

# BEYOND THE COMMON RULE: ETHICAL STRUCTURES FOR DATA RESEARCH IN NON-ACADEMIC SETTINGS

JULES POLONETSKY, OMER TENE, & JOSEPH JEROME\*

ABSTRACT ..... 333

INTRODUCTION ..... 334

I. BIG DATA BENEFITS—AND RISKS ..... 337

II. CONTEXTUAL AND NON-CONTEXTUAL DATA USE ..... 339

III. WHAT SHOULD BE COVERED BY A CSRB ..... 341

*A. Data experimentation* ..... 341

*B. Non-contextual data uses* ..... 346

*C. Disparate impact* ..... 349

IV. CSRB STRUCTURE ..... 352

*A. Who is covered* ..... 352

*B. External (FTC-approved) or Internal (Organization-led)* ... 353

V. GUIDELINES FOR CSRB ENGAGEMENT ..... 356

*A. Guiding Substantive Principles* ..... 356

        1. Respect for persons ..... 357

        2. Beneficence ..... 358

        3. Justice ..... 360

        4. Respect for law ..... 361

*B. Promoting Independence and Trust* ..... 362

*C. Documentation and Accountability* ..... 365

CONCLUSION ..... 366

## ABSTRACT

As scientific knowledge advances, new data uses continuously emerge in a wide variety of contexts, from combating fraud in the payment card industry, to reducing the time commuters spend on the road, detecting harmful drug interactions, improving marketing mechanisms, personalizing the delivery of education in K-12

---

\* Jules Polonetsky is Co-chair and Executive Director, Omer Tene Senior Fellow, and Joseph Jerome Policy Counsel at the Future of Privacy Forum. We would like to thank Arielle Brown for excellent assistance with research and drafting.

schools, encouraging exercise and weight loss, and much more.<sup>1</sup>

At corporations, not-for-profits, and academic institutions, researchers are analyzing data and testing theories that often rely on data about individuals. Many of these new uses of personal information are natural extensions of current practices, well within the expectations of individuals and the boundaries of traditional Fair Information Practice Principles.<sup>2</sup> In other cases, data use may exceed expectations, but organizations can provide individuals with additional notice and choice. However, in some cases enhanced notice and choice is not feasible, despite the considerable benefit to consumers if personal information were to be used in an innovative way. This article addresses the processes required to authorize non-contextual data uses at corporations or not-for-profit organizations in the absence of additional notice and choice. Although many of these challenges are also relevant to academic researchers, their work will often be guided by the oversight of Internal Review Boards (which are required for many—but not all— new research uses of personal information).

## INTRODUCTION

One of the defining features of the big data economy is the increase in research taking place outside of universities and traditional academic institutions. With information becoming the raw material for production, more organizations are exposed to—and closely examining—vast amounts of personal data about citizens, consumers, patients, and employees. This includes not only companies in industries ranging from technology and education to financial services and healthcare, but also nonprofit entities, which seek to advance societal causes, and even political campaigns. Such

---

1. See generally EXEC. OFFICE OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES (2014) [hereinafter WHITE HOUSE BIG DATA REPORT], available at [http://www.whitehouse.gov/sites/default/files/docs/big\\_data\\_privacy\\_report\\_may\\_1\\_2014.pdf](http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf).

2. Rooted in the United States Department of Health, Education and Welfare's seminal 1973 report, *Records, Computers and the Rights of Citizens* (1973), the Fair Information Practice Principles are at the core of the Privacy Act of 1974, the Organization for Economic Co-operation (OECD) Privacy Guidelines 1980, and numerous global privacy laws. See Org. for Econ. Co-operation & Dev. [OECD], *Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, OECD Doc. C(80)58/FINAL (Sept. 23, 1980), available at <http://www.oecd.org/internet/ieconomy/oecdguidelinesontheProtectionofPrivacyandTransborderFlowsOfPersonalData.htm>; Council Directive 95/46/EC of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, 1995 O.J. (L 281) 31.

research initiatives, which transcend the scope of activities subject to existing ethical frameworks, should be subject to clear principles and guidelines. An uncertain regulatory terrain could jeopardize the value of important research, which may be perceived as ethically tainted or become hidden from the public domain to prevent scrutiny. Concerns over data ethics could diminish collaboration between researchers and private sector entities, restrict funding opportunities, and lock research projects in corporate contributing to the development of new products without furthering generalizable knowledge.

In its May 2014 report *Big Data: Seizing Opportunities, Preserving Values*, the White House recognized the tremendous value of data research in healthcare, education, energy conservation, law enforcement, and national security.<sup>3</sup> At the same time, the White House acknowledged the negative implications big data could have for privacy and other civil rights.<sup>4</sup> While calling for legislative reforms in areas ranging from consumer and student privacy to electronic surveillance, the White House Report came short of proposing solutions for balancing of big data benefits against risks to privacy and civil liberties. Such solutions are sorely needed by organizations that assess on a daily basis whether to proceed with data experimentation and novel data uses.

This paper picks up where our article, *Big Data for All: Privacy and User Control in the Age of Analytics*,<sup>5</sup> left off, proposing organizational structures for addressing common dilemmas pitting compelling data benefits against associated risks. Increasingly, big data is not only fueling business intelligence but also informing decision-making around some of the world's toughest social problems.<sup>6</sup> It is changing the face of philanthropy and research

---

3. WHITE HOUSE BIG DATA REPORT, *supra* note 1, at 6-7; *see also* EXEC. OFFICE OF THE PRESIDENT, PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT: BIG DATA AND PRIVACY: A TECHNOLOGICAL PERSPECTIVE (2014), *available at* [https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast\\_big\\_data\\_and\\_privacy\\_-\\_may\\_2014.pdf](https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf).

4. WHITE HOUSE BIG DATA REPORT, *supra* note 1.

5. Omer Tene & Jules Polonetsky, *Big Data for All: Privacy and User Control in the Age of Analytics*, 11 NW. J. TECH. & INTELL. PROP. 239 (2013); *see also* Jules Polonetsky & Omer Tene, *Privacy and Big Data: Making Ends Meet*, 66 STAN. L. REV. ONLINE 25 (2013), <http://www.stanfordlawreview.org/online/privacy-and-big-data/privacy-and-big-data>; Omer Tene & Jules Polonetsky, *Judged by the Tin Man: Empowering Individuals in an Age of Big Data*, 11 J. TELECOM. & HIGH TECH. L. 351 (2013); Omer Tene & Jules Polonetsky, *Privacy in the Age of Big Data: A Time for Big Decisions*, 64 STAN. L. REV. ONLINE 63 (2012).

6. Kevin C. Desouza & Kendra L. Smith, *Big Data for Social Innovation*, STAN. SOC. INNOVATION REV., Summer 2014, *available at* [http://www.ssireview.org/articles/entry/big\\_data\\_for\\_social\\_innovation](http://www.ssireview.org/articles/entry/big_data_for_social_innovation).

alike.<sup>7</sup> The benefits of such research accrue not only to organizations but also to affected individuals, communities, and society at large. Though it is often industry—and government—use of big data that have received the most scrutiny, ethical considerations around data use must become standardized throughout society.

When personal information is involved, in some cases a proposed use of data is inappropriate, either due to heightened risks to affected individuals or ethical concerns about societal impact. In other cases, despite attendant privacy risks, society should not forfeit the tremendous promise of big data innovations, nor should the results of data research remain locked away for fear of public backlash or regulatory action. In these cases, ethical analysis should allow data use to proceed.

The review process required of organizations that measure data benefits against privacy risks could build upon existing ethical frameworks, namely the seminal Belmont Report on ethical principles for human subject research in the biomedical and behavioral sciences,<sup>8</sup> and the more recent Menlo Report for computer and information security research.<sup>9</sup> These reports call on researchers to respect the basic *dignity and autonomy* of their subjects, demonstrate *beneficence* by balancing the benefits of data use against any potential harm, and respect *justice* by ensuring that the value of new data research accrues to many different segments of the public.<sup>10</sup>

This article proposes to adapt the Belmont principles to the fast-paced reality of today's data-driven world through the development of what Ryan Calo called "Consumer Subject Review Boards" (CSRBs).<sup>11</sup> Calo suggested that organizations should "take a page

---

7. Lucy Bernholz, *How Big Data Will Change the Face of Philanthropy*, WALL ST. J. (Dec. 15, 2013, 4:00 PM), <http://www.wsj.com/articles/SB10001424052702304243904579197652066923202>.

8. NAT'L COMM'N FOR THE PROT. OF HUMAN SUBJECTS OF BIOMEDICAL AND BEHAVIORAL RESEARCH, BELMONT REPORT: ETHICAL PRINCIPLES AND GUIDELINES FOR THE PROTECTION OF HUMAN SUBJECTS OF RESEARCH (1979) [hereinafter BELMONT REPORT], *available at* <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>.

9. DAVID DITTRICH ET AL., U.S. DEP'T OF HOMELAND SEC., THE MENLO REPORT: ETHICAL PRINCIPLES GUIDING INFORMATION AND COMMUNICATION TECHNOLOGY RESEARCH (Aug. 2012) [hereinafter MENLO REPORT], *available at* <https://www.predict.org/%5CPortals%5C0%5CDocuments%5CMenlo-Report.pdf>.

10. *Id.* The Menlo Report adds a fourth principle, Respect for Law and Public Interest, which requires researchers to engage in legal due diligence, be transparent in methods and results, and be accountable for actions. *Id.* at 11.

11. Ryan Calo, *Consumer Subject Review Boards: A Thought Experiment*, 66 STAN. L. REV. ONLINE 97 (2013), <http://www.stanfordlawreview.org/online/privacy-and-big-data/consumer-subject-review-boards>.

from biomedical and behavioral science” and create small committees with diverse expertise that would operate according to predetermined principles for ethical use of data.<sup>12</sup> His brainchild, the CSRB, has recently found its way into the “discussion draft” the White House circulated of a Consumer Privacy Bill of Rights Act under the title “Privacy Review Board.”<sup>13</sup>

We discuss below a possible framework for CSRBs, including issues for consideration, organizational structure and basic substantive rules. We maintain the label CSRB, as opposed to Institutional Review Board (IRB) or Privacy Review Board, to account for the new institution’s remit, which is broader than oversight of human subject research and transcends privacy considerations to address fairness, equality, and other civil liberties concerns. Such broader ethical issues—which are often linked to data use—are inadequately addressed by traditional privacy principles, yet are increasingly central to debates about data use.

#### I. BIG DATA BENEFITS—AND RISKS

For organizations operating at the cutting edge of technological innovation, it has become clear over the past few years that legal compliance and sound security practices are insufficient to meet consumer expectations and societal demands.<sup>14</sup> Data research, once limited to marketing-oriented A/B testing, now encompasses information about how humans live, eat, sleep, consume media, move about, and behave in the seclusion of their home. It produces inferences about individuals’ tastes and preferences, social relations, communications, transportation, and work habits. It implies pervasive testing of products and services that are an integral part of intimate daily life, ranging from connected home products to social networks to smart cars. At the same time, it presents risks to individuals’ privacy, on the one hand, and to organizations’ legal compliance, reputation, and brand, on the other hand.

With organizations developing vast laboratories for big data research, data ethics have become a critical component of

---

12. *Id.* at 102

13. CONSUMER PRIVACY BILL OF RIGHTS, §103(c) (Administration Discussion Draft 2015), *available at* <https://www.whitehouse.gov/sites/default/files/omb/legislative/letters/cpbr-act-of-2015-discussion-draft.pdf>.

14. J. Trevor Hughes & Omer Tene, *The Truth is Out There: Compliance and Security are not Enough*, PRIVACY PERSPECTIVES (Oct. 3, 2014), <https://privacyassociation.org/news/a/the-truth-is-out-there-for-big-data-privacy-compliance-and-security-are-not-enough>.

organizational governance frameworks.<sup>15</sup> Organizations can no longer view privacy strictly as a compliance matter to be addressed by legal departments or a technical issue handled by IT. Rather, to avert public embarrassment and consumer backlash, they must employ ethical review processes and instill issue-spotting skills in employees throughout the organization.

Some may argue that ethical review boards are merely industry subterfuge intended to subvert legislation. However, if structured right, they could provide a meaningful mechanism to shed light on organizations' decision-making processes and ensure that data practices are not only legal but also fair and ethical. Oversight by an ethical review board would not shield an organization from regulatory action, much less from media coverage. But it would put an organization in a position to explain why its practices are meritorious and create a reviewable audit trail in case of investigation.

When organizations process personal data outside of their original context, individuals may in some cases greatly benefit, but in other cases may be surprised, outraged, or even harmed.<sup>16</sup> Soliciting consent from affected individuals can be impracticable. Organizations collect data indirectly or based on identifiers that do not match individuals' contact details. Moreover, by definition, some non-contextual uses—including the retention of data for longer than envisaged for purposes of a newly emergent use—cannot be foreseen at the time of collection. As Crawford and Schultz write, “how does one give notice and get consent for innumerable and perhaps even yet-to-be-determined queries that one might run that create ‘personal data’?”<sup>17</sup>

Indeed, even the federal rules governing human subject research recognize an exception to informed consent requirements where: (a) research involves no more than minimal risk to the subjects; (b) the waiver will not adversely affect the rights and welfare of the subjects; (c) the research could not practicably be carried out without the waiver or alteration; and (d) whenever appropriate the subjects are provided additional pertinent

---

15. VIKTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, *BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK, AND THINK*, 181-82 (2013) (suggesting that firms employ “internal algorithmists” to vet big data projects for ethical and societal impact).

16. HELEN NISSENBAUM, *PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE* (2009).

17. Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C.L. REV. 93, 108 (2014); see also Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249 (2008).

information after participation.<sup>18</sup>

With the dawning of the Internet of Things, big data has become a tool not only for organizations in the technology and computing sectors but also in retail, banking, healthcare, and even home appliances. Companies are now engaged in the type of research once the mainstay of academic and research institutions. Yet unlike clinical trials, data-driven research is seldom based on information collected from a small subset of consenting participants; its raw material is the vast trove of data found in user databases spanning thousands or even millions of individuals. As the Menlo Report notes:

This environment complicates achieving ethically defensible research . . . . It results in interactions with humans that are often indirect, stemming from an increase in either logical or physical “distance” between researcher and humans . . . . The relative ease in engaging multitudes of distributed human subjects (or data about them) through intermediating systems speeds the potential for harms to arise, and extends the range of stakeholders who may be impacted.<sup>19</sup>

Here, CSRBs could inject a measure of objective, independent, ethical oversight, not captured by bottom-line considerations and isolated from pressures imposed by management and marketing teams.

## II. CONTEXTUAL AND NON-CONTEXTUAL DATA USE

The principle of “respect for context” was introduced into the privacy policy debate by the 2012 White House report, *Consumer Data Privacy in a Networked World*.<sup>20</sup> The White House observed that “[c]onsumers have a right to expect that companies will collect, use, and disclose personal data in ways that are consistent with the context in which consumers provide the data.”<sup>21</sup> The notion of context as a limiting principle for data use has academic pedigree, underlying Helen Nissenbaum’s theory of socially responsible

---

18. 45 C.F.R. §46.116(d) (2009).

19. MENLO REPORT, *supra* note 9, at 3.

20. THE WHITE HOUSE, CONSUMER DATA PRIVACY IN A NETWORKED WORLD: A FRAMEWORK FOR PROTECTING PRIVACY AND PROMOTING INNOVATION IN THE GLOBAL DIGITAL ECONOMY (2012), available at [https://www.whitehouse.gov/sites/default/files/email-files/privacy\\_white\\_paper.pdf](https://www.whitehouse.gov/sites/default/files/email-files/privacy_white_paper.pdf).

21. *Id.* at 1.

technology policy.<sup>22</sup> It is closely related to, but more textured and nuanced than, the cornerstone of American consumer privacy protection, the principle of “notice and choice.”<sup>23</sup> As with notice and choice, however, in certain contexts organizations need an escape valve to authorize uses of data that create compelling benefits with proportionally smaller costs, where soliciting consent from widely dispersed or partially anonymized consumers would be impractical or prohibitively expensive.

In its recent report, *Internet of Things: Privacy & Security in a Connected World*,<sup>24</sup> the Federal Trade Commission (FTC) staff recognized that “applied aggressively, a notice and choice approach could restrict unexpected new uses of data with potential societal benefits.”<sup>25</sup> Yet the staff was reluctant to support out-of-context data uses absent additional consumer consent, without prior adoption of legislative use limitations or widely accepted codes of conduct. The FTC staff pointed to the Commission’s prior enforcement actions, as well as private litigation, as evidence of the risks for consumers and organizations of non-contextual data uses. It noted that under the FTC’s unfairness authority, the Commission could prohibit uses of data that cause or are likely to cause substantial injury to a consumer, where that injury was neither reasonably avoidable by the consumer nor outweighed by a benefit to consumers or competition.

In contrast, the White House’s proposed Consumer Privacy Bill of Rights, while recognizing a context-based privacy framework, introduced a process for balancing data benefits against privacy risks in circumstances when additional notice and consent—positioned in the new bill as “heightened transparency and individual control”—is not practicable.<sup>26</sup> It would establish Privacy Review Boards to weigh and potentially authorize the use of personal information in those cases where benefits outweigh risks. Both industry and the privacy advocacy community have criticized the proposed White House legislation. Consumer advocates claimed that the bill gave organizations too much leeway to evaluate risks

---

22. NISSENBAUM, *supra* note 16.

23. See M. Ryan Calo, *Against Notice Skepticism in Privacy (and Elsewhere)*, 87 NOTRE DAME L. REV. 1027, 1028, 1032 (2012).

24. FTC STAFF REPORT, *INTERNET OF THINGS: PRIVACY & SECURITY IN A CONNECTED WORLD* (2015), available at <http://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf>.

25. *Id.* at 43.

26. CONSUMER PRIVACY BILL OF RIGHTS, *supra* note 13, at §103(c).



and determine the protections that consumers will obtain.<sup>27</sup> Industry argued the bill was too constrained to have practical value, saddling high tech entrepreneurs with red tape.<sup>28</sup>

Without dismissing these concerns, some form of CSRB could provide a useful mechanism for infusing larger ethical and equality considerations into organizations' review of non-contextual data uses. The merits of the CSRB concept deserve broader discussion beyond the legislative debate. Accordingly, this paper invokes the underlying issues involved with the concept as well as the factors that could help make a CSRB a practical solution for challenging data decisions in organizations.

One key question needs to be addressed at the outset: should CSRBs be internal corporate organs or external oversight bodies? On the one hand, organizations would hesitate to pass confidential business decisions to an external body. On the other hand, advocates would not be satisfied with a process that is governed internally and opaque. The feasibility of CSRBs thus hinges on the development of a model that can ensure rapid response and business confidentiality while at the same time guaranteeing transparency and accountability.

### III. WHAT SHOULD BE COVERED BY A CSRB

Which issues must an organization escalate for review and decision by a CSRB? Clearly, a CSRB cannot be charged with second-guessing every operational decision. At the same time, CSRB activities should not be limited strictly to projects involving human subject research. This part explores the categories of data use that would benefit from CSRB review.

#### A. *Data experimentation*

The Belmont principles, which emerged in the 1970s from a U.S. Department of Health, Education, and Welfare commission, address the ethics of human subject research in the fields of biomedical and behavioral science. They require academic researchers who intend to conduct experiments involving individuals to vet their proposal with an IRB comprised of peers and structured according to specific

---

27. Letter from Consumer Groups to President Barack Obama (Mar. 3, 2015), available at <http://www.consumerwatchdog.org/resources/ltrobamagroups030315.pdf>.

28. See, e.g., Press Release, Consumer Elecs. Ass'n, Government Must Not Stifle Innovation While Protecting Privacy (Feb. 27, 2015), available at <http://www.ce.org/News/News-Releases/Press-Releases/2015-Press-Releases/CEA-Government-Must-Not-Stifle-Innovation-While-Pr.aspx>.

federal regulations that apply ethical principles and guidelines for the protection of human subjects.<sup>29</sup>

In 1981, with the Belmont Report as foundational background, the Department of Health and Human Services (HHS) and the Food and Drug Administration revised then-existing human subjects regulations. A decade later, in 1991, the Federal Policy for the Protection of Human Subjects, better known as the “Common Rule,” was published and codified in separate regulations by fifteen federal departments and agencies, including HHS, Department of Education, Department of Commerce and National Institute of Standards and Technology (NIST), and National Science Foundation.<sup>30</sup> Other departments and government agencies, including the Department of Homeland Security and the Social Security Administration, comply with the Common Rule without having issued specific regulations.

Ryan Calo observed that currently, a private company conducting similar experiments, sometimes involving thousands of consumers, faces no such obligations—even where its purpose is to profit at the expense of its research subjects.<sup>31</sup> Yet as Calo himself recognized, “Subjecting companies to the strictures of the Belmont Report and academic institutional review would not be appropriate. Firms must operate at speed and scale, protect trade secrets, and satisfy investors.”<sup>32</sup> Moreover, the Common Rule itself exempts research “involving the collection or study of existing data, documents, records . . . if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.”<sup>33</sup> Hence, big data research based on preexisting de-identified information falls outside the remit of the Common Rule.

There are countless examples of big data research helping to create compelling societal value. For example, in the field of education, a recent report by the National Center for Education Statistics, based on wide-scale national statistics, helped uncover what could otherwise have remained discreet but pervasive race-based discrimination. The report shows that despite comprising

---

29. Protection of Human Subjects, 45 C.F.R. §§ 46.103, 46.108 (2012), *available at* <http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html>.

30. *See id.*; *see also* U.S. DEP’T OF HEALTH & HUMAN SERVS, FEDERAL POLICY FOR THE PROTECTION OF HUMAN SUBJECTS (‘COMMON RULE’), <http://www.hhs.gov/ohrp/humansubjects/commonrule/>.

31. Calo, *supra* note 11, at 101; *see also* Ryan Calo, *Digital Market Manipulation*, 82 GEO. WASH. L. REV. 995, 1046 (2014).

32. Calo, *supra* note 11, at 101.

33. 45 C.F.R. § 46.101(b)(4) (2012).

fifteen percent of all college students in the United States (and 13.1% of the general population), African-Americans obtained just seven percent of the nation's science, technology, engineering and mathematics (STEM) bachelor's degrees, four percent of master's degrees, and two percent of PhDs.<sup>34</sup> The report further demonstrates that even after earning all of those degrees, African American scientists attracted markedly less funding than their white counterparts.<sup>35</sup>

Another recent report, by National School Boards Association, offers novel policy solutions for increasing education rates in the U.S. The report, *Partnerships, not Pushouts*, combines census data with data collected by various organizations to identify factors—known as “pushouts”—that may be responsible for driving young people away from education. Pushout factors can be more common among different segments of the population. For example, school suspensions, considered a major pushout factor, affect one out of five black students and only one out of twenty white students, which may partly explain the large discrepancy between graduation rates of those two groups.<sup>36</sup>

According to another recent report, *Building a GradNation: Progress and Challenge in Ending the High School Dropout Epidemic*, fewer than one in four students with disabilities earns a high school diploma in Nevada, compared to eighty-one percent in Montana.<sup>37</sup> Further, the Report shows that, in Minnesota, just fifty-nine percent of low-income students graduate compared with eighty-seven percent of their wealthier peers. Such striking disparities, which surface as a result of data analysis, help school districts, states, and the federal government craft appropriate policy responses.<sup>38</sup>

In each of these examples, students' information is analyzed to help fight discrimination, an important societal goal, which surely

---

34. Liana Christin Landivar, *Disparities in STEM Employment by Sex, Race and Hispanic Origin*, AM. COMTY. SURVEY REPORTS (Sept. 2013), <http://www.census.gov/prod/2013pubs/acs-24.pdf>.

35. Donna K. Ginther et al., *Race, Ethnicity, and NIH Research Awards*, 333 Sci. 1015 (2011), available at <http://www.sciencemag.org/content/333/6045/1015.full>.

36. NAT'L SCH. BDS. ASS'N, PARTNERSHIPS, NOT PUSHOUTS: A GUIDE FOR SCHOOL BOARD MEMBERS: COMMUNITY PARTNERSHIPS FOR STUDENT SUCCESS 7 (2014), available at [http://www.nsba.org/sites/default/files/reports/Partnerships\\_Not\\_Pushouts\\_Guide.pdf](http://www.nsba.org/sites/default/files/reports/Partnerships_Not_Pushouts_Guide.pdf).

37. ROBERT BALFANZ ET AL., BUILDING A GRAD NATION: PROGRESS AND CHALLENGE IN ENDING THE HIGH SCHOOL DROPOUT EPIDEMIC, ANNUAL UPDATE 4 (2014), available at [http://gradnation.org/sites/default/files/17548\\_BGN\\_Report\\_FinalFULL\\_5.2.14.pdf](http://gradnation.org/sites/default/files/17548_BGN_Report_FinalFULL_5.2.14.pdf).

38. See, e.g., Jules Polonetsky & Omer Tene, *Who is Reading Whom Now: Privacy in Education from Books to MOOCs*, 17 VAND. J. ENT. & TECH. L. (forthcoming), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2507044](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2507044).

justifies some degree of privacy risk. Of course, data experimentation is used not only in the field of education. In a different context, the Internet of Things, with its ubiquitous sensors in the foreground and data analysis machinery in the background, provides a fertile backdrop for data research. By collecting a wide variety of information—including telematics, behavioral, biometric, and location—smart cars provide valuable insights into road safety issues. Manufacturers, technicians, and drivers can use diagnostic and vehicle performance information generated by a car to obtain feedback about how vehicles are performing on the road. Car manufacturers can chart vehicle performance in order to plan safety and performance improvements in the future.<sup>39</sup> At the same time, data capture and analysis impact the privacy interests of drivers whose vehicles increasingly rely on software and data to provide basic services.

The social networking environment is a significant arena for data research. Facebook's recent announcement establishing guidelines, review processes, training, and enhanced transparency for research projects demonstrates the emergence of data-ethics as a crucial component of corporate governance programs.<sup>40</sup> Facebook's move came on the heels of a wave of negative public reaction<sup>41</sup> to the publication of a research paper<sup>42</sup> documenting a large-scale experiment conducted on its user base.<sup>43</sup> In that experiment, researchers sought to learn the effects on users' sentiments of tweaking the dosage of positive or negative comments on their News Feeds. Critics viewed the exercise as a real-life experiment on human

---

39. FUTURE OF PRIVACY FORUM, *THE CONNECTED CAR AND PRIVACY: NAVIGATING NEW DATA ISSUES* 6 (2014), available at [http://www.futureofprivacy.org/wp-content/uploads/FPF\\_Data-Collection-and-the-Connected-Car\\_November2014.pdf](http://www.futureofprivacy.org/wp-content/uploads/FPF_Data-Collection-and-the-Connected-Car_November2014.pdf).

40. Mike Schroepfer, *Research at Facebook*, FACEBOOK NEWSROOM (Oct. 2, 2014), <http://newsroom.fb.com/news/2014/10/research-at-facebook>.

41.

42. Adam D. I. Kramer, Jamie E. Guillory & Jeffrey T. Hancock, *Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks*, 24 *PROC. NAT'L ACAD. SCI.* (June 17, 2014), <http://www.pnas.org/content/111/24/8788.full.pdf>.

43. For an empirical analysis of public reaction to such covert data research, see Stuart Schechter & Cristian Bravo-Lillo, *Using Ethical-Response Surveys to Identify Sources of Disapproval and Concern with Facebook's Emotional Contagion Experiment and Other Controversial Studies* (Oct. 30, 2014) (unpublished draft), available at <http://research.microsoft.com/pubs/220718/CURRENT%20DRAFT%20-%20Ethical-Response%20Survey.pdf> ("Regardless of what processes evolve to govern the set of individuals who must decide whether research is approved or rejected, those tasked with making the decisions will have tough choices. Most of the rules that govern research, such as the requirement for participant consent, give review boards considerable discretion.").

subjects without their knowledge or informed consent. *Forbes* columnist Kashmir Hill lamented what she called “a new level of experimentation, turning Facebook from a fishbowl into a petri dish.”<sup>44</sup> Arthur Caplan wrote that the experiment “should send a shiver down the spine of any Facebook user or anyone thinking about becoming one,” and that it should never have been performed.<sup>45</sup>

Clearly, Facebook would have benefitted from prior review and clearance of the research project by a CSRB.<sup>46</sup> Pursuant to the ensuing backlash, Facebook announced the appointment of “a panel including our most senior subject-area researchers, along with people from our engineering, research, legal, privacy and policy teams, that will review projects falling within these guidelines. This is in addition to our existing privacy cross-functional review for products and research.”<sup>47</sup> Alas, as discussed below, skeptics criticized the Facebook announcement demanding that the company provide additional transparency into the decision making process of its newly minted board.<sup>48</sup>

Data experimentation in this setting could further compelling societal goals. Last year, working with researchers from Johns Hopkins University, Facebook adjusted its profile settings so users could announce their status as an organ donor, or sign up if they were not already registered. Over a single day, the new feature prompted more than 13,000 individuals to sign up as organ donors—more than twenty-one times the daily average. Most observers would agree that increasing organ-donation rates is a laudable goal, but clearly, some kinds of social influence must be considered off-limits or subject to special disclosures.

Of course, research based on customer data is not restricted to

---

44. Kashmir Hill, *Facebook Doesn't Understand the Fuss About Its Emotion Manipulation Study*, FORBES (June 29, 2014), <http://www.forbes.com/sites/kashmirhill/2014/06/29/facebook-doesnt-understand-the-fuss-about-its-emotion-manipulation-study/>.

45. Arthur Caplan & Charles Seife, *Facebook Experiment Used Silicon Valley Trickery*, NBC NEWS (June 30, 2014), <http://www.nbcnews.com/health/mental-health/opinion-facebook-experiment-used-silicon-valley-trickery-n144386>.

46. Michelle N. Meyer, *How an IRB Could Have Legitimately Approved the Facebook Experiment—and Why that May Be a Good Thing*, FACULTY LOUNGE (June 29, 2014), <http://www.thefacultylounge.org/2014/06/how-an-irb-could-have-legitimately-approved-the-facebook-experimentand-why-that-may-be-a-good-thing.html>.

47. Schroepfer, *supra* note 40.

48. Vindu Goel, *Facebook Promises Deeper Review of User Research, but Is Short on the Particulars*, N.Y. TIMES (Oct. 2, 2014), <http://www.nytimes.com/2014/10/03/technology/facebook-promises-a-deeper-review-of-its-user-research.html>.

the social networking space. In *The Facebook Experiment: Gambling? In This Casino?*, we wrote that many companies are engaged in A/B testing to assess users' reaction to subtle changes in interface design or delivery methods.<sup>49</sup> By dividing users into two buckets, a variant and a control, A/B testing online allows organizations to randomly divide the user experience into two and measure any difference in online behaviors between the two experiences. Such testing has long been seen as an essential means to create new products, improve existing features and to sometimes advance scientific research when breakthroughs are reported to the public.<sup>50</sup> In fact, popular myth credits one early A/B test to the famous discovery of the cure for scurvy. In the mid-1700s, the British Royal Navy charged ship surgeon James Lind with finding a way to ward off scurvy among sailors. Lind conducted a clinical test, giving groups of sailors, but not a specially designated control group, various possible remedies, including citrus fruits, vinegar and cider. Eventually he proved that citrus was the answer, and it was thereafter incorporated into all sailors' rations.<sup>51</sup>

In sum, organizations' data research should be subject to structured ethical review procedures to prevent the risk of regulatory arbitrage—academics “laundering” research through a corporation or nonprofit to escape the strictures of academic IRBs. At the same time, by setting practicable oversight procedures that facilitate virtuous fast-moving research projects, policymakers would avoid incentivizing researchers and corporations to withdraw knowledge from the public sphere.

### B. Non-contextual data uses

To address the shortcomings of the existing privacy framework, particularly in view of big data opportunities and the rapid deployment of an Internet of Things, a CSRB could be tasked with considering non-contextual uses of data in the absence of additional notice and choice. It would do so by weighing potential risks, broadly defined, against benefits to various stakeholders as well as possible

---

49. Jules Polonetsky & Omer Tene, *The Facebook Experiment: Gambling? In This Casino?*, RE/CODE (July 2, 2014), <http://recode.net/2014/07/02/the-facebook-experiment-is-there-gambling-in-this-casino>.

50. Brian Christian, *The A/B Test: Inside the Technology That's Changing the Rules of Business*, WIRED (Apr. 25, 2012, 8:47 PM), [http://www.wired.com/2012/04/ff\\_abtesting/](http://www.wired.com/2012/04/ff_abtesting/).

51. JAMES LIND, A TREATISE OF THE SCURVY, IN THREE PARTS: CONTAINING AN INQUIRY INTO THE NATURE, CAUSES, AND CURE, OF THAT DISEASE (Cambridge Univ. Press 2014) (1753), available at <http://www.jameslindlibrary.org/lind-j-1753/>.

mitigating actions.

Examples of non-contextual data uses frequently emerge from the field of social media, where social norms are slow to catch up with the dizzying pace of technological innovation. What is technically possible is often surprising or frowned upon. What may not be illegal can be distasteful. Consider the commotion caused in 2012 by an app called “Girls Around Me,” which mapped and disclosed the location and information of “girls” who publicly “checked-in” through their social networks to locations in that user’s vicinity.<sup>52</sup> Although the app did not violate any privacy settings or surface data that was not otherwise publicly available, its startling re-contextualization of otherwise benign geolocation check-ins caused an outrage that led major social networks to block its access to their APIs.<sup>53</sup>

Quite distinct from a cavalier app developer, British Airways, the venerable global carrier, was castigated in the media for its “Know Me” program, which was intended to provide a more personalized service to frequent fliers.<sup>54</sup> Under this program, airline personnel googled passengers to learn more about their profiles and preferences. Instead of being flattered, passengers were taken aback when flight attendants addressed them with friendly, personal greetings. Similarly, a report that the government employs private companies to do background checks on people applying for the PreCheck security program, including by sifting through applicants’ social media postings, led to media outcry.<sup>55</sup> Trying to explain the backlash against the re-contextualization of personal information, one official suggested that while individuals have become accustomed to companies creating comprehensive profiles about them for marketing purposes, “people are far less comfortable with the prospect of a company sifting through the vast amount of

---

52. Caroline Howard, *Stalking Women Online? 'Girls Around Me' is an Alarming App for That*, FORBES (Mar. 31, 2012, 8:35 PM), <http://www.forbes.com/sites/carolinehoward/2012/03/31/stalking-women-online-girls-around-me-is-an-app-for-that/>.

53. *Id.* For a more recent attempt by a dating app to leverage geolocation proximity, see John Paul Titlow, *Creepy Or Brilliant? This Dating App Is Like Tinder... For Strangers You See in Public*, FAST COMPANY (Mar. 25, 2015), <http://www.fastcompany.com/3044295/fast-feed/creepy-or-brilliant-this-dating-app-is-like-tinder-for-strangers-you-see-in-public>.

54. Tim Hume, *BA Googles passengers: Friendlier Flights or Invasion of Privacy?*, CNN (Aug. 22, 2012, 6:02 AM), <http://www.cnn.com/2012/08/22/travel/ba-google-image-passengers/>.

55. Joe Sharkey, *PreCheck Expansion Plan Raises Privacy Concerns*, N.Y. TIMES (Mar. 9, 2015), [http://www.nytimes.com/2015/03/10/business/precheck-expansion-plan-raises-privacy-concerns.html?\\_r=0](http://www.nytimes.com/2015/03/10/business/precheck-expansion-plan-raises-privacy-concerns.html?_r=0).

commercially available personal data for the purpose of creating a potentially intrusive security profile that would be collected under the aegis of a government agency.”<sup>56</sup>

In certain cases, the results of non-contextual data use can be difficult to foresee, even where individuals *are* provided with ample notice and choice. In a recent *Vox* article, an author writing under a pseudonym recounted his story of buying a personalized genetic test for himself and, as a gift, for his parents through the popular genetic testing service 23andMe.<sup>57</sup> As part of the service, 23andMe offers a social media-type function that connects people based on results of their genomic tests, which suggest that they have common ancestry or are directly related. In a dramatic twist, the author learned through the feature about the existence of a previously unknown half-brother sired in secret by his father. The revelation brought great emotional pain to his entire family, leading eventually to his parents’ divorce.

But non-contextual data uses can also deliver compelling benefits. In *Big Data for All*,<sup>58</sup> we wrote about Dr. Russ Altman, a professor of medicine and bioengineering at Stanford University, who discovered that when taken together, Paxil®—the blockbuster antidepressant prescribed to millions of Americans—and Pravachol®—a highly popular cholesterol-reducing drug—have a dreadful side effect, increasing patients’ blood glucose to diabetic levels. They did so sifting through thousands of de-identified Bing search engine logs to test whether a higher proportion of users who searched for *both* “Paxil” and “Pravachol” also typed in words related to a symptomatic footprint of diabetes (think words such as “sweaty palms” or “fatigue”) than those who searched for just “Paxil” or “Pravachol” separately.

By implementing a novel signal detection algorithm that identifies statistically significant correlations, the researchers were thus able to parse out latent adverse effect signals from spontaneous reporting systems. In 2009, for example, “there were an estimated 15 million prescriptions for paroxetine [Paxil] and 18 million prescriptions for pravastatin [Pravachol] in the United States”; there

---

56. *Id.*

57. George Doe, *With Genetic Testing, I Gave My Parents the Gift of Divorce*, *Vox* (Sep. 9, 2014, 7:50 AM), <http://www.vox.com/2014/9/9/5975653/with-genetic-testing-i-gave-my-parents-the-gift-of-divorce-23andme>; see J. Trevor Hughes, *Consent and Forgetting: What Privacy Pros Can Learn from One Family’s Unexpected Experience*, *PRIVACY PERSPECTIVES* (Sep. 12, 2014), <https://privacyassociation.org/news/a/consent-and-forgetting-what-privacy-pros-can-learn-from-one-familys-unexpected-experience>.

58. Tene & Polonetsky, *supra* note 5, at 245.



were an estimated one million individuals who used both drugs in combination. For these individuals, the work of Altman and his colleagues was potentially life-saving.<sup>59</sup>

In order to help distinguish between meritorious and offensive new data initiatives, CSRBs, comprising internal and external experts, would apply a set of agreed-upon ethical principles and maintain an auditable decision trail.

### C. *Disparate impact*

In an influential blog post, online commentator Alistair Croll wrote, “[b]ig data is our generation’s civil rights issue, and we don’t know it.”<sup>60</sup> Innovative data uses may impact not only individuals’ privacy but also additional rights and civil liberties. In their article *Big Data’s Disparate Impact*, Solon Barocas and Andrew Selbst observed that by definition, data mining is always a form of statistical discrimination; indeed, the very point of data mining is to provide a rational basis upon which to distinguish between individuals.<sup>61</sup> The problem, they posited, is that data mining can reproduce existing patterns of discrimination, inherit the prejudice of prior decision makers, or simply reflect the widespread biases and inequalities that persist in society.

In fact, big data could have the perverse result of *exacerbating* existing inequalities by suggesting that historically disadvantaged groups actually *deserve* less favorable treatment. Moreover, big data research can impact entire populations whose data is not even part of examined datasets. The Menlo Report notes, “An evolved paradigm for applying ethical principles to protect humans who may be impacted by research considers activities having human-harming potential rather than simply looking at whether the research does or does not involve human subjects.”<sup>62</sup>

---

59. See Nicholas Tatonetti et al., *Detecting Drug Interactions from Adverse-Event Reports: Interaction Between Paroxetine and Pravastatin Increases Blood Glucose Levels*, 90 CLINICAL PHARMACOLOGY & THERAPEUTICS 133, 133, 139 (2011); Nicholas Tatonetti et al., *A Novel Signal Detection Algorithm for Identifying Hidden Drug-Drug Interactions in Adverse Event Reports*, 12 J. AM. MED. INFORMATICS ASS’N 79, 79–80 (2011); see generally David N. Reshef et al., *Detecting Novel Associations in Large Data Sets*, 334 SCIENCE 1518, 1520 (2011).

60. Alistair Croll, *Big Data is Our Generation’s Civil Rights Issue, and We Don’t Know It*, SOLVE FOR INTERESTING (July 31, 2012, 12:40 PM), <http://solveforinteresting.com/big-data-is-our-generations-civil-rights-issue-and-we-dont-know-it/>.

61. Solon Barocas & Andrew D. Selbst, *Big Data’s Disparate Impact*, 104 CALIF. L. REV. (forthcoming Feb. 2016), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2477899](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899).

62. MENLO REPORT, *supra* note 9, at 4.

Consider, for example, Boston's adoption of an innovative solution to combat the common municipal problem of road potholes. Boston's municipality introduced "Street Bump," an app using the motion-sensing capabilities of smart phones to automatically report information to the city about the condition of the streets users drive on. When a user's car hit a pothole, their phone recorded the shock and sent it to a data hub, which combined the information from many other phones to pinpoint problem areas on streets to be repaired. Surprisingly, "Street Bump" discovered more potholes in wealthy areas of the city than in poor ones. The result, which could have regressively diverted urban resources from the poor to the rich, was actively brought about by the unequal distribution of smart phones and app usage across the population. Wealthier neighborhoods had more smart phone and app users than poorer ones, causing the discrepancy.

Another example is Latanya Sweeney's research demonstrating that Google queries for black-sounding names were more likely to return contextual advertisements for arrest records than searches for white-sounding names.<sup>63</sup> Sweeney confirmed that it was not that advertisers were paying to place their ads that set out to place their inventory on black-sounding names. Apparently, the fact that black-sounding names were more likely to yield such advertisements resulted from the algorithmic process that Google employs to determine which advertisements to place. While proprietary, the algorithm is known to rely in its placement decisions on the tendency of users to actually click on an ad. Over time, as people click one version of an ad more often than others, the weights assigned by the algorithm change and the ad text getting the most clicks eventually displays more frequently. In other words, the differential delivery of ads simply reflected the kinds of prejudice already held by those exposed to the ads.

When should an organization attempt to eliminate disparities that are generated by algorithms, putting a thumb on the scale to provide more just results? In cases like the above, the decision seems clear. But in other cases, adjusting for biases could easily be considered surreptitious manipulation of social values. Google itself, for example, chose a different path when it came to its organic search results. Users who searched for the word "Jew" obtained results linking to hate groups, presumably reflecting what users who

---

63. Latanya Sweeney, *Discrimination in Online Ad Delivery*, 56 COMM. ACM 44 (2013).

searched for that term tended to click on.<sup>64</sup> Different results would come up in searches for terms like “Jewish” or “Judaism,” reflecting the divergent nomenclature of different user groups. Recognizing this, Google elected not to alter the accuracy of its algorithm. Instead, it provided a disclosure at the top of the page that displayed the search results for “Jew,” explaining the reason for the offensive – but accurate results.

To be sure, algorithmic decision-making cannot solve all that is unequal and unfair in society, but technology can solve certain problems. For example, computer scientist Cynthia Dwork and her colleagues have been working on algorithms that assure racial fairness in decision making.<sup>65</sup> At the same time, Dwork and Deirdre Mulligan expressed concern that privacy solutions could hinder efforts to identify classifications that produce discriminatory outcomes, by limiting the availability of data about sensitive attributes.<sup>66</sup> Indeed, a recent white paper, *Big Data: A Tool for Fighting Discrimination and Empowering Groups*, offers more than a dozen case studies of deploying data mining to unearth and remedy unjust discrimination in areas ranging from segregation in schools to allocation of public works and employment in the high-tech sector.<sup>67</sup>

Who should be charged with conducting disparate impact analyses and what are the criteria for such a review? Whether an organization conducts a formal analysis or simply becomes aware of disparate impact of a product or service, how and when should it act to remedy that result? When does an initial analysis warrant the collection of additional, possibly sensitive data to help advance solutions to societal problems at a cost to privacy? And conversely, when do the privacy risks outweigh the benefits the other compelling societal interests? Clearly, organizations must address disparate impacts that run afoul of antidiscrimination laws. But in many other cases disparate treatment may be reasonable and appropriate. In all of these cases, a CSRB can help navigate

---

64. David Becker, *Google Caught in Anti-Semitism Flap*, CNET (Apr. 7, 2004, 10:27 AM), [http://news.cnet.com/2100-1038\\_3-5186012.html](http://news.cnet.com/2100-1038_3-5186012.html).

65. Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, & Rich Zemel, *Fairness Through Awareness*, arXiv:1104.3913 [cs.CC] (2011), <http://arxiv.org/pdf/1104.3913v2.pdf>.

66. Cynthia Dwork & Deirdre K. Mulligan, *It's Not Privacy, and It's Not Fair*, 66 STAN. L. REV. ONLINE 35 (2013), <http://www.stanfordlawreview.org/online/privacy-and-big-data/its-not-privacy-and-its-not-fair>.

67. FUTURE OF PRIVACY FORUM, *BIG DATA: A TOOL FOR FIGHTING DISCRIMINATION AND EMPOWERING GROUPS* (2014), available at <http://www.futureofprivacy.org/wp-content/uploads/Big-Data-A-Tool-for-Fighting-Discrimination-and-Empowering-Groups-Report1.pdf>.

organizational decisions based on a structured process for ethical review.

#### IV. CSRB STRUCTURE

##### A. *Who is covered?*

While the discussion above focuses on examples from corporate data practices, CSRBs could similarly benefit nonprofit entities as well as political campaigns.<sup>68</sup> Nonprofits regularly engage in big data analysis for purposes ranging from improving education to reducing greenhouse emissions. Political campaigns have become data-driven and increasingly “micro-targeted” to highly individualized potential voter groups.

Consider the work done by the Urban Institute, a Washington, D.C.-based think tank, to document continued segregation of black and Hispanic students in public schools.<sup>69</sup> Drawing from the Department of Education’s National Center for Education Statistics, the Urban Institute provides interactive county-level maps that track and visualize public school segregation. The maps aggregate primary and secondary public school enrollment by county and identify where white children predominantly attend majority white schools and where minorities attend schools with predominantly minority classmates. The data is compiled using demographic information and a combination of five school surveys, covering the universe of all free public schools and school districts in the United States. It shows that despite the country’s growing diversity, even extremely diverse regions of the country still maintain segregated school systems.<sup>70</sup> Such big data analysis has significant societal value, yet it also raises privacy questions that should be addressed by a CSRB.

In the past several election cycles, presidential campaigns and other well-funded races for major political offices have become big data operations.<sup>71</sup> Presidential campaign organizations and the two main parties have assembled and maintained extraordinarily

---

68. Cf. CONSUMER PRIVACY BILL OF RIGHTS, *supra* note 13, at §201(c)(2).

69. Reed Jordan, *America’s Public Schools Remain Highly Segregated*, METROTRENDS (Aug. 27, 2014), <http://blog.metrotrends.org/2014/08/americas-public-schools-remain-highly-segregated>.

70. See FUTURE OF PRIVACY FORUM, *supra* note 67.

71. Sasha Issenberg, *How President Obama’s Campaign Used Big Data to Rally Individual Voters*, MIT TECH. REV. (Dec. 16, 2012), <http://www.technologyreview.com/featuredstory/508836/how-obama-used-big-data-to-rally-voters-part-1/>.

detailed political dossiers on practically every American voter.<sup>72</sup> To maximize voter registration, engagement, and balloting, data-driven political campaigns micro-targeted voters based on statistical correlations between their observable patterns of offline and online behavior and the likelihood of their supporting a candidate.<sup>73</sup> Daniel Kreiss warns, “[t]his means that campaigns can develop narrow appeals based on ideology and self-interest and direct them to different groups of voters, appearing to be all things to all people.”<sup>74</sup> These practices raise concerns about loss of voter anonymity, political speech, freedom of association, and the transparency of the political process. Although non-commercial in nature, election data analysis, which fomented fundamental implications for the future of democratic discourse, merits structured review by a CSRB.

*B. External (FTC-approved) or Internal (Organization-led)*

A threshold question is whether a CSRB should be an internal corporate organ or an external body. On the one hand, it would be challenging to create an internal process with sufficient accountability and transparency to make credible decisions that assuage public concerns. On the other hand, it would be equally hard for organizations to hand over a high volume of strictly confidential, sometimes highly technical business decisions to an external decision-making body. The White House legislative effort sidesteps these questions by tasking the FTC with setting down detailed rules and procedures for the operation of CSRBS, including “whether differing requirements are appropriate for Boards that are internal or external to covered entities.”<sup>75</sup>

---

72. Ira Rubinstein, *Voter Privacy in the Age of Big Data*, 2014 WIS. L. REV. 861, 863 (2014).

73. Daniel Kreiss, *Yes We Can (Profile You): A Brief Primer on Campaigns and Political Data*, 64 STAN. L. REV. ONLINE 70 (2012), <http://www.stanfordlawreview.org/online/privacy-paradox/political-data>; Natasha Singer & Charles Duhigg, *Tracking Voters' Clicks Online to Try to Sway Voters*, N.Y. TIMES (Oct. 28, 2012), <http://www.nytimes.com/2012/10/28/us/politics/tracking-clicks-online-to-try-to-sway-voters.html>.

74. Kreiss, *supra* note 73, at 74; *see also* Daniel Kreiss & Philip N. Howard, *New Challenges to Political Privacy: Lessons from the First U.S. Presidential Race in the Web 2.0 Era*, 4 INT'L J. COMM'N 1032 (2010).

75. CONSUMER PRIVACY BILL OF RIGHTS, *supra* note 13, at § 103(e) (“Rulemaking.— Within 180 days after enactment of this Act, the Commission shall promulgate regulations under 5 U.S.C. § 553 to establish the minimum requirements for Privacy Review Boards to qualify for Commission approval, forms and procedures for submission of applications for approval, and a process for review and revocation of such approval. When promulgating regulations under this subsection, the Commission shall consider, among other factors: the range of evaluation processes suitable for covered entities of

This article recommends that any legislative effort to establish CSRBs allow for the creation of both external and internal review boards, which would provide organizations with different levels of legal comfort about their data processing operations. A dual-track approach that permits both options would ensure that CSRBs are scalable to address the needs of both small and large companies and organizations.

*External CSRBs* would be independent entities operating under a clear set of legal and ethical guidelines and subject to formal reporting requirements and regulatory oversight. An external CSRB would be a reasonable option for an organization that lacks the resources, ability, or expertise to develop methodical internal processes. These bodies, which would essentially be an extension of currently existing third-party privacy accountability agents, could serve multiple companies in an industry or sector, thus solving the problem of small- and medium-size enterprises that lack the necessary scale to create an internal CSRB.<sup>76</sup> The FTC could play a guiding role in articulating substantive principles and requirements for CSRB deliberations, and could subject independent, external review mechanisms to detailed reporting requirements. Undergoing an external review process would offer an organization a presumption of legitimacy in cases of challenges to the fairness of data research or non-contextual uses of personal information.

However, a number of important practical limitations confront the implementation of external CSRBs. As discussed below, one significant benefit a CSRB could offer is ongoing monitoring of and consultation about non-contextual data uses over time. In contrast to an internal organ, external CSRBs will likely lack the ability—or capacity—for such ongoing review. Furthermore, with ethical reviews being made in a virtual vacuum, specific decisions may not reflect the full spectrum of risks and rewards, as CSRB members focus on particular scrutinized projects without overall knowledge and understanding of the organization's broader goals and

---

various sizes, experiences, and resources; the range of evaluation processes suitable for the privacy risks posed by various types of personal data; the costs and benefits of levels of independence and expertise; the costs and benefits of levels of transparency and confidentiality; the importance of mitigating privacy risks; the importance of expedient determinations; and whether differing requirements are appropriate for Boards that are internal or external to covered entities. Within 90 days of receipt, following public comment, the Commission shall approve or deny an application for Privacy Review Board approval, and explain in writing the reasons for any denial.”)

76. Todd Ruback, *A Brief Look at Self-Regulation and European Data Protection*, PRIVACY PERSPECTIVES (Feb. 4, 2015), <https://privacyassociation.org/news/a-a-brief-look-at-self-regulation-and-european-data-protection/>.

operations. Moreover, a large number of project approval requests could overwhelm a limited number of external bodies, particularly given that membership in an external CSRB would require broad technical and business expertise. Finally, organizations may have legitimate concerns about confidentiality, intellectual property, and trade secrets when asked to share technical data about their pipeline of innovative projects with an external oversight board.

As a practical matter, large, more established organizations would likely prefer to set up *internal review boards*, yet these would necessarily provide less external transparency. Appropriately, for purposes of any regulatory framework, an internal CSRB would also offer less legal certainty for an organization. An internal CSRB would expand upon existing efforts by the privacy profession to create a culture where privacy is a considered part of doing business, and shift considerations of context away from compliance, notice and choice, toward broader ethical determinations. In addition, the review by an internal board would extend to concerns about fairness and equality that typically exceed the remit of a privacy officer.

To be sure, critics and consumer advocates may not view an internal CSRB as trustworthy or independent. After all, one important goal of a CSRB is enhanced transparency of decision-making processes for the public and regulators. Consequently, internal CSRBS will require mechanisms to ensure accountability, such as detailed documentation requirements to permit regulatory oversight and enforcement *ex post*. In addition, the composition and structure of an internal CSRB would have to be regulated for it to be an effective gatekeeper.

The discussion here echoes the time-honored debates in legal, financial, and economic policy circles concerning the optimal composition of corporate boards of directors.<sup>77</sup> On the one hand, increasing the proportion of outside directors on the board strengthens board independence, enhances transparency, and reduces conflicts of interests. On the other hand, outside directors are inexorably less knowledgeable than corporate officers about the company's operations, goals, and prospects. Consequently, they may be less influential and effective than inside directors in disciplining

---

77. Jeffrey N. Gordon, *The Rise of Independent Directors in the United States, 1950-2005: Of Shareholder Value and Stock Market Prices*, 59 STAN. L. REV. 1465 (2007); but see Sanjai Bhagat & Bernard Black, *The Non-Correlation Between Board Independence and Long-Term Firm Performance*, 27 J. CORP. L. 231 (2002); Laura Lin, *The Effectiveness of Outside Directors as a Corporate Governance Mechanism: Theories and Evidence*, 90 NW. U. L. REV. 898 (1996).

management. More troubling, given their nomination by management and remuneration by the company, outside directors are often not quite as independent as they are made up to be.

To sum, internal CSRBs would require structural guarantees to ensure their members' expertise and independence and prevent capture by management and driven by bottom line concerns. This would include enhanced documentation requirements to allow for regulatory oversight. Approval by an internal CSRB under a documented procedure would not afford an organization a presumption of fairness but would serve as evidence in case of a later challenge to the legitimacy of its data use.

## V. GUIDELINES FOR CSRB ENGAGEMENT

In order to establish trust when implementing a CSRB, organizations will have to embrace a common set of principles and procedures. This section examines how the Belmont principles might be adapted to evaluate innovative data uses, as well as lessons from existing IRB practice and privacy cost-benefit analysis.

### A. Guiding Substantive Principles

Ethical codes traditionally develop in response to disruptions that pose new questions about responsibility, trust, and institutional legitimacy.<sup>78</sup> Increasingly, civil society has come to perceive the vast collection and use of personal information in the age of big data as such a disruptive force.<sup>79</sup> Some of the ethical principles that have emerged as a result of similar disruptions in the past include: (a) a foundational respect for individuals, (b) a careful balancing of risks to individuals against benefits to society, (c) rules for participant selection, (d) independent review of research proposals, (e) self-regulating communities of professionals, and (f) funding contingent on adherence to ethical standards.<sup>80</sup> As discussed above, some of these principles may be poorly suited to evaluate data use at the

---

78. Jake Metcalf, *Ethics Codes: History, Context, and Challenges*, COUNCIL FOR BIG DATA, ETHICS, AND SOC'Y (Nov. 9, 2014), <http://bdes.datasociety.net/council-output/ethics-codes-history-context-and-challenges/>.

79. See, e.g., The Leadership Conference on Civil and Human Rights, *Civil Rights Principles for the Age of Big Data*, 2014, <http://www.civilrights.org/press/2014/civil-rights-principles-big-data.html>; Jeffrey F. Rayport, *What Big Data Needs: A Code of Ethical Practices*, MIT TECH. REV., May 26, 2011, [www.technologyreview.com/news/424104/what-big-data-needs-a-code-of-ethical-practices/](http://www.technologyreview.com/news/424104/what-big-data-needs-a-code-of-ethical-practices/).

80. Metcalf, *supra* note 78.



cutthroat pace of today's business landscape, yet they could guide and inform the values that an internal CSRB try to emulate.

A number of these principles derive from the 1979 Belmont Report, which provides the foundation for the Common Rule, regulating ethical review of government-funded research proposals. Any comprehensive review process needs a benchmark for evaluation. The Belmont principles were an early effort to establish "broader ethical principles [to] provide a basis on which specific rules may be formulated, criticized, and interpreted."<sup>81</sup> The Belmont Report introduced the guiding principles of *respect for persons*, *beneficence*, and *justice*. In 2012, focusing on research in the field of information and communication technologies, the Menlo Report added a fourth principle, *respect for law and the public interest*.<sup>82</sup>

These principles, which were crafted to address the ethics of human subjects research in the fields of biomedical and behavioral science, capture concerns about the potentially ghastly consequences of unfettered research.<sup>83</sup> They require academic researchers who intend to conduct experiments involving individuals to vet their proposals with an IRB comprised of peers, structured according to specific federal regulations, and applying ethical principles and guidelines for the protection of human subjects.<sup>84</sup>

### 1. Respect for persons

The ethical foundation of privacy fits well into the broader principle of *respect for persons*.<sup>85</sup> This principle embraces the need to respect the basic dignity and autonomy of individuals.<sup>86</sup> The Menlo Report expanded this notion to include consideration of data that directly interface, integrate with, or otherwise impact individuals.<sup>87</sup> Dignity and individual autonomy are also the bedrock of privacy rights. More than fifty years ago, Edward Bloustein emphasized the importance of dignity in his classic exposition of the inviolate

---

81. BELMONT REPORT, *supra* note 8.

82. MENLO REPORT, *supra* note 9, at 8.

83. *See, e.g.*, SUSAN M. REVERBY, EXAMINING TUSKEGEE: THE INFAMOUS SYPHILIS STUDY AND ITS LEGACY (2009); EILEEN WELSON, THE PLUTONIUM FILES: AMERICA'S SECRET MEDICAL EXPERIMENTS IN THE COLD WAR (2000); Marcia Angell, *The Ethics of Clinical Research in the Third World*, 337 NEW ENG. J. MED. 847 (1997).

84. Protection of Human Subjects, 45 C.F.R. §§ 46.103, 46.108 (2012).

85. Lawrence O. Gostin, *Ethical Principles for the Conduct of Human Subject Research: Population-Based Research and Ethics*, 19 L. MED. & HEALTH CARE 191 (1991).

86. *Cf.* ALAN F. WESTIN, PRIVACY AND FREEDOM (1970).

87. MENLO REPORT, *supra* note 9, at 9–10.

personality that “defines man’s essence as a unique and self-determining being.”<sup>88</sup> In his work, Bloustein argued that the four privacy torts restated by Dean Prosser all shared a concern with safeguarding human dignity.<sup>89</sup> In his seminal piece, *The Two Western Cultures of Privacy*, James Whitman identified dignity as the unifying concept for European privacy law.<sup>90</sup>

Fittingly, the policy debates around privacy values and practices already closely track the Belmont and Menlo principles. The principle of respect for persons mirrors privacy law’s emphasis on individual autonomy, manifest in the principles of fair notice and informed choice. While notice and choice have historically sufficed to mitigate most legal concerns about the use of personal information, the arrival of new technologies, which reduce the practicability of traditional choice mechanisms, have made the need for separate CSRB reviews more salient. While CSRBS could be charged with considering new mechanisms for offering consumers notice, they would typically focus on exploring non-consensual out-of-context data uses. In doing so, CSRBS should weigh the principles of beneficence and justice in light of existing public policy concerns.

## 2. Beneficence

*Beneficence* requires researchers and their overseers to take a broader view of prospective research. This principle seeks to promote a balance between the benefits of data use against potential harms. Both the Belmont and Menlo Reports recognize the difficulty of this balancing act; the Belmont Report recognizes that “precise judgments” are challenging and further, “only on rare occasions will quantitative techniques be available for the scrutiny of research proposals.”<sup>91</sup> The Menlo Report further accepts the additional challenges in identifying risks and rewards in the digital space, considering the scale and rapidity of information, inability to attribute risks to specific individuals, and limits to how we understand the causal relationship between the online and offline worlds.<sup>92</sup> The Menlo Report states, “Beneficence does not require that all harm be completely eliminated and every possible benefit be

---

88. Edward J. Bloustein, *Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser*, 39 N.Y.U. L. REV. 962, 971 (1964).

89. *Id.*; see William L. Prosser, *Privacy*, 48 CALIF. L. REV. 383 (1960).

90. James Q. Whitman, *The Two Western Cultures of Privacy: Dignity Versus Liberty*, 113 YALE L.J. 1151 (2004).

91. BELMONT REPORT, *supra* note 8.

92. MENLO REPORT, *supra* note 9, at 12.

identified and fully realized. Rather, researchers should systematically assess risks and benefits across all stakeholders.”<sup>93</sup>

Despite these difficulties, beneficence calls for organizations to perform a rational, non-arbitrary, systematic assessment. Thus, CSRBs will need to engage in a broad and holistic consideration of prospective benefits. Our whitepaper, *Benefit-Risk Analysis for Big Data Projects*, proposes a framework for factoring data benefits and risks into a decision-making process, which we call “data benefit analysis.”<sup>94</sup> It introduces data benefit analysis as a two-step process: First, organizations should assess the “raw value” of a benefit, which consists of the *nature* of the benefit, the identity of potential *beneficiaries*,<sup>95</sup> and the *degree* (or size and scope) of the benefit. Second, organizations should discount the raw value score by the probability that the benefit can be achieved to obtain a discounted value score. This score, in turn, should be weighed against the risks a project presents for privacy and individual liberties, reduced by any mitigating actions that can be employed by the organization, such as de-identification, retention limitations, and opt-out rights.

The whitepaper reasons that so far, there has been little analytical work to assess big data benefits in a way commensurate with existing privacy risk frameworks.<sup>96</sup> Seen this way, the discussion is incomplete, since accounting for costs is only part of a balanced cost-benefit equation. In order to complete cost-benefit analysis, organizations need tools to help them assess, prioritize, and—to the extent possible—quantify a project’s rewards. Not all benefits are or should be treated as equal. A potentially big benefit with a high likelihood of success must be treated differently than a smaller benefit with a similarly high likelihood of success—or a big benefit that is unlikely to ever be accomplished.

Although underdeveloped in practical literature, the balancing of benefits against risks is firmly rooted in existing legal doctrine. Section 5(n) of the FTC Act sets up the determination of “unfairness” as a balancing of consumer injury against “countervailing benefits to

---

93. *Id.* at 9.

94. JULES POLONETSKY, OMER TENE & JOSEPH JEROME, FUTURE OF PRIVACY FORUM, BENEFIT-RISK ANALYSIS FOR BIG DATA PROJECTS (Sept. 2014), *available at* [http://www.futureofprivacy.org/wp-content/uploads/FPF\\_DataBenefitAnalysis\\_FINAL.pdf](http://www.futureofprivacy.org/wp-content/uploads/FPF_DataBenefitAnalysis_FINAL.pdf).

95. The CPBR recognizes the importance of benefits outside of an organization when evaluating the privacy impact of non-contextual data uses. The bill would provide an exception from heightened notice and choice where a “Privacy Review Board determines that the goals of the covered entity’s analysis are likely to provide substantial benefits that do not exclusively accrue to the covered entity.”

96. PRIVACY IMPACT ASSESSMENT (David Wright & Paul de Hert eds., 2012).

consumers or to competition.”<sup>97</sup> Similarly, the European Article 29 Data Protection Working Party applied a balancing test in its opinion interpreting the “legitimate interest” clause of the European Data Protection Directive. The Article 29 Working Party warns against decision-making based on mechanical weighing of benefits against risks. Rather it stresses that factors to consider when carrying out a balancing test include: (a) the impact on the individual and “their reasonable expectations about what will happen to their data, as well as the nature of the data and how they are processed,” (b) the interest of the public and any affected community, and (c) additional safeguards which could limit undue impact on the individual, “such as data minimization, privacy enhancing technologies, increased transparency, general and unconditional right to opt-out, and data portability.”<sup>98</sup>

### 3. Justice

*Justice*, in short, would require a CSRB to take fairness and equity into account and consider all stakeholders’ interests, ensuring that the value of research accrues to different segments of the public.<sup>99</sup> In the context of big data, the Menlo Report suggests that the principle of justice implies that research must “not arbitrarily target persons or groups based on attributes including (but not limited to): religion, political affiliation, sexual orientation, health, age, technical competency, national origin, race, or socioeconomic status.”<sup>100</sup> In other words, a CSRB should engage in disparate impact analysis, focusing on whether vulnerable populations, such as minorities, kids, the elderly, the disabled, or those suffering from health conditions, are targeted or affected by a project.

---

97. 15 U.S.C. § 45(n). A similar analysis is required under the European Data Protection Directive and the new draft Regulation. For example, the European Parliament has proposed exceptions to consent requirements where data use “serves high public interests, cannot possibly be carried out otherwise, and other safeguards are applied.” Article 29 Working Party, Annex – Health Data in Apps and Devices, Feb. 5, 2015, [http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2015/20150205\\_letter\\_art29wp\\_ec\\_health\\_data\\_after\\_plenary\\_annex\\_en.pdf](http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf).

98. Article 29 Working Party, WP 217, Op. 06/2014 on the Notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, Apr. 9, 2014, [http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217\\_en.pdf](http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf).

99. This principle finds additional support in the recent White House Consumer Bill of Rights, which directs that Privacy Review Boards ensure that any benefits do “not exclusively accrue” to the testing organization. CONSUMER PRIVACY BILL OF RIGHTS, *supra* note 13, at § 103(c)(2).

100. MENLO REPORT, *supra* note 9, at 14.

For entities that operate in an environment governed by fair lending and employment law, CSRBs are well positioned to consider whether a project could have disparate impact on specific populations—particularly if empowered to monitor a project over time.<sup>101</sup> Yet, as discussed above,<sup>102</sup> it is arguable that current law cannot adequately address disparate impacts in data, and further, that existing technology may not detect—or may even perpetuate—existing imbalances and unfair discrimination. But while stakeholders pursue both technological solutions and best practices around algorithmic discrimination, internal CSRBs are an important mechanism to weed out discriminatory impacts and ensure fair deployment of new technologies.

At the same time, as has become evident in other contexts, data limitations could have the perverse effect of constraining regulators' ability to enforce against unfair discrimination as well as industry's capacity to respond to discriminatory decision-making.<sup>103</sup> A CSRB may need to collect external data or otherwise observe project data flows, sometimes over an extended time period, in order to make a proper evaluation and determination of decisional fairness. Accordingly, CSRBs should be empowered to grant conditional approval to innovative data projects. This would allow for organizations to proceed with potentially beneficial projects, even as decisions are reviewed and modified as new data trends emerge.

#### 4. Respect for law

The fourth principle added by the Menlo Report, *Respect for Law and Public Interest*, builds on these aforementioned substantive requirements and stresses the need for organizations to engage in legal due diligence, be transparent in methods and results, and accountable for their actions.<sup>104</sup> As a basic matter, of course, organizations should comply with the law. A lack of transparency and accountability risks undermining the credibility of, trust and confidence in, and ultimately support for organizations' endeavors in general and novel research settings.<sup>105</sup>

When evaluating big data proposals, a CSRB should assess

---

101. See Peter Swire, *Lessons from Fair Lending Law for Fair Marketing and Big Data* (Sept. 2014), [http://www.futureofprivacy.org/wp-content/uploads/FairMarketingLessons\\_WhitePaperFTC.pdf](http://www.futureofprivacy.org/wp-content/uploads/FairMarketingLessons_WhitePaperFTC.pdf).

<sup>102</sup> *Supra* note 61 and accompanying text.

103. *Id.* at 11.

104. MENLO REPORT, *supra* note 9, at 15.

105. *See id.*

measures deployed by an organization to minimize privacy impact. In doing so, a CSRB could ask whether: *ex post* notification to individuals could replace the need for consent; retention limitations could be put in place to restrict the likelihood of future data repurposing; and the information collected is especially sensitive (e.g., health or mental conditions, intimate relations, political views). Additionally, a CSRB should explore the nature of collaboration with the research community and whether the results of the research will be published.

### B. Promoting Independence and Trust

One of the key mechanisms for promoting trust that emerged out of the Belmont Report was the establishment of independent IRBs. These institutions were designed to ameliorate conflicts of interest that could appear in research, serve as a check over researchers and avoid allowing the ends to always justify the means.<sup>106</sup> An IRB acts as an administrative body with a degree of independence and distance “from the investigator and the rewards of research” that has the authority to approve, require modifications in, or disapprove all research activities under its purview.<sup>107</sup>

While the IRB system has been subject to criticism over the ensuing decades,<sup>108</sup> there is widespread agreement that IRBs have served an important role not only in ensuring ethics in research but maintaining public confidence in the research system at large.<sup>109</sup> As Robert Levine, a medical ethicist at Yale, has explained, the single most important factor that contributes to the successful functioning of the IRB is its credibility within its institution and within the community that institution serves.<sup>110</sup> One benefit of the IRB system

---

106. Margaret R. Moon & Felix Khn-Maung-Gyi, *The History and Role of Institutional Review Boards*, 11 *AMA J. ETHICS* 311 (2009), available at <http://journalofethics.ama-assn.org/2009/04/pfor1-0904.html>.

107. *Id.* Also see IRB Guidebook, Chapter 1: [http://www.hhs.gov/ohrp/archive/irb/irb\\_chapter1.htm](http://www.hhs.gov/ohrp/archive/irb/irb_chapter1.htm).

108. DEP'T OF HEALTH & HUMAN SERVS, INSTITUTIONAL REVIEW BOARDS: A TIME FOR REFORM (1998), <https://oig.hhs.gov/oei/reports/oei-01-97-00193.pdf>; DEP'T OF HEALTH & HUMAN SERVS., PROTECTING HUMAN RESEARCH SUBJECTS: STATUS OF RECOMMENDATIONS (2000), <https://oig.hhs.gov/oei/reports/oei-01-97-00197.pdf>; Ann Hamilton, *The Development and Operation of IRBs: Medical Regulations and Social Science*, 33 *J. APPLIED COMM. RESEARCH* 189 (2005).

109. Beth Azar, *Ethics at the Cost of Research?*, *MONITOR ON PSYCHOL.*, Feb. 2002, at 38, available at <http://www.apa.org/monitor/feb02/ethicscost.aspx>; see also Richard S. Saver, *Medical Research Oversight from the Corporate Governance Perspective: Comparing Institutional Review Boards and Corporate Boards*, 46 *WM. & MARY L. REV.* 619, 693 (2004).

110. ROBERT J. LEVINE, *ETHICS AND REGULATION OF CLINICAL RESEARCH* (1986).

is to serve the dual purpose of protecting the rights and welfare of research subjects while ensuring fair application of policies to researchers.<sup>111</sup> CSRBs should work to fulfill these aspirations, while taking structural lessons from the IRB system to adapt it to a non-academic setting.

Internal CSRBs should reflect multiple viewpoints within an organization and have the necessary expertise to understand the ethical and technical challenges of an organization's innovative data use. The review process must also be funded adequately to ensure that a CSRB has the required resources, including, if necessary, counseling by external experts.

Although comprising mostly insiders, traditional IRBs are shielded from commercial and political pressures by the innate independence of tenured academic faculty, academics' reputational concerns, and the nonprofit nature of major research institutions.<sup>112</sup> In addition, under the Common Rule, "Each IRB shall include at least one member who is not otherwise affiliated with the institution . . ."<sup>113</sup> The Common Rule requires IRB membership to be diverse through the inclusion of individuals with varying backgrounds and "consideration of race, gender, and cultural backgrounds and sensitivity to such issues as community attitudes . . ."<sup>114</sup> At least one member of an academic IRB must have a scientific background and at least one must come from a nonscientific area.<sup>115</sup>

Similarly, the makeup of an internal CSRB must reflect multiple viewpoints within an organization with the necessary expertise to weigh complex operational issues. This would include stakeholders such as an organization's privacy officer, legal counsel, and ethics expert, along with representatives of product, engineering, marketing, and PR departments.<sup>116</sup> The privacy officer should chair and anchor the discussions to reflect the primacy of privacy concerns in the analysis. One way an organization can demonstrate

---

111. Protection of Human Subjects, 43 Fed. Reg. 56174, 56175–76 (Nov. 30, 1978). Further, IRBs can help to "contribute to the education of the research community and the public regarding the ethical conduct of research. The committees can become resource centers for information concerning ethical standards and federal requirements and can communicate with federal officials and with other local committees about matters of common concern." *Id.*

112. *Cf. Saver, supra* note 109.

113. 45 C.F.R. § 46.107 (2009).

114. *Id.* § 46.107(a).

115. *Id.* § 46.107(c).

116. Facebook's research review panel, for example, includes senior subject area researchers and representatives from engineering, research, legal, privacy, and policy teams. *See Schroeffer, supra* note 40.

its commitment to independent review is by ensuring the seniority and job security of the privacy officer. Over the past few years, the appointment of privacy officers has heralded the emergence of a new profession comprising individuals who have risen in ranks to become in-house guardians of personal data.<sup>117</sup> Organizations could buttress the job security of privacy officers by committing to terminate them only for cause. Securities regulators could require publicly traded companies to disclose in their financial statements the reasons for any such termination.

A key question is what role external participants should play on an internal CSRB. As a practical matter, external participants may in some cases be useful to ensure that an internal CSRB has the necessary expertise to evaluate an innovative data proposal. Increasingly, companies are already establishing outside “privacy advisory boards” that call upon experts in law, ethics, statistics, and PR to advise them on complex questions of privacy and ethics.<sup>118</sup> While including outside voices adds to the objectivity and public legitimacy of any review process, most existing boards act in a purely advisory role, not as decision makers. The law should not dictate to organizations the makeup of an internal CSRB, yet if the FTC chooses to challenge the decision of an internal CSRB, the proportion of independent voices on the CSRB could weigh on the credibility of its decision to authorize a particular data use.

Internal CSRBs must be structured to provide an organization with candid, substantive advice, comparable to a board of directors’ audit committee. A CSRB’s reporting structure must be designed to promote objectivity and accountability. Accordingly, at least one member of an internal CSRB should be a member of or accountable to the organization’s senior management. Internal CSRBs must also have their own budget to ensure adequate resources to investigate data uses, obtain expert advice and counsel, and follow projects over time.<sup>119</sup>

---

117. Andrew Clearwater & J. Trevor Hughes, *In the Beginning... An Early History of the Privacy Profession*, 74 OHIO ST. L. J. 897 (2013); A. Bamberger & Deirdre K. Mulligan, *Privacy on the Books and on the Ground*, 63 STAN. L. REV. 247, 253 (2011).

118. See, e.g., Schroepfer, *supra* note 40; John Grant, *Announcing the Palantir Council on Privacy and Civil Liberties*, THE PALANTIR BLOG (Nov. 2, 2012), <https://www.palantir.com/2012/11/announcing-the-palantir-council-on-privacy-and-civil-liberties/>; *The Advisory Council to Google on the Right to be Forgotten*, GOOGLE (2014), <https://www.google.com/advisorycouncil/>. For ethical review boards of government projects, see <http://www.surveille.eu/>.

119. See CENTER FOR INFO. POLICY LEADERSHIP, THE ROLE AND FUNCTION OF A DATA PROTECTION OFFICER IN THE EUROPEAN COMMISSION’S PROPOSED GENERAL DATA PROTECTION REGULATION (Sept. 25, 2013) (listing the requirements for independence of the data



One challenge that must still be addressed is whether CSRBs should function as monitoring mechanisms or rather as mediators across different stakeholder groups both inside and outside of industry. One of the key benefits of an internal CSR is its capacity to play a monitoring role, particularly with regard to identifying and scrutinizing potentially unfair discriminatory impacts over time. However, comparing between the existing IRB system and corporate governance more generally, Richard Safer has cautioned that formalized and intensive monitoring can impair an IRB's ability to mediate different stakeholder interests.<sup>120</sup> As Robert Levine explained, "IRBs were established to work collaboratively with investigators," rather than as overly aggressive overseers.<sup>121</sup> Finding the optimal balance to enable an internal CSR to function with a degree of independence while at the same time promoting trust among internal stakeholders, regulators, and the public will require additional discussion.

To sum, an internal CSR will need to reflect multiple viewpoints within an organization and have the necessary expertise to understand the ethical and technical challenges of innovative data use. This will require establishing a review process that is adequately funded and resourced. An organization must provide its internal CSR with the requisite resources to perform its task including, if necessary, counseling by external experts and various outside stakeholders, and to set forth transparent rules and procedures for developing trust and authority.

### C. Documentation and Accountability

A key challenge for internal CSRs is striking a balance between the secrecy required to facilitate information sharing and open discussion and the transparency needed to enhance trust and promote accountability. In order for internal CSRs to be accepted as a legitimate alternative to external oversight, they must employ detailed documentation requirements. Proper documentation will help protect all stakeholders, including consumers from undue risk of harm and organizations from potential liability.<sup>122</sup> Procedure-

---

protection officer (DPO) under the reformed EU privacy regulation).

120. Saver, *supra* note 109, at 693-94.

121. Carl Elliott & Trudo Lemmens, *Ethics for Sale*, SLATE (Dec. 13, 2005), [http://www.slate.com/articles/health\\_and\\_science/medical\\_examiner/2005/12/ethics\\_for\\_sale.html](http://www.slate.com/articles/health_and_science/medical_examiner/2005/12/ethics_for_sale.html).

122. See MARILYN BERNER, WRITE SMARTER, NOT LONGER, IN THE MENTAL HEALTH PRACTITIONER AND THE LAW: A COMPREHENSIVE HANDBOOK 54-71 (1998).

oriented documentation should be thorough, with organizations documenting board constitution, membership, reporting structure, and ongoing monitoring and review procedures.<sup>123</sup>

In addition to extensive process-oriented documentation, organizations should be required to document the rationale for their ultimate decision-making. The precise contours of these requirements merit further discussion. Organizations, after all, should not be placed in the position of having to disclose information that is legally protected or could compromise valuable trade secrets. But transparency and documentation will be necessary in order to facilitate potential after-the-fact regulatory investigations. It will also be important for organizations to communicate their use of an internal CSRB to the public. At a minimum, organizations should strive to make the general procedures used by an internal review process available for public scrutiny.

In this respect, the Common Rule guidance of the Office for Human Research Protections is instructive.<sup>124</sup> It requires organizations to set forth detailed policies and procedures for IRBs, including a step-by-step description with key operational details about: a primary reviewer system, lists of specific documents distributed to reviewers, the timing of document distribution prior to IRB meetings, the range of possible actions taken by the IRB, etc. To enhance transparency and accountability, CSRBs could be required to undertake similar obligations.

#### CONCLUSION

In an era of big data, data research has escaped the confines of academic institutions. It pervades daily business decisions not only in the technology sector but also in more traditional industries like finance, healthcare, retail, and utilities. It underlies decision-making processes in government and private sector, including nonprofits, and political campaigns. It calls for a new set of ethical guidelines, based on the foundational Belmont principles for human subject research, to help organizations make careful decisions balancing big data opportunities against risks to individuals' privacy and civil

---

123. Some of the records that traditional IRBs are required to maintain, include (1) research proposals, (2) meeting minutes, (3) review activities, (4) certain categories of correspondence, (5) panel membership information, and (6) written procedures that they use. 21 C.F.R. § 56.115(a) (2015).

124. DEP'T OF HEALTH & HUMAN SERVS, OFFICE FOR HUMAN RESEARCH PROTECTIONS, GUIDANCE ON WRITTEN IRB PROCEDURES, July 1, 2011, *available at* <http://www.hhs.gov/ohrp/policy/irbgd107.pdf>.

liberties.

Ever since its genesis in the seminal article by Samuel Warren and Louis Brandeis, the right to privacy has had to be balanced against competing interests and norms.<sup>125</sup> Warren and Brandeis wrote, "It remains to consider what are the limitations of this right to privacy . . . . To determine in advance of experience the exact line at which the dignity and convenience of the individual must yield to the demands of the public welfare or of private justice would be a difficult task . . . ." <sup>126</sup>

A CSRB will be tasked with operationalizing these lofty principles by conducting a delicate balancing job, weighing risks to dignitary values against the prospect for big data rewards.

Establishment of a CSRB is not a silver bullet. It needs to be accompanied by clear procedural guidelines ensuring independent expert review, transparency, and accountability, to prevent captured boards that rubber-stamp management decisions. It requires substantive principles to guide decision-makers as they debate what are, ultimately, not legal but rather ethical, moral questions. This article launches a discussion about the skill set and toolbox required to undertake this mission in order to guarantee the responsible generation of knowledge and data innovation.

---

125. Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 214 (1890).

126. *Id.*