend to be more focused on areas of the business that are more likely to need them, so analytics groups . . . [or] research groups within organizations (*Interviewee #22).

In one illustrative example, a data ethics committee considered whether the company should sell information technology to a customer who might, in turn, share it with the Chinese government for use in surveilling its population (Interviewee #2).[37]

Internal data ethics committees, while growing in popularity, remained a minority approach at the time of our 2019 survey. Even in survey sample, in which companies that took data ethics seriously were likely over-represented, only 33 percent of the companies used such a committee for formal review of data ethics concerns. Over 47 percent of respondents indicated that their company had only an informal review process, or no process at all. We expect the use of data ethics committees to increase as a growing number of companies confront the risks that their use of advanced analytics can create.

Table 6: What is your company's process for identifying ethical risks?

	Percent
We do not have a process set up currently	18.5
Informal screening or review--by a person or office (such as a data ethics executive or team)	22.2
Formal screening or review--by a person or office (such as a data ethics executive or team)	11.1
Formal screening or review by an internal committee, advisory board, or specialized body (e.g. ethics committee, IRB, etc.)	33.3
Screening or review of another sort: Please specify	11.1
I do not know	3.7

[37] As noted above, this ethical question arose before the United States added these companies to the Entity List and so made such sales to them illegal.

The make-up of the data ethics committee varies from company to company but generally includes representatives from the legal, privacy, security, communications, data analytics and engineering departments, as well as the affected business unit (Interviewees #6, 14). Some companies include individuals from government affairs (Interviewee #19), or from corporate social responsibility, (Interviewee #2). The committee may also seek input from an external advisory board of the type described above or, where necessary, from C-Suite executives including the CEO.

The data ethics committee often operates by consensus, with all members required to agree before an ethically challenging project can move forward (Interviewee #17). The group may tweak the project until all members are comfortable with it.[38] Some data ethics committees have the power to cancel projects or contracts where the committee believes that the risks are too high. In one important example of this approach, publicly reported in the media, Microsoft's AI and Ethics in Engineering and Research (AETHER) Committee vetoed significant sales contracts on ethical grounds and put significant limits on others.[39]

Companies that wish to create a data ethics committee would do well to consider thoughtfully some important design choices. These include:

- What types of expertise does this particular company need on its data ethics committee? Which perspectives are most important?
- Should the committee be able to consult with and get input from an external advisory group?
- Where should the committee be located within the organization? Privacy? Legal? Risk management? Strategy?
- To whom should it report? This person needs to be sufficiently high in the corporate hierarchy for the committee's judgements to carry weight.

---

[38] "I don't want this body . . . the cross functional review team . . . to start voting on things, because it just goes the wrong way, I think. So, I don't think that we've ever approved a project with a bigger team where people have not signed off. So, we essentially, will talk about it for as long as it takes for everybody to be okay with it. And it always happens." Where one member objects, the group "can massage it, we can massage it, we can massage it, and hopefully we'll reach a place where they can say, 'Yeah, okay. Now I'm okay with that.'" Interviewee #14.

[39] Geekwire, *Microsoft is Turning Down Some Sales Over AI Ethics, Top Researcher Eric Horvitz Says,* https://www.geekwire.com/2018/microsoft-cutting-off-sales-ai-ethics-top-researcher-eric-horvitz-says/

- What standards should the committee use in making its decisions?
- Should the committee have the power to cancel projects or contracts, or only to make recommendations?
- Should the committee operate by consensus, or majority vote?
- How should issues be elevated to the committee? What process should be followed? What materials provided for committee consideration?
- How does the company define success for this committee? More ethical products? Fewer "incidents" that damage reputation?

4. Broader themes

Several broader themes emerge from the interviewees' statements about management practices.

a) Streamlined vs. deliberative

To begin with, one can see two basic corporate data ethics management approaches. The first is quick and streamlined. It sends decision-makers out into the business units where they spot issues and make just-in-time data ethics decisions. Where these executives do escalate thorny ethical problems back to the center, they come directly to a senior decision-maker who has a direct line to the C-Suite or CEO and is able to reach quick decisions about even the most complex issues.

The majority approach, however, is more deliberative and structured. It involves a hub-and-spokes approach to issue spotting; triage and escalation with respect to issue resolution; and a cross-functional data ethics committee to consider and reach decisions about the most difficult ethical issues, sometimes with input from an external advisory board. We loosely characterize these as "streamlined" and "deliberative" approaches to data ethics issue spotting and resolution.

Based on the interview data, we hypothesize that faster-paced, Silicon Valley-type companies tend to utilize the streamlined process. This gives them speed. However, it both increases the risk that the company may fail to spot certain ethical issues and arguably decreases the thoroughness, and so the quality, of the company's ethical decision-making. By contrast, more

established companies appear to prioritize decision-making quality over speed. They insist that privacy or ethics officers in the field escalate difficult ethical decisions to the more senior executives at the center. They build a cross-functional data ethics committee to deliberate on and decide these complex issues. This takes longer. But it ensures that each decision is the product of a sustained, multi-perspective debate which can, in the most difficult cases, include referral to and input from an external advisory board. This should yield higher quality decisions.

Based on our rather limited interview sample, we further hypothesize that companies in the most highly regulated industries (e.g., heath care, pharmaceutical, financial, transportation, etc.) are more likely to have deliberative ethics decision-making systems, whereas those in newer, technology-oriented industries disproportionately adopt the streamlined approach. This may be because highly regulated companies are able to take legacy management structures developed for existing regulatory requirements and adapt them for the beyond compliance data ethics function.

Finally, we anticipate that the deliberate approaches will narrow the speed gap when compared to streamlined ones. They are likely to take the precedents that their cross-functional ethics committees produce and turn them into guidance for "spoke" decision-makers operating in the business units. This will, over time, enable the dispersed decision-makers to make more and more decisions, while referring fewer issues back to the committee itself. The speed differential between streamlined and deliberate decision-making processes should thus reduce over time, while the quality difference will remain. This should lead companies, even those in fast-paced industries, to prefer the more deliberate approach over the streamlined one.

b) Internal focus vs. system-wide

The interview data also suggests another important divide in corporate data ethics management processes. Some companies focus their data ethics efforts on their internal

operations. Others look not only at what the company itself is doing, but at the behavior of its suppliers and customers. They seek to achieve data ethics throughout the entire production system and value chain of which they are part.

Data ethics started with an internal focus. Soon after corporations began widely to use "big data" and advanced analytics, academics and privacy managers analogized this corporate activity to human subjects research in the university context. In an influential 2013 article, Professor Ryan Calo argued that companies should establish Consumer Subject Review Boards that would serve the same vetting function as Institutional Review Boards do in the university context (Calo 2013). This article helped to frame corporate data ethics management as a kind of private sector IRB focused on the company's own "human subjects" research.[40] One interviewee recounted that, as they started to build their company's data ethics process, "I really thought about the IRB model." (Interviewee #14).

Several interviewees expressed concerns about using the IRB model for data ethics. For one thing, IRBs in the university setting are notoriously slow. "It's not a fast and flexible system, and in the world of data driven applications, a month can be a killer for a project." (Interviewee #1). An IRB faces internally. It focuses on and considers research projects that bubble up from the company itself. That is a vital function. But, according to some interviewees, it is not sufficient. In today's connected world, one party's misbehavior profoundly impacts its business partner's. "You could have everybody doing the right thing, and you introduce one party into that process, whether it's the supplier of certain data or a processor that does a piece of the whole, and the weakness in that link is what's going to bring the whole thing down. The

[40] For example, in 2013 the Future of Privacy Forum, a leading tech industry think tank, posted an interview with Professor Calo on the topic of Consumer Subject Review Boards, https://fpf.org/2013/08/28/podcast-talking-consumer-subject-review-boards-with-ryan-calo/. In 2015, it hosted a Roundtable titled *Beyond IRB's: Ethical Review Processes for Big Data Research*, https://fpf.org/2015/12/10/beyond-irbs-designing-ethical-review-processes-for-big-data-research/, that was attended by over 60 academics and industry researchers and at which Professor Calo gave a keynote address.

reputational impact ... forget the compliance impact or the business continuity impact or investment impact." (Interviewee #21).

This same interviewee explained that, in order to truly account for important risks and protect its own reputation, a company's data ethics initiative must extend beyond its own ranks to include all entities in its value chain. It must seek to "ensure that each link in a chain or each part of the solution that's provided, that either contributes to or benefits from the predictive analytics, has to subject themselves to a certain competency and a certain set of diligence and a certain moral or ethical commitment to be part of that chain or ecosystem." (Interviewee #21). This suggests another distinction between those companies that focus their data ethics initiatives internally through an IRB model or otherwise; and those that take a system-wide approach that includes their suppliers, business partners and, in some cases, even customers.

c) From Compliance, to Strategy

The growth of data ethics management, as personified by the Chief Data Ethics Officer, may signify a fundamental change in the way that companies manage data-related risks. Traditionally, the Privacy Officer's role was to make sure that the company complied with governing privacy laws. This made the Privacy Officer a type of internal cop, and the privacy function a drag on the business operation, even if a very necessary one. Data ethics is not about compliance. It is about going beyond compliance in order to mitigate risk and maintain the company's reputation as a good steward of people's data. Its goal is to build customer trust, which is essential to business success. That makes it similar to other business units—those focused on quality and reliability, communications, or customer relations—whose ultimate goal is to build and preserve the company's trusted relationships with customers, regulators and other important stakeholders. While corporate staff have tended to view the privacy function as a box that the business units need to check, they are increasingly coming to appreciate the data ethics

as contributing to the core business mission of building trust and goodwill. If privacy was a compliance function then, increasingly, data ethics is a *strategic* activity. One interviewee who had made the change from privacy officer to data ethics officer spoke about the transition in just this way:

> *[The shift from privacy officer to ethics officer] is reflective of a really different way of approaching the subject . . . [R]eframing the whole discussion around customer trust has transformed the way I'm able to talk to the business. Before . . . the goal was to simply to get it by me, to check the compliance function. . . . [Then] I went in and I said, hey, this is about whether our customers trust us. . . . So that was the lens that the business understood. They understood how important it is to keep customer trust. They want more customers. So when I talked to them about the customer experience and customer trust, it completely turned it around. . . . The reality is we're ending up going so much farther and building things that are far superior in terms of the customers experience around privacy. Just because I started with how the business wants to design products and services (*Interviewee #20).

Another interviewee whose position had grown from privacy to data ethics explained the distinction in strikingly similar terms: "Privacy became more of an operational function for the organization. . . . we became an enterprise solution*."* (Interviewee #16).

One interviewee told us that it was neither the compliance, privacy nor legal offices that pushed for the establishment of a data ethics function; it was the *strategy office*, "they were the ones that saw the need for it and created it." (Interviewee #2). The fact that data ethics springs from the strategy group further suggests the changing role of data risk governance from a legal or compliance function to one linked much more closely to enterprise strategy. We anticipate that, in the years to come, more companies will adopt the data ethics function and that it will be the strategy office that drives this change.

VIII. TECHNOLOGICAL SOLUTIONS

Trust is an important lubricant of the modern economy. It is not only a nice thing to have and foster; it has, as the economist Arrow said back in the 1970s, a very important pragmatic value. As noted previously in this Report, trust serves as a central motivation for corporate data ethics generally and, more specifically, for why corporations and organizations are increasingly turning to technological solutions to data ethics issues. In this Section we focus on some of these technological solutions, specifically on data privacy, fairness in AI algorithms, and analytic and management tools that several of our interviewees touch upon as they relate to ethical governance through technology. Process based technology solutions, discussed elsewhere, are outside the scope of this section.

A. DATA PRIVACY AND ANONYMIZATION

Almost all of our interviewees bring up the importance of data privacy and anonymization. Modern technology-based efforts to protect privacy include research efforts on k-anonymity, (Sweeney 2002), l-diversity (Machanavajjhala, et al. 2007) and epsilon-differential privacy (or differential privacy for short) ((Dwork, et al. 2014). The interviews make clear that the last of these, differential privacy, is the current de-facto standard for privacy preserving data analysis.: As one interviewee stated:

> *We have a policy called differential privacy which is heavily utilized by a lot of the largest tech companies but way less utilized by the rest of the world. We basically automated our own implementation of it. What differential privacy does is it provides mathematical guarantees that you could never actually get down to the value of a specific cell based on the answer you're getting from any query. So it's a mathematical way to try to guarantee privacy (*Interviewee #5).

Several interviewees also discussed simpler rule-based aggregation strategies for anonymization and de-identification aligned with the classical notions of k-anonymity. For

example, one, explaining that their company only provides aggregate information to others, explained that the company had rules in place to protect individual privacy. For instance, "[I]f there were only fewer than five people that had a particular issue then we wouldn't share even the analysis, or the outcome, or the summary of that, because we thought it would be too close to identifying a particular set of users." (Interviewee #4). Another interviewee summarized a number of privacy-protective strategies:

> *[T]hose vary from masking, which hides the values of particular cells in a table, in different ways. It can be hashing, it can be generalization, it can be replacing with one specific value. So, for example, replace the last four digits of every social security number . . . with four zero's. . . . We can restrict what . . . data is usable based on times. We could create a rule that says only show the last six months of data. We have what's called minimization polices where you could say, "Only allow users to access statistically representative example of X percentage of this data set." You can create that policy and you can set that percentage to whatever you want (*Interviewee #5).

It is worth noting that privacy and ethics are overlapping, but not synonymous. Organizations that protect privacy can still utilize data in ways that are ethically problematic. As one interviewee put it:

> *I think data ethics is much larger than privacy. It's not just about whether I keep somebody's data private. I can aggregate people's data in ways where they remain private, but they become part of a cohort where predictive analytics uses their data, with respect to that group, in ways that an ethicist might say are not appropriate, or does projects, that while using anonymous data . . . [is] a project that an ethicist might say, "You shouldn't be doing."* (Interviewee #5).[41]

[41]Another interviewee made a similar point: *"I can still misuse data using differential privacy - it doesn't ask me what my query is, it's going to keep me from identifying anybody, but if I use the tool to identify who is gay based on certain data points, differential privacy isn't going to say to me: 'Hey, that's not a good [ethical] research element," even though it would prevent me from identifying any individual.'"* (Interviewee #19).

In short, while tools that support privacy and anonymization -- including ideas like privacy by design (Cavoukian 2012) -- are essential to an ethical corporate data framework, they are not sufficient in themselves to ensure ethical application of advanced analytics.

B. ALGORITHMIC FAIRNESS

Several interviewees brought up the importance of algorithmic fairness, fair data use, and its connection to data ethics governance. (Kleinberg, et al, 2018; Wilkinson, et al., 2016). They point out a clear need for technological solutions to facilitate fair artificial intelligence (AI) and Machine Learning (ML). For example, companies to make an effort to find and use datasets that are both inclusive of marginalized groups (so that the resulting AI does not treat them less accurately or well) and, at the same time, are not themselves shaped by harmful societal bias. One data ethics manager explained how her company tried to address this:

> *These could be questions like does this dataset impact communities of color or- does this data set include information about people who are already disproportionately disadvantaged, that sort of thing. If the data set had the potential to... cause more or less bias, there was an escalation path… A set of specific labels. You know, low, medium and high that has different escalation path as to how or when that data set will be approved for purchase, work, or use (*Interviewee #2).

Some interviewees also point out the need to have fair algorithms that minimize risk and yet have utility, pointing to the inherent pareto-optimality often involved. One interviewee explained that, "*with the Ebola epidemic, there could have been more accurate predictive models of the spread, if the cell phone companies had been willing to share data, but they were so afraid to do so . . . [They were afraid because of] [p]rivacy. An epidemiologist was arguing that, "No, you can't do that because there's a risk of re-identification."* (Interviewee #3).

C. THE CLEAR AND PRESSING NEED FOR EXPLAINABLE ALGORITHMS

Somewhat related, and yet distinct from the notion of fair algorithms is the ability to understand what these complex ML and AI algorithms are exactly doing under the hood. (Samek et al. 2019; Gunning and Aha 2019). Explainability and model transparency facilitates trust and ensures regulatory policies are being met. Several interviewees point this out as a key technology need to support ethical data governance.

> *we get into AI and machine learning, sometimes it's pretty challenging to describe, to understand what transparency means. In the old days you could say, we take an email address and we look at your purchase history. And we decide what products you might be interested in buying based on your past purchase history, and we will send you targeted marketing based on that. That was pretty straight forward. People can understand that, but when you have thousands of data inputs developing and machines discovering correlations that might not be intuitively obvious. Building profiles and customizing a variety of experiences based on that. Not only is it harder to explain, in some cases it might be impossible to explain to that same level, because there is no human who understands what correlations are being drawn (*Interviewee #12).

Interviewees also point out that explainable algorithms are critical to engender trust (Interviewee #9). Additionally, some point out that one may have to go beyond simple explainability and consider explanations as to why models fail and the need to understand risks associated with such failure.

> *usually the first thing they spend most of their time focuses on is explainability . . . A lot of that is, quite frankly, true, but focusing only on explainability, I think, obscures the larger picture. Our point … is risk management and failure mitigation.... From a risk management perspective, there are a variety of different ways and processes and things we can do to help govern these models even if we don't explain them....Think about failure. What does failure mean to you? As a company how would you react? What processes are in place? ….. One of the downsides of models that are hard to explain is when they fail, it's hard to understand why (*Interviewee #5).

D. ALGORITHMIC AUDITING OF DATA USE

The modern economy is increasingly reliant on our ability to generate and store large tracts of data and realize actionable insights from this data. The algorithmic steps by which these insights are discerned and subsequently shared is often complex -- requiring multiple transformative steps. These complex multi-step processes in turn can lead to several sources of risk at each step. The ability to mitigate such risk requires the ability to audit the algorithms to ensure that stated ethical data governance policies and regulatory requirements are being met. (Raji, et al., 2020). While research in this space is still relatively recent, organizations are beginning to examine such ideas.

> *Next year, we're actually going to begin auditing. Part of making sure we're doing what we need to do is to make sure another set of eyes comes in, and we're going to be opening up to audit the request, the conditions, the compliance with those requests, and the assurance the data is being used the way we directed them to use it"* (Interviewee #16).

Some interviewees also talk about the role of data provenance (Buneman et al. 2001) in the development of technological solutions for auditing data use.

> *[W]e characterize all the data sources, understand the data provenance, how we're bringing the data together, how we're transforming it, how we're activating it, what it's going to used for, what the controls are... we go through and we measure for things like hidden bias or hidden discrimination. We measure for accuracy [of algorithms] and accuracy occurs on a continuum. If you're delivering a fraud product, you need to be very accurate. If it's a marketing product, the accuracy is not quite as imperative as it is when you're doing identity authentication.* (Interviewee #6).

E. SYSTEMS TECHNOLOGIES TO ENABLE GOVERNANCE

Several interviewees talked about technical systems that they use to enable or enhance governance of data science. Some employ controlled data warehouses or data lakes (a federated set of datasets including both structured and unstructured data in raw form) (Jarke, et al., 1999;

Ramakrishnan, et al., 2017), as a means of ensuring credentialed accessibility to trusted individuals within an organization.

> *when we talk to the data lake team or the data governance team, this is a point that we make absolutely clear to them, that at no point would this data transfer or be pulled from the data lake environment or from a secondary source from data lake and back into any of the credit systems... There are administrative and physical controls… different users are basically confined to play in their space.* (Interviewee #17).

Another talked about the use of "virtualization . . . virtually representing data without actually using it or copying it" as a way to "represent all data across an organization in one single place" and so provide an effective access control mechanism. (Interviewee #5) (Singh, et al., 2008; Soror, et al., 2007).

Virtualization can also assist in addressing the "reproducibility crisis" in machine learning,[42] the situation in which data scientists often cannot reproduce results across teams. (Interviewee #5). One interviewee explained that increased formalization of machine learning, through virtualization and other means, is essential to establishing governance of it. The question is how to do this without constraining data scientists and stifling creativity.

> *[T]here's a crisis in the world of data science and data scientists are basically incapable of reproducing their results across teams. And that's a product of many things. But largely it's just a product of how informal the worlds they live in are. And so, if one data scientist leaves teams and leaves the organization and another data scientist comes in, frequently they have to start from scratch. And, it's a crisis as you start to rely on the models they're developing more and more. But it is squarely a governance topic in my thinking and it's squarely a governance topic in my thinking because if data scientists can't confidently reproduce what's going on, how can lawyers and governance and people thinking about risk, how can all of these personnel justify these decisions and justify the risks embedded in them? So, I think there has to be, for a variety of reasons, a move towards formalization*

[42] https://petewarden.com/2018/03/19/the-machine-learning-reproducibility-crisis/

> *and the trick is just going to be can organizations do it without over correcting and can they do it without stifling creativity (*Interviewee #5).

In sum, we find that technology support for privacy and anonymization is necessary and essential but is alone not sufficient. We find that both algorithmic fairness as well as explainability of AI and ML algorithms are important to ensure ethical governance. The ability to understand the provenance of data transformations as well as the ability to audit algorithms that process such data is also essential as are systems technology that provision for access control and data virtualization and management.

IX. PURSUING THE SOCIAL GOOD

As companies grapple with the need to go "beyond compliance" in their data ethics practices, some of them have welcomed opportunities to promote the social good, either directly for their customers through their company's services or products, or for the public by working with municipalities or providing information. As one study participant puts it: "there's something intuitive in the idea that this is still everybody's data and that it should somehow and someway benefit everybody." (Interviewee #20). Efforts to promote the social good exemplify the oft-repeated refrain among our interviewees that, in going beyond compliance, their companies are trying to do "the right thing" and not merely embracing morally good options because they enhance customer or public trust, or otherwise advance their business interests.

The main examples of these efforts seem to fall into two baskets. First, data analytics may allow companies to warn individuals of risks or to help them discover opportunities to improve their lives. A well-known and controversial example of this sort is Facebook's suicide prevention program. Facebook used advanced analytics and AI to identify, based on a user's postings, whether that individual was potentially suicidal. In March 2017, the company began scanning users' posts for pre-suicidal signals, and then sharing with the user information about how to obtain support, or, in some cases, notifying emergency responders.[43] In the second category of "social good" projects, companies produce socially valuable information for municipalities to improve their planning or efficiency. Our interviews reveal a number of such examples, often utilizing location data: improving evacuation planning during natural disasters, tracking infectious diseases, or relieving traffic congestion in cities.

[43] https://www.businessinsider.com/facebook-is-using-ai-to-try-to-predict-if-youre-suicidal-2018-12

Companies' involvement in these efforts to promote the social good raise a series of important questions. To what extent are these efforts to promote the social good motivated primarily by the companies' long-term interest in gaining customer or public trust? Will companies continue the efforts at such time as promoting the social good no longer benefits the company? How do such initiatives get started within companies, and how can they be sustained? Finally, are companies morally obligated to try to "do the right thing," or must they simply respect certain moral and legal constraints as they pursue their business interests? These questions have long been at the heart of debates about corporate social responsibility.

The Business Roundtable made headlines in August 2019 when it stated that companies should not focus primarily on shareholder value, but instead should embrace a broader commitment to all stakeholders. Alex Gorsky, Chairman of the Board and Chief Executive Officer of Johnson & Johnson, said about the statement: "It affirms the essential role corporations can play in improving our society when CEOs are truly committed to meeting the needs of all stakeholders."[44] At a minimum, this suggests that moral values and regard for the interests of a broad set of affected individuals (and not just shareholder or executive self-interest bounded only by legal compliance) should inform corporate decision-making in an integrated way. But what are the limits of this commitment, in principle and in practice?[45]

Consider the competing motivations driving many corporate social responsibility (CSR) programs. Philanthropy and other ways of connecting with the community that go beyond core business services or products may be motivated by a desire to promote the social good, but they

---

[44] https://www.businessroundtable.org/business-roundtable-redefines-the-purpose-of-a-corporation-to-promote-an-economy-that-serves-all-americans.

[45] Decisions by major corporations during the COVID epidemic to cut their workforces while continuing to pay out hundreds of millions in dividends to shareholders suggests the Business Roundtable statement has not had much effect. https://www.washingtonpost.com/business/2020/05/05/dividends-layoffs-coronavirus/

are often also motivated by a company's own interest in enhancing its reputation. What might it mean for a company to pursue CSR in the spirit of the Business Roundtable statement? Rangan et al. have argued that although an effective and appropriate CSR program must reflect a company's overall "business purpose and values," it should avoid being consciously directed by narrow business aims. Improved business outcomes "should be spillover, not their reason for being:"

> Some [CSR] initiatives indeed create shared value; some, though intended to do so, create more value for society than for the firm; and some are intended to create value primarily for society. Yet all have one thing in common: They are aligned with the companies' business purpose, the values of the companies' important stakeholders, and the needs of the communities in which the companies operate. These companies, of course, stand in stark contrast to those that are focused solely on creating value for their shareholders (Rangan, et al., 2015).

Our study reinforces the thought that doing "the right thing" is typically meant in this spirit. Companies are cognizant of the need to attend to moral values and to the interests of a broad set of stakeholders, and they may welcome opportunities to pursue the social good independent of any immediate business aim. But the pursuit of the social good cannot be entirely divorced from the company's business purpose. Doing "the right thing," then, typically refers to the company's conscious willingness to go beyond legal compliance into less certain moral territory, to try to live up to what a responsible company is expected to do, and to welcome opportunities where its business purpose and values coincide with the social good.

This idea of the motivation for, and limits, of pursuing the social good for its own sake is reflected in the ambiguities in many study participants' comments on such efforts. Take these examples from two interviewees and a recent report:

> *To achieve loyalty and trust from users while constantly evolving and offering new products and services, companies must do more than implement good data*

> *practices—they must build a culture of privacy and security that embeds and formalizes values of digital dignity and data stewardship, and contributes to the social good (*de Mooy and Yuen 2016).
>
> *Yes, we definitely have [talked publicly about social good projects]. It ranges from … press releases that we have done where we've talked about it or in conjunction with a university or a city or things of that nature. When I speak externally, I always talk about it if the forum presents itself because I think it shows how you can build a program and try to enhance your reputation as a company that cares for data and you develop that trust factor or that transparency factor with the external environment, whether it's your customers or regulators or the press or the media or whoever is part of the audience (*Interviewee #9).
>
> *I do see us as a single corporate culture about putting our customers first and doing the right thing. And giving back to the community and being trustworthy (*Interviewee #20).

These passages highlight the key question of motivation and the limiting case where a company's interest and the social good diverge.

Our study offers some clues about how companies may remain committed to the social good even at some cost to business interests. A lot depends on individuals within the company themselves remaining committed to the projects and offering their time, effort, and leadership. Sometimes this can work in a bottom-up fashion: "I would say it usually starts with individual teams, individual people. And then, they escalate it and say, 'We think that we should do this.' . . . And then, obviously, that goes up the chain ... It started from the bottom up." (Interviewee #14). This in turn suggests that companies that want to remain committed to the social good have reason to bring in employees who will be sensitive to moral issues or other stakeholder concerns. "[H]ire individuals with a background or experience in . . . sociology, ethics, and/or human subject research. . . . Distributing this talent throughout the organization will embed a value of data stewardship throughout the decision-making and review processes." (de Mooy and Yuen 2016, at 17).

But more commonly companies that pursue the social good are able to maintain that stance because of the commitment of their CEO, which in turn informs the corporate culture. Study participants often invoked the influence of leadership on the way the whole company functions.

> *I've heard our CEO bemoan the fact that some of the really large data companies internationally have done virtually nothing to help with their data, and [the CEO] believes that's profoundly wrong (*Interviewee #20).
>
> *I don't know why some companies care about things they're not legally required to care about,. . . As I looked inside of [company], at the time, there was a philosophy in the company around responsibility and doing the right thing. It changed, but it was there. And kind of part of DNA of the company. . . . We had an incredibly strong CEO at the time that cared about responsible practices, and going above and beyond. We had an even stronger general counsel. . . I almost think that it comes down to just this perfect storm of the company's history and philosophy around social responsibility (*Interviewee #2).

It perhaps goes without saying that privately-held companies have even more flexibility to reflect the values of the owner or other leadership.

In summary, then, companies have long recognized claims of corporate social responsibility that require going beyond compliance. As companies enter the world of data ethics, they will encounter many opportunities to benefit their communities through data analytics. Some companies are already looking for these opportunities and they recognize that, in competitive field where customer and public trust is vital, pursuing the social good often coincides with their long-term interests. Whether most companies will integrate moral values and the broader interests of the public into their decision-making *for their own sake*, and not because of the coincidence of morality and interest, remains an open question.

REFERENCES

Balkin, Jack. 2016. "Information Fiduciaries and the First Amendment." *UC Davis Law Review* 49(4):1183–1234.

Barocas, Solon and Andrew D. Selbst. 2014. "Big Data's Disparate Impact." *California Law Review* 104(3):671–732.

Bartley, Tim. 2018. *Rules Without Rights: Land, Labor, and Private Authority in the Global Economy*. Oxford: Oxford University Press.

Solon Barocas and Nissenbaum H. 2014. "Big Data's End Run Around Anonymity and Consent," in *Privacy, Big Data and the Public Good* (Julia Lane et al. eds): 44.

Beauchamp T, Childress J. *Principles of Biomedical Ethics*, 7th Edition. New York: Oxford University Press, 2013.

Beckert, Jens. 2016. *Imagined Futures: Fictional Expectations and Capitalist Dynamics*. Cambridge, MA: Harvard University Press.

Biernacki, Patrick and Dan Waldorf. 1981. "Snowball Sampling: Problems and Techniques of Chain Referral Sampling." *Sociological Methods & Research* 10(2):141–63.

Peter Buneman, Sanjeev Khanna, Wang Chiew Tan. 2001. "Why and Where: A Characterization of Data Provenance." *ICDT 2001*: 316-330

Calo, Ryan. 2013. "Consumer Subject Review Boards: A Thought Experiment." *Stanford Law Review Online* 66: 97 (2013).

Calo, Ryan. 2014. "Digital Market Manipulation." *The George Washington Law Review* 82(4):995–1051.

Cate, Fred H. and Viktor Mayer-Schonberger. 2013. "Notice and Consent in a World of Big Data." *International Data Privacy Law* 3(2):67–73.

Cavoukian, Ann. 2012. "Privacy by Design." *IEEE Technol. Soc. Mag.* 31(4): 18-19

Citron, Danielle Keats. 2008. "Technological Due Process." *Washington University Law Review* 85(6):1249–1313.

Citron, Danielle Keats. 2016. "Big Data Should Be Regulated by 'Technological Due Process.'" *The New York Times*. Retrieved April 23, 2020 (https://www.nytimes.com/roomfordebate/2014/08/06/is-big-data-spreading-inequality/big-data-should-be-regulated-by-technological-due-process).

Citron, Danielle Keats and Frank Pasquale. 2014. "The Scored Society: Due Process for Automated Predictions." *Washington Law Review* 89(1):1–34.

Crawford, Kate and Jason Schultz. 2014. "Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms." *Boston College Law Review* 55(1):93–128.

Cycyota, Cynthia S. and David A. Harrison. 2002. "Enhancing Survey Response Rates at the Executive Level : Are Employee- or Consumer-Level Techniques Effective ?" *Journal of Management* 28(2):151–76.

Cycyota, Cynthia S. and David A. Harrison. 2006. "What (Not to Expect A Meta-Analysis of Top Manager Response Rates and Techniques Over Time." *Organizational Research Methods* 9(2):133–60.

Dastin, Jeffrey. 2018. *Amazon scraps secret AI recruiting tool that showed bias against women*, Reuters (October 9, 2018).

de Mooy, Michelle and Shelten Yuen. 2016. *In Wearable Health: Towards Privacy-Aware Research and Development* 20.

Drosou, Marina, H. V. Jagadish, Evaggelia Pitoura, and Julia Stoyanovich. 2017. "Diversity in Big

Data: A Review." *Big Data* 5(2):73–84.
Dwork, Cynthia and Aaron Roth. 2014. "The Algorithmic Foundations of Differential Privacy." *Found. Trends Theor. Comput. Sci.* 9(3-4): 211-407
Ebeling, Mary F. E. 2016. *Healthcare and Big Data*. New York: Pagrave Macmillan.
Esty, Daniel and Andrew Winston, *Green to Gold* (2006)
European Data Protection Supervisor's Ethics Advisory Committee (2018). "Towards Digital Ethics"
Faber, Jacob W. 2019. "Segregation and the Cost of Money: Race, Poverty, and the Prevalence of Alternative Financial Institutions." *Social Forces* 98(2):819–48.
Faroukhi, Abou Zakaria, Imane El Alaoui, Youssef Gahi, and Aouatif Amine. 2020. "Big Data Monetization Throughout Big Data Value Chain: A Comprehensive Review." *Journal of Big Data* 7(3):1–22.
Felzmann, Heike, Eduard Fosch Villaronga, Christoph Lutz, and Aurelia Tamo-Larrieux. 2019. "Transparency You Can Trust: Transparency Requirements for Artificial Intelligence Between Legal Norms and Contextual Concerns." *Big Data & Society* (January-June):1–14.
Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar. 2020. *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI*.
Floridi, Luciano and Josh Cowls. 2019. "A Unified Framework of Five Principles for AI in Society." *Harvard Data Science Review* 1(1):1–15.
Flyverbom, Mikkel, Ronald Deibert, and Dirk Matten. 2019. "The Governance of Digital Technology, Big Data, and the Internet: New Roles and Responsibilities for Business." *Business and Society* 58(1):3–19.
Fosso Wamba, Samuel, Shahriar Akter, Andrew Edwards, Geoffrey Chopin, and Denis Gnanzou. 2015. "How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study." *International Journal of Production Economics* 165:234–46.
Fourcade, Marion and Kieran Healy. 2017. "Seeing like a Market." *Socio-Economic Review* 15(1):9–29.
Gawande, Atul. 2009. *The Checklist Manifesto: How to Get Things Right*. New York: Metropolitan Books.
Glaeser, Edward L., Hyunjin Kim, and Michael Luca. 2017. *Nowcasting the LocalEconomy: Using Yelp Data to Measure Economic Activity*. NBER Working Paper 24010.
Gunningham, Neil, Robert Kagan and Dorothy Thornton, 2006. "Social License and Environmental Protection: Why Businesses Go Beyond Compliance," *Law and Social Inquiry.*
Gregg, Aaron. 2018. "Microsoft, Amazon Pledge to Work with Pentagon Following Anonymous Online Rebukes," *The Washington Post* (October 26, 2018).
Gunning, David and David W. Aha. 2019. "DARPA's Explainable Artificial Intelligence" *(XAI) Program. AI Mag.* 40(2): 44-58.
Hartzog, Woodrow and Daniel J. Solove. 2015. "The Scope and Potential of FTC Data Protection." *George Washington Law Review* 83(6):2230–2300.
Harwell, Drew. 2018. "Google to Drop Pentagon AI Contract After Employee Objections to the 'Business of War,'" *The Washington Post* (June 1, 2018),
Hellman, Deborah. 2020. "Measuring Algorithmic Fairness." *Viriginia Law Review* 106:Forthcoming.
Herschel, Richard and Virginia M. Miori. 2017. "Ethics & Big Data." *Technology in Society* 49:31–36.
Hirsch, Dennis D. 2011. "The Law and Policy of Online Privacy: Regulation, Self-Regulation, or

Co-Regulation." *Seattle University Law Review* 34(2):439–80.
Hirsch, Dennis D. 2015. "That's Unfair ! Or Is It? Big Data, Discrimination and the FTC's Unfairness Authority." *Kentucky Law Journal* 103:345–61.
Hirsch, Dennis D. 2020. "From Individual Control to Social Protection: New Paradigms for Law and Policy in the Age of Predictive Analytics." *Maryland Law Review.* 79: 439-505.
Hubbard, Douglas. 2009. *The Failure of Risk Management: Why It's Broken and How to Fix It*.
Hutson, Matthew. 2018. "Artificial intelligence faces reproducibility crisis," *Science* 16 Feb 2018 : 725-726.
Information Accountability Foundation. 2015. "Unified Ethical Frame for Big Data Analysis"
Jagadish, H. V., Johannes Gehrke, Alexandros Labrinidis, Yannis Papakonstantinou, Jignesh M. Patel, and Raghu Ramakrishnan. 2014. "Big Data and Its Technical Challenges." *Commun. ACm* 57(7):86–94.
Jarke, M. and M. Lenzerini, Y Vassiliou, eds. 1999. *Fundamentals of Data Warehousing.* Springer-Verlag, Berlin-Heidelberg, Germany. 1999.
Jobin, Anna, Marcello Ienca, and Effy Vayena. 2019. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1:389–99.
Kerry, Cameron. 2019. "Breaking Down Proposals for Privacy Legislation: How Do They Regulate?" (Brookings, March 8, 2019).
Khoury, Muin J. and John P. A. Ioannidis. 2014. "Big Data Meets Public Health." *Science* 346(6213):1054–55.
Kleinberg, Jon, and Jens Ludwig, Sendhil Mullainathan, and Ashesh Rambachan. 2018. "Algorithmic Fairness." *AEA Papers and Proceedings*, 108: 22-27.
Kroll, Joshua A., Solon Barocas, Edward W. Felten, and Joel R. Reidenberg. 2017. "Accountable Algorithms." *University of Pennsylvania Law Review* 165(3):633–706.
Lavalle, Steve, Eric Lesser, Rebecca Shockley, Michael S. Hopkins, and Nina Kruschwitz. 2011. "Big Data, Analytics and the Path From Insights to Value Big Data." *MIT Sloan Management Review* 52(2):21–31.
Madaio, Michael, Luke Stark, Jennifer Wortman Vaughan and Hannah Wallach. 2020. *Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI*.
Machanavajjhala, Ashwin and Daniel Kifer, Johannes Gehrke, Muthuramakrishnan Venkitasubramaniam. 2007. "L-diversity: Privacy beyond k-anonymity." *ACM Trans. Knowl. Discov. Data* 1(1): 3.
March, James G. and Herbert A. Simon. 1958. *Organizations*. New York: Wiley.
Marx, Vivian. 2013. "Biology: The Big Challenges of Big Data." *Nature* 498(7453):255–60.
Mayer-Schonberger, Viktor and Yann Padova. 2016. "Regime Change? Enabling Big Data Through Europe's New Data Protection Regulation." *Columbia Science and Technology Law Review* 17(2):315–35.
McAfee, Andrew and Erik Brynjolfsson. 2012. "Big Data: The Management Revolution." *Harvard Business Review* (October).
Mcdermott, Yvonne. 2017. "Conceptualising the Right to Data Protection in an Era of Big Data." *Big Data & Society* (January-June):1–7.
Mittelstadt, Brent. 2019. "Principles Alone Cannot Guarantee Ethical AI." *Nature Machine Intelligence* 1:501–7.
Mittelstadt, Brent D., Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. "The Ethics of Algorithms: Mapping the Debate." *Big Data and Society* 3(2):1–21.
Mulligan, Deirdre K. and Kenneth A. Bamberger. 2015. *Privacy on the Ground: Driving Corporate Behavior in the United States and Europe*. Cambridge, MA: MIT Press.

Nuaimi, Eiman Al, Hind Al Neyadi, Nader Mohamed, and Jameela Al-jaroodi. 2015. "Applications of Big Data to Smart Cities." *Journal of Internet Services and Applications* 6(25).
O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.
Ohm, Paul. 2010. "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization." *UCLA Law Review* 57:1701–77.
Pasquale, Frank (2016). *The Black Box Society: The Secret Algorithms That Control Money and Information.* CAMBRIDGE: HARVARD UNIVERSITY PRESS.
Prakash, Aseem, "Why Do Firms Adopt 'Beyond Compliance' Environmental Policies," *Business Strategy and the Environment* 10:286-299 (2011).
PwC. 2016. Responsibly Leveraging Data in the Marketplace: Key Elements of a Leading Approach to Data Use Governance.
Raji, Inioluwa Deborah, and Andrew Smart, Rebecca N. White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, Parker Barnes. 2020. *Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. FAT* '20: Conference on Fairness, Accountability, and Transparency*, 2020: 33-44.
Richards, Neil M. and Woodrow Hartzog. 2015. "Taking Trust Seriously in Privacy Law." *Stanford Technology Law Review* 19:431–72.
Richards, Neil M. and Jonathan H. King. 2014. "Big Data Ethics." *Wake Forest Law Review* 49(2):393–432.
Rona-tas, Akos. 2017. "The Off-Label Use of Consumer Credit Ratings." *Historical Social Research* 42(1):52–76.
Rosen, Rebecca J. 2013. "Armed With Facebook 'Likes' Alone, Researchers Can Tell Your Race, Gender, and Sexual Orientation," *The Atlantic* (Mar. 12, 2013)
Rosenberg, Matthew, et al. 2018, "How Trump Consultants Exploited the Facebook Data of Millions," New York Times (Mar. 17, 2018)
Polonetsky, Jules, Omer Tene & Joseph Jerome (2014). *Benefit-Risk Analysis for Big Data Projects.*
Ramakrishnan, Raghu, and Baskar Sridharan, John R. Douceur, Pavan Kasturi, Balaji Krishnamachari-Sampath, Karthick Krishnamoorthy, Peng Li, Mitica Manu, Spiro Michaylov, Rogério Ramos, Neil Sharman, Zee Xu, Youssef Barakat, Chris Douglas, Richard Draves, Shrikant S. Naidu, Shankar Shastry, Atul Sikaria, Simon Sun, Ramarathnam Venkatesan. 2017. "Azure Data Lake Store: A Hyperscale Distributed File Service for Big Data Analytics." *SIGMOD Conference 2017*: 51-63
Rangan, Kasturi, Lisa Chase, and Sohel Karim. 2015). "The Truth About CSR." *Harvard Business Review* 93, nos. 1/2 (January–February 2015): 40–49.
Rothstein, Richard. 2017. *The Color of Law: A Forgotten History of How Our Government Segregated America*. New York: Liveright.
Rubinstein, Ira S. 2013. "Big Data: The End of Privacy or a New Beginning?" *International Data Privacy Law* 3(2):74–87.
Samek, Wojciech, and Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, Klaus-Robert Müller. 2019. "Explainable AI: Interpreting, Explaining and Visualizing Deep Learning." Lecture Notes in Computer Science 11700, Springer 2019, ISBN 978-3-030-28953-9
Sandler, Ronald and Basl, John, *Building Data and AI Ethics Committees* (Northeastern University Ethics Institute & Accenture, 2019)
Selbst, Andrew D. and Solon Barocas. 2018. "The Intuitive Appeal of Explainable Machines." *Fordham Law Review* 87(3):1085–1139.
Schneider, Giulia. 2018. "European intellectual property and data protection in the digital-

algorithmic economy: a role reversal(?)" *Journal of Intellectual Property Law and Practice*, 13(3): 229-237.
Scott Shane and Wakabayashi, Daisuke (2018). "'The Business of War,' Google Employees Protest Work for the Pentagon," New York Times (April 4, 2018),
Singh, Aameek, and Madhukar R. Korupolu, Dushmanta Mohapatra. 2008). "Server-storage virtualization: integration and load balancing in data centers." *SC 2008*: 53
Singleton, Royce A. and Bruce C. Straits. 2010. *Approaches to Social Research*. New York: Oxford University Press.
Soror, A. A., and A. Aboulnaga and K. Salem. 2007. "Database Virtualization: A New Frontier for Database Tuning and Physical Design," 2007 *IEEE 23rd International Conference on Data Engineering Workshop*, Istanbul, 2007, pp. 388-394
Stoyanovich, Julia, Ke Yang, and H. V. Jagadish. 2018. "Online Set Selection with Fairness and Diversity Constraints." *Proceedings of the 21st International Conference on Extending Database Technology* 241–52.
Suddaby, Roy. 2006. "From the Editors: What Grounded Theory is Not," *Academy of Management Journal*, 49(4): 633-642.
Sweeney, Latanya. 2002. "k-Anonymity: A Model for Protecting Privacy." *Int. J. Uncertain. Fuzziness Knowl. Based Syst.* 10(5): 557-570
Tene, Omer and Jules Polonetsky. 2012. "Privacy in the Age of Big Data: A Time for Big Decisions." *Stanford Law Review Online* 64:63–69.
Tene, Omer and Jules Polonetsky. 2014. "A Theory of Creepy : Technology , Privacy , and Shifting Social Norms." *Yale Journal of Law and Technology* 16(1):59–102.
Tene, Omer and Jules Polonetsky. 2016. "Beyond IRBs : Ethical Guidelines for Data Research." *Washington and Lee Law Review Online* 72(3):458–71.
Wachter, Sandra and Brent Mittelstadt. 2019. "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI." *Columbia Business Law Review* (2):494–620.
Wachter, Sandra, Brent Mittelstadt, and Chris Russell. 2018. "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR." *Harvard Journal of Law & Technology* 31(2):841–88.
Waldman, Ari E. 2018a. "Designing Without Privacy," *Houston Law Review* 55(3):659-727.
Waldman, Ari E. 2018b. *Privacy as Trust: Information Privacy for An Information Age.* Cambridge University Press.
Waters, Richard. 2019. "Google Scraps Ethics Council for Artificial Intelligence." *Financial Times* (April 4, 2019).
Whittlestone, Jess, Rune Nyrup, Anna Alexandrova, and Stephen Cave. 2019. "The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions." *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* 195–200.
Wilkinson, M., and Dumontier, M., Aalbersberg, I. et al. 2016. "The FAIR Guiding Principles for scientific data management and stewardship." *Sci Data* 3, 160018.
Yang, Ke, Julia Stoyanovich, Abolfazl Asudeh, Bill Howe, H. V. Jagadish, and Gerome Miklau. 2018. "A Nutritional Label for Rankings." *Proceedings of the International Conference on Management of Data (SIGMOD '18)* 1773–76.
Zarsky, Tal. 2017. "Incompatible: The GDPR in the Age of Big Data." *Seton Hall Law Review* 47(4):2.
Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York, NY: PublicAffairs.
Zwitter, Andrej. 2014. "Big Data Ethics." *Big Data & Society* ((July-December)):1–6.