JUDGED BY THE TIN MAN: INDIVIDUAL RIGHTS IN THE AGE OF BIG DATA

OMER TENE AND JULES POLONETSKY*

"*'How about my heart?' asked the Tin Woodman.*
'Why, as for that,' answered Oz, 'I think you are wrong to want a heart. It makes most people unhappy. If you only knew it, you are in luck not to have a heart.'"[1]

INTRODUCTION

Big data—the enhanced ability to collect, store and analyze previously unimaginable quantities of data in tremendous speed and with negligible costs[2]—delivers immense benefits in marketing efficiency,

* Omer Tene is Deputy Dean of the College of Management Haim Striks School of Law, Israel and a Senior Fellow at the Future of Privacy Forum; Jules Polonetsky is Co-chair and Executive Director of the Future of Privacy Forum.

1. L. FRANK BAUM, THE WONDERFUL WIZARD OF OZ 190 (1900).

2. Neil Versel, *Big Data Helps Kaiser Close Healthcare Gaps*, INFORMATION WEEK (Mar. 7, 2013), http://www.informationweek.com/healthcare/electronic-medical-records/big-data-helps-kaiser-close-healthcare-g/240150269 (Kaiser Permanente defines big data as data for which the "size is beyond the ability of typical database software tools to capture, store, manage, and analyze."). *See also*, McKinsey Global Institute, *Big Data: The Next Frontier for Innovation, Competition and Productivity* (2011), *available at* http://www.mckinsey.com/~/media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx. Big data, however, is typically characterized not only by volume but also by velocity (speed of processing) and variety (the capability to link diverse data sets and process unstructured data)

healthcare, environmental protection, national security and more.[3] While some privacy advocates may dispute the merits of sophisticated behavioral marketing practices or debate the usefulness of certain data sets to efforts to identify potential terrorists,[4] few remain indifferent to the transformative value of big data analysis for government, science, and society at large.[5] At the same time, even big data evangelists should recognize the potentially ominous social ramifications of a surveillance society governed by heartless algorithmic machines.[6]

In this essay, we present some of the privacy and non-privacy risks of big data as well as directions for potential solutions. In a previous paper, we argued that the central tenets of the current privacy framework, the principles of data minimization and purpose limitation, are severely strained by the big data technological and business reality.[7] Here, we assess some of the other problems raised by pervasive big data analysis. To highlight the ethical and moral dilemmas, we sometimes refer to big data algorithms as "the machine" (which is more elegant than "zombie," though less animated than the "tin man" in the title).[8] In their book, *A Legal Theory for Autonomous Artificial Agents*, Samir Chopra and Larry White note that "as we increasingly interact with these artificial agents in unsupervised settings, with no human mediators, their seeming autonomy and increasingly sophisticated functionality and behavior, raises legal and philosophical questions."[9] In this article, we argue that the focus on the machine is a distraction from the debate surrounding data driven ethical dilemmas, such as privacy, fairness, and discrimination. The

(together referred to as the "three v's"). *See also* Press Release, Gartner, Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data (June 27, 2011), *available at* http://www.gartner.com/it/page.jsp?id=1731916.

3. Omer Tene & Jules Polonetsky, *Privacy in the Age of Big Data: A Time for Big Decisions*, 64 STAN. L. REV. 63 (2012), http://www.stanfordlawreview.org/online/privacy-paradox/big-data.

4. Julia Anguin, *U.S. Terrorism Agency to Tap a Vast Database of Citizens*, WALL ST. J. (Dec. 13, 2012), *available at* http://online.wsj.com/article/SB10001424127887324478304578171623040640006.html; R. Jeffrey Smith, *Senate Report Says National Intelligence Fusion Centers Have Been Useless*, THE CENTER FOR PUBLIC INTEGRITY (Oct. 3, 2012), *available at* http://www.foreignpolicy.com/articles/2012/10/03/senate_report_says_national_intelligence_fusion_centers_have_been_useless.

5. *See* WORLD ECON. FORUM, PERSONAL DATA: THE EMERGENCE OF A NEW ASSET CLASS (2011), *available at* http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf.

6. Paul Ohm, *Don't Build a Database of Ruin*, HARV. BUS. REV. BLOG (Aug. 12, 2012, 10:00 AM), http://blogs.hbr.org/cs/2012/08/dont_build_a_database_of_ruin.html.

7. 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).

8. *Cf.* Chopra & White's use of "artificial agents." SAMIR CHOPRA & LAURENCE WHITE, A LEGAL THEORY FOR AUTONOMOUS ARTIFICIAL AGENTS (2011).

9. *Id.* at 2.

machine may exacerbate, enable, or simply draw attention to the ethical challenges, but it is humans who must be held accountable.

Instead of vilifying machine-based data analysis and imposing heavy-handed regulation, which in the process will undoubtedly curtail highly beneficial activities,[10] policymakers should seek to devise agreed-upon guidelines for ethical data analysis and profiling. Such guidelines would address the use of legal and technical mechanisms to obfuscate data; criteria for calling out unethical, if not illegal, behavior; categories of privacy and non-privacy harms; and strategies for empowering individuals through access to data in intelligible form.

HUMAN SUBJECT RESEARCH UNFETTERED

Big data has expanded the scope of human subject research far beyond anything envisaged by social science or medical researchers just a few years ago. Today, everyone—including businesses, governments, private citizens and platform operators—is a "researcher," analyzing the data exhaust produced by individuals' daily lives to identify useful patterns and correlations. In most cases, these research activities are not tempered by the procedural and ethical safeguards, which were traditionally required to conduct human subject research. To the contrary, the machine is often driven by entrepreneurs, app developers, or data scientists who seek innovation at any cost. Although in many large companies chief privacy officers and legal teams play an oversight role, today's start-up app developers can rapidly amass vast amounts of data with little, if any, oversight. This type of research impacts not only the privacy of individuals whose data is examined, but also the rights of those subject to social sorting as a consequence.[11]

Like any interpretative process, big data analysis is prone to error and far from objective. Data crunching may appear to be an exact science; yet it is laden with subjective input from researchers who decide which data to analyze, questions to examine, and purposes to pursue. As danah boyd put it: "[d]o numbers speak for themselves? The answer, we think, is a resounding 'no'. . . . All researchers are interpreters of data."[12] The same numbers tell different stories depending on the methodologies and theories of those who set the research agenda. Furthermore, the machine is not immune to error based on inaccurate input, skewed

10. *See* discussion *infra*, notes 60 to 66 and accompanying text.

11. For example, if research demonstrates that men between age 40 and 50 who smoked for 10 years have a high instance of heart disease, the insurance premiums charged to an individual who meets these criteria will rise regardless of whether or not his data was in the original dataset.

12. danah boyd & Kate Crawford, *Six Provocations for Big Data*, A DECADE IN INTERNET TIME: SYMPOSIUM ON THE DYNAMICS OF THE INTERNET AND SOC'Y (2011), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431.

samples, or faulty algorithms. Some commentators go so far as arguing that most current published research findings are false or at best inaccurate measures of the prevailing bias.[13] When viewing ads tailored by intricate behavioral tracking infrastructures, we often sense that the machine gets us all wrong. Eli Pariser called this "a bad theory of you," based on there being no single set of criteria that describes who we are.[14] While relatively benign when the decision is whether to show a web surfer a sports ad or a fashion ad, erroneous results may have profound adverse effects on individuals' lives in other contexts, such as healthcare, credit, employment or law enforcement.[15]

Even more fundamentally, big data analysis is inapposite to traditional methods of scientific research (i.e., define a research question; gather information; form an explanatory hypothesis; test the hypothesis; etc.). While these earlier paradigms were characterized by experimentation and reasoning, big data analysis is driven by the availability of data at an unprecedented scale as well as the computational resources enabling rapid value extraction. As Julie Cohen observes, "the idea of the scientific research program as a series of limited data collections for the purpose of testing and possibly falsifying a particular hypothesis."[16] Some regard this challenge to traditional scientific method a groundbreaking revolution, heralding a "fourth paradigm" of scientific research.[17] Others question the rigor of scientific investigations that are both open-ended and ongoing.[18] One commentator notes that "[r]elaxed practices regarding the communication of computational details is creating a credibility crisis in computational science, not only among scientists, but as a basis for policy decisions and in the public mind."[19]

13. *See* John P. Ioannidis, *Why Most Published Research Findings Are False*, 2 PLOS MED. 696 (2005), *available at* http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.0020124; *see also* David H. Freedman, *Lies, Damned Lies, and Medical Science*, THE ATLANTIC (Oct. 4, 2010, 6:16 PM), http://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/308269/.

14. ELI PARISER, THE FILTER BUBBLE: WHAT THE INTERNET IS HIDING FROM YOU (2011).

15. *See, e.g.*, Nassim Taleb, *Beware the Big Errors of 'Big Data'*, WIRED (Feb. 8, 2013), *available at* http://www.wired.com/opinion/2013/02/big-data-means-big-errors-people.

16. Julie E. Cohen, *What Privacy is For*, 126 HARV. L. REV. (forthcoming 2013), *available at* http://www.harvardlawreview.org/symposium/papers2012/cohen.pdf.

17. *See, e.g.*, Gordon Bell, Tony Hey & Alex Szalay, *Beyond the Data Deluge*, 323 SCIENCE 1297 (2009), *available at* https://www.sciencemag.org/content/323/5919/1297.full.

18. *See* Mark Birkin, *Big Data Challenges for Geoinformatics*, *in* GEOINFOR GEOSTAT: AN OVERVIEW 1 (2012), *available at* http://www.scitechnol.com/2327-4581/2327-4581-1-e101.pdf.

19. VICTORIA STODDEN, ESTABLISHING SCIENTIFIC FACTS, COLUM. UNIV. DEPT. OF STATISTICS (Sept. 2011), *available at* http://www.stanford.edu/~vcs/talks/VictoriaStoddenFQXiSept2011.pdf; *see also* David Berry,

Finally, attention must be given to the accessibility of big data sets to the research community at large.[20] Traditionally, when scientists published their research, they also made the underlying data available so that other scientists could verify the results. Yet with big data, it is often only the employees of certain organizations that benefit from access, conducting analysis and publishing results without making the underlying data publicly available.[21] Such scientists may argue, first, that the data are a proprietary asset of their business. Indeed, they may claim that disclosing the data could infringe customers' privacy.[22] Who gets access to big data sets; for what purposes; in what contexts; and with what constraints—are fundamental questions that must be addressed by future research.[23]

DISCRIMINATION – TELLING RIGHT FROM WRONG

Significantly, big data analysis allows for granular distinctions to be made between individual characteristics, preferences and activities. Whether such distinctions are made for the sake of personalization, research or public planning, they facilitate discrimination based on a wide (in fact, infinite) spectrum of characteristics. We refer here to "discrimination" in a value-neutral sense; i.e., drawing distinctions between individuals and treating them differently based on such distinctions.[24] To assess the ethical implications of discrimination, we need to unpack the meaning of the term, which is, of course, highly

The Computational Turn: Thinking About the Digital Humanities, 12 CULTURE MACH. 1 (2011); David L. Donoho et al., *Reproducible Research in Computational Harmonic Analysis*, COMPUTING IN SCI. & ENG'G, Jan./Feb. 2009, at 8.

20. *See* John Markoff, *Troves of Personal Data, Forbidden to Researchers*, N.Y. TIMES, May 2, 2012, at D1.

21. *See* Lev Manovich, *Trending: The Promises and the Challenges of Big Social Data*, *in* DEBATES IN THE DIGITAL HUMANITIES (Matthew Gold ed., 2012) (claiming that "only social media companies have access to really large social data – especially transactional data. An anthropologist working for Facebook or a sociologist working for Google will have access to data that the rest of the scholarly community will not.").

22. *See* Bernardo Huberman, *Sociology of Science: Big Data Deserve a Bigger Audience*, 482 NATURE 308 (2012) (warning that privately held data was threatening the very basis of scientific research, and complaining that "[m]any of the emerging 'big data' come from private sources that are inaccessible to other researchers. The data source may be hidden, compounding problems of verification, as well as concerns about the generality of the results.").

23. boyd & Crawford, *supra* note 12, at 12.

24. Merriam-Webster.com defines the intransitive verb "discriminate" as "1) to make a distinction; 2) to make a difference in treatment or favor on a basis other than individual merit." Black's Law Dictionary defines "discrimination" as: "1) The effect of a law or established practice that confers privileges on a certain class or that denies privileges to a certain class because of race, age, sex, nationality, religion, or disability; 2) Differential treatment; esp., a failure to treat all persons equally when no reasonable distinction can be found between those favored and those not favored." BLACK'S LAW DICTIONARY 1886 (9th ed. 2009).

charged. Discrimination could be socially desired (e.g., treating minors as children and not as adults); generally acceptable (e.g., applying Amazon's recommendation system to enhance consumers' shopping experience); or morally reprehensible (e.g., not hiring individuals of a certain age or race). In our daily life, we draw distinctions (i.e., discriminate) all the time. A person sitting next to us on a plane is tall or short, agitated or relaxed, attractive or unattractive, young or old—there is an endless list of such adjectives; and our attitudes and actions towards that person will vary accordingly.

The machine can instantly make millions of such distinctions working with vast pools of personal data. But an ethical assessment of machine-driven distinctions requires a coherent theory of discrimination. The machine is incapable of determining whether a distinction is ethical or not. Unless we come up with a comprehensive theory of discrimination that can be represented algorithmically, we have no rigorous way to distinguish between ethical and non-ethical machine-based discrimination.[25] We certainly should not expect the machine to make moral decisions that *we* have yet to make.

Have we decided why it is legitimate to market to pregnant women in one context (e.g., based on subscription to a magazine) but morally distasteful to do so in another (e.g., Target's compilation of a "pregnancy score" for shoppers)?[26] Can an employer ethically decline to interview a job candidate because they see a picture of them drinking a beer on a social media site?[27] Is price discrimination, the offering of different prices to different people based on their perceived willingness to pay, good or bad? Does it favor the wealthy[28] or the less privileged?[29] Is it fair

25. There have been attempts of statistical testing for discrimination in big data analysis. *See, e.g.*, Salvatore Ruggieri, Dino Pedreschi & Franco Turini, *Data Mining for Discrimination Discovery*, 4 ACM TRANSACTIONS ON KNOWLEDGE DISCOVERY FROM DATA, Art. 9 (May 2010); Binh Thanh Luong, Salvatore Ruggieri & Franco Turini, *k-NN as an Implementation of Situation Testing for Discrimination Discovery and Prevention*, *in* PROCEEDINGS OF THE 17TH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING 502 (Aug. 2011). These efforts too must first coalesce around an agreed upon delineation of legitimate vs. illegitimate discrimination.

26. *See* Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES MAG. (Feb. 16, 2012), http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html.

27. Jeffrey Rosen, *The Web Means the End of Forgetting*, N.Y. TIMES (July 21, 2010), http://www.nytimes.com/2010/07/25/magazine/25privacy-t2.html.

28. *See* Jennifer Valentino-DeVries, Jeremy Singer-Vine & Ashkan Soltani, *Websites Vary Prices, Deals Based on Users' Information*, WALL ST. J. (Dec. 24, 2012), http://online.wsj.com/article_email/SB10001424127887323777204578189391813881534-lMyQjAxMTAyMDIwNDEyNDQyWj.html#12 (reporting that "areas that tended to see the discounted prices had a higher average income than areas that tended to see higher prices"); *see also* Omer Tene, *Privacy: For the Rich or for the Poor?*, CONCURRINGOPINIONS.COM (July 26, 2012), http://www.concurringopinions.com/archives/2012/07/privacy-for-the-rich-or-for-the-poor.html.

29. *See* Dana Mattioli, *On Orbitz, Mac Users Steered to Pricier Hotels*, WALL ST. J.

that companies can exploit price sensitivity on an individualized basis, thereby usurping the entire value surplus available in a transaction by pricing goods or services as close as possible to an individual's reservation price?[30] What is the fault line between legitimate (or at least not illegal) price discrimination and price discrimination that effectively excludes entire groups of individuals (who are viewed as not being "worth enough" to bother with) from the market? And what if the makeup of such excluded groups is positively correlated with racial or gender bias?

It is difficult enough to decide which forms of discrimination are illegal. Deciding whether discrimination that is not illegal is unethical or morally undesired may become daunting. Robert Fullinwider explains:

> Many may be led to the false sense that they have actually made a moral argument by showing that the practice discriminates (distinguishes in favor of or against). The temptation is to move from 'X distinguishes in favor of or against' to 'X discriminates' to 'X is wrong' without being aware of the equivocation involved.[31]

Should we preempt any form of discrimination by requiring companies to mail Porsche catalogs to everyone regardless of income? Should Victoria's Secret or Pampers be required to target all shoppers regardless of gender or age? Or perhaps offers should always be *available* to all but not *promoted* to all? But then again, that may deny the benefit of the bargain to those who do not know about it.

Some of our ethical and moral criteria are so fragile, nuanced, and culturally dependent that it is not clear that the machine will *ever* be capable of appropriately weighing them. Indeed, it is far from clear that we would even *want* the machine to obtain the ability to distinguish right from wrong. Such an anthropomorphized machine—a "technological singularity"[32]—would likely cause more privacy and moral angst than the current dumbed-down version.[33] Artificial intelligence has yet to

(Aug. 23, 2012), http://online.wsj.com/article/SB10001424052702304458604577488822667325882.html.

30. In a prior article we likened transacting with a big data platform to a game of poker where one of the players has his hand open and the other keeps his cards close. The online company knows the preferences of the transacting individual inside out, perhaps better than the individual knows him or herself. Tene & Polonetsky, *supra* note 7.

31. ROBERT FULLINWIDER, THE REVERSE DISCRIMINATION CONTROVERSY: A MORAL AND LEGAL ANALYSIS 11–12 (1980).

32. RAY KURZWEIL, THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY (2006).

33. Some would say such a machine is "creepy." *See generally* Omer Tene & Jules Polonetsky, *A Theory of Creepy: Technology, Privacy and Shifting Social Norms*, ___ YALE J. L. & TECH. (forthcoming 2014).

produce systems that approach human-cognition.[34] Far from it, the only morality that can currently be attributed to the machine is what Ian Kerr calls "slave morality," the proclivity to fulfill human orders to inhuman perfection. Bruce Boyden recently argued that it is precisely the inhuman nature of the machine that allays privacy concerns in the context of machine-based communications surveillance.[35] "What people who worry about privacy are trying to prevent is changed beliefs about themselves, changed behavior by other people, or changed attributions of social status resulting from a disclosure of private information—in other words, changed mental states."[36] For the sake of privacy, it may be best to leave the tin man without a heart.

DON'T BLAME THE MACHINE

A complicating factor is that the machine's unrestricted ability to identify patterns in endless piles of data facilitates the masking of illegitimate or illegal discrimination behind layers upon layers of mirrors and proxies.[37] A clever programmer can embed bias in a complex algorithm such that discrimination will be very difficult to detect.[38] The machine can find strong correlations, which result in discriminatory outcomes that are based on neutral factors. It is wrong to discriminate based on race; yet it will be exceedingly difficult to detect such discrimination if it is based on a dozen factors that through big data analysis are found to be positively correlated to race. And sometimes it will be difficult to discern whether the category used for profiling is

34. Note Harry Surden's observation that "these statistical and probability-based machine-learning models (often combined with logical-knowledge based rules about the world) often produce high-quality and effective results (not quite up to the par of nuanced human translators at this point), without any assertion that the computers are engaging in profound understanding with the underlying "meaning" of the translated sentences or employing processes whose analytical abilities approach human-level cognition." Harry Surden, *Autonomous Agents and Extension of Law: Policymakers Should be Aware of Technical Nuances*, CONCURRINGOPINIONS.COM (Feb. 16, 2012), http://www.concurringopinions.com/archives/2012/02/autonomous-agents-and-extension-of-law-policymakers-should-be-aware-of-technical-nuances.html.

35. Bruce Boyden, *Can a Computer Intercept Your Email?*, 34 CARDOZO L. REV. 669 (2012).

36. *Id.* at n.188.

37. *Id.*; *see, e.g.*, Carter Jernigan & Behram F.T. Mistree, *Gaydar: Facebook friendships expose sexual orientation*, 14(10) FIRST MONDAY (Oct. 2009), http://www.firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2611/2302 (demonstrating a method for accurately predicting the sexual orientation of Facebook users by analyzing friendship associations).

38. The Supreme Court has ruled that under the Civil Rights Act of 1964, a policy that was fair in form but discriminatory in impact is illegal. *See* Griggs v. Duke Power Co., 401 U.S. 424, 431 (1971) ("Congress has now provided that tests or criteria for employment or promotion may not provide equality of opportunity merely in the sense of the fabled offer of milk to the stork and the fox.").

legitimate or just a façade for another, less wholesome agenda.[39] This just goes to say that the machine can be a powerful tool for discrimination, just as it is a potent tool for healthcare research, environmental sustainability and economic efficiency. It does not discriminate any more or less legitimately than the people who use it.

There is nothing new about the fact that people discriminate based on unethical criteria, some of which are not illegal. For example, employers may (or may not) prefer to hire attractive job candidates.[40] That today they can satisfy such bias by sifting through candidates' Facebook profiles is a phenomenon that has nothing to do with the morality of the technology itself. To be sure, the machine enables the scaling of such discrimination to entire populations. But should we outlaw distinctions drawn by the machine in cases where those same distinctions are legal (albeit subject to moral disdain) if drawn by individuals? Some laws aspire to resolve machine-based discrimination by requiring the involvement of a human operator at certain decision-making junctures.[41] However, it is far from clear that human intervention mitigates discrimination risk; in fact the opposite may be true. Indeed, when technical risk-based profiles were first introduced in the mortgage industry, they were hailed as a definitive answer to the unequal treatment loan officers give borrowers.[42]

Consider, for example, recent research by Latanya Sweeney demonstrating that a greater percentage of ads having the word "arrest" in their text appear for searches on Google and Reuters.com for black-identifying first names (such as DeShawn, Darnell and Lakisha) than for white-identifying first names (such as Brad, Dustin and Jill).[43] Surely it is not the machine that independently decided to discriminate on a first

39. Consider the Federal Reserve Board report asserting that credit card companies adjusted consumers' rates and credit limits based in part on where they shopped, what they bought, and whom they bought from. What could such criteria be correlated to or disguise? *See* BD. OF GOVERNORS OF THE FED. RESERVE SYS., REPORT TO THE CONGRESS ON REDUCTIONS OF CONSUMER CREDIT LIMITS BASED ON CERTAIN INFORMATION AS TO EXPERIENCE OR TRANSACTIONS OF THE CONSUMER 19 (2010), *available at* http://www.federalreserve.gov/BoardDocs/RptCongress/creditcard/2009/consumercreditreductions.pdf.

40. *See Hiring Hotties*, THE ECONOMIST (July 21, 2012), http://www.economist.com/node/21559357 (attractiveness discrimination); *cf. Don't hate me because I'm beautiful*, THE ECONOMIST (Mar. 31, 2012), http://www.economist.com/node/21551535.

41. *See* Directive 95/46/EC of the European Parliament and of the Council 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, art. 15, 1995 O.J. (L 281) 31 [hereinafter European Data Protection Directive].

42. *See* U.S. DEPARTMENT OF JUSTICE, CIVIL RIGHTS DIVISION, FAIR LENDING ENFORCEMENT PROGRAM (2001).

43. Latanya Sweeney, Discrimination in Online Ad Delivery 11 (Jan. 28, 2013) (unpublished manuscript), *available at* http://dataprivacylab.org/projects/onlineads/1071-1.pdf.

name basis; rather as Sweeney observes, online ad delivery is a "socio-technical construct."[44] In her research findings, Sweeney could not determine whether the documented bias was caused by advertisers providing ad templates suggestive of arrest disproportionately to black-identifying names, or by the Google Ad Sense algorithm simply reflecting *society's* bias by preferring to place ads that obtain higher clickthrough rates.[45] Sweeney posits, "technology can do more to thwart discriminatory effects and harmonize with societal norms."[46] Hence, she calls for "fairness by design" to complement the increasingly prevalent requirement for "privacy by design."[47] Indeed, Cynthia Dwork and others suggest innovative ways to bake fairness into algorithms to prevent overt or covert discrimination.[48] At the same time, if we believe certain distinctions are worthy of legal restriction, law should bar their use in decisions regardless of whether they are made by human or machine. In other words, as long as humans continue to be biased and discriminating, machine-made decisions will reflect (and may very well amplify) such discrimination.

FRAGMENTATION OF PUBLIC DISCOURSE

An additional and somewhat related problem, which was exposed by Joe Turow,[49] Cass Sunstein[50] and others, concerns the risks to free speech and democratic discourse that are inherent in the fragmentation of the information commons. Personalization technologies channel content into "filter bubbles," enabling platform providers and inevitably governments to "divide and conquer" by manipulating public opinion.[51] For example, during the last U.S. Presidential elections, political campaigns were "micro-targeted" delivering individualized messages to potential voters based on their narrow interests, causes, and fears.[52] The data that support this micro-targeting are increasingly being merged with information about the online identities and behavior of voters. These practices raise concerns about loss of voter anonymity, political speech,

44. *Id.* at 3.

45. *Id.* at 34.

46. *Id.* at 35.

47. *Id.*; Ira Rubinstein, *Regulating Privacy by Design*, 26 BERKELEY TECH. L.J. 1409 (2011).

48. CYNTHIA DWORK ET AL., FAIRNESS THROUGH AWARENESS (2011), *available at* http://www.cs.toronto.edu/~zemel/documents/fairAwareItcs2012.pdf.

49. JOSEPH TUROW, THE DAILY YOU: HOW THE NEW ADVERTISING INDUSTRY IS DEFINING YOUR IDENTITY AND YOUR WORTH (2012).

50. CASS SUNSTEIN, REPUBLIC.COM (2001).

51. PARISER, *supra* note 14.

52. Daniel Kreiss, *Yes We Can (Profile You): A Brief Primer on Campaigns and Political Data*, 64 STAN. L. REV. ONLINE 70 (2012); Natasha Singer & Charles Duhigg, *Tracking Voters' Clicks Online to Try to Sway Voters*, N.Y. TIMES, Oct. 28, 2012, at A16.

freedom of association, and the transparency of the political process. "This 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."[53]

Not only political speech but also artistic and creative freedoms may be affected in a big data environment. For example, recent reports describe how Netflix harvests data from millions of users to produce content that best fits their tastes.[54] On the one hand, the television market has for many years thrived on a ratings system assessing the popularity of shows based on eyeballs. On the other hand, the ability to amass granular information regarding individuals' viewing habits and target specially tailored content at them raises concerns over siloization and narrowcasting. As Joseph Turow puts it, "the industrial logic behind the[se] activities makes clear that the emerging marketplace will be far more an inciter of angst over social difference than a celebration of the 'American salad bowl.'"[55]

Quite disturbing in this context,[56] is the fact that the machine is covered by an opaque veil of secrecy, which is backed by corporate claims of trade secrecy and intellectual property. In the analogue world, we could typically understand the logic underlying political advertising, credit or employment decisions; whereas in the big data environment, we are cowed into submission by a powerful data infrastructure, a "surveillant assemblage,"[57] delivering practically uncontestable results. This sense of being judged by the tin man, a heartless machine that operates based on incomprehensible criteria, is troubling.[58] It raises the specter of vulnerability and helplessness that accompanied Franz Kafka's anti-hero Joseph K., who was confounded by an opaque, logically

53. Kreiss, *supra* note 52, 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. COMMC'N 1032 (2010).

54. Andrew Leonard, *How Netflix is Turning Viewers into Puppets*, SALON (Feb. 1, 2013, 5:45 AM), http://www.salon.com/2013/02/01/how_netflix_is_turning_viewers_into_puppets.

55. JOSEPH TUROW, NICHE ENVY: MARKETING DISCRIMINATION IN THE DIGITAL AGE 2 (2006).

56. *See* Allison Brennan, *Microtargeting: How Campaigns Know You Better than You Know Yourself*, CNN (Nov. 5, 2012, 6:45 PM), http://www.cnn.com/2012/11/05/politics/voters-microtargeting ("When asked if they wanted political advertising tailored to your interests, 86% of Americans surveyed said they did not . . . 64% said their support for a candidate would decrease if they found out a candidate was micro-targeting them differently than their neighbor.").

57. Cohen, *supra* note 16, at 10.

58. Valentino-DeVries, *supra* note 28 ("It is difficult for online shoppers to know why, or even if, they are being offered different deals from other people. Many sites switch prices at lightning speed in response to competitors' offerings and other factors, a practice known as 'dynamic pricing.'").

baffling bureaucracy trying him for an unknown charge.[59] And while perhaps tolerable when restricted to the marketing context, such opaque decision-making tools threaten to pose a risk to democracy and free speech when introduced into the political sphere.

(LACK OF) REGULATORY REFORM

Against the backdrop of these challenges, policymakers have struggled to come up with a coherent regulatory response. Over the past two years, the OECD, EU, and US have launched extensive processes for comprehensive reform of their privacy frameworks.[60] Yet the result of these processes remains strongly anchored in the existing policy framework, which is rooted in an architecture dating back to the 1970s.[61] The major dilemmas and policy choices for informational privacy in the age of big data remain unresolved.

Specifically, privacy and data protection laws are premised on individual control over information and on principles such as data minimization and purpose limitation. Yet it is not clear that minimizing information collection is always a practical approach to privacy in the age of big data. To the contrary, data minimization appears inimical to the very concept of big data. And the discussion over individual control, which is closely linked (through the consent requirement) to principle of purpose limitation, too often transforms into an arena for highly charged polemics between industry and privacy advocates over what the public "really" wants.[62] The recent legislative reform proposals in Europe,

59. FRANZ KAFKA, THE TRIAL (Oxford Univ. Press 2009) (1925).

60. *See* Omer Tene, *Privacy Law's Midlife Crisis: A Critical Assessment of the Second Wave of Global Privacy Laws*, 74 OHIO ST. L. J. (forthcoming 2013) (describing the reform processes); *see also* ORG. FOR ECON. CO-OPERATION & DEV., THIRTY YEARS AFTER THE OECD PRIVACY GUIDELINES 93 (2011), *available at* http://www.oecd.org/sti/interneteconomy/49710223.pdf; 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* http://www.whitehouse.gov/sites/default/files/privacy-final.pdf; FED. TRADE COMM'N., PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS (2012), *available at* http://ftc.gov/os/2012/03/120326privacyreport.pdf; EUROPEAN COMM'N, PROPOSAL FOR A REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL ON THE PROTECTION OF INDIVIDUALS WITH REGARD TO THE PROCESSING OF PERSONAL DATA AND ON THE FREE MOVEMENT OF SUCH DATA (GENERAL DATA PROTECTION REGULATION) (2012), *available at* http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf.

61. Tene, *supra* note 60; *see also* Omer Tene, *There is No New Thing Under the Sun*, CONCURRING OPINIONS (July 30, 2012, 7:47 PM), http://www.concurringopinions.com/archives/2012/07/there-is-no-new-thing-under-the-sun.html.

62. Natasha Singer, *Do Not Track? Advertisers Say 'Don't Tread on Us'*, N.Y. TIMES, Oct. 13, 2012, at BU 3 (discussing the "correct" default setting for the "do not track" mechanism in the W3C Tracking Protection Working Group).

which all but outlaw data-based profiling, appear detached from technological and business realities and impossible to operationalize.[63]

When trying to solve the big data conundrum, it is easy to swing to extremes ranging from techno-utopianism on the one hand[64] to alarmist fear mongering on the other. Alas, technological, business, social, and ethical realities will force us to more carefully tread a path towards a nuanced reconciliation of big data benefits with individual rights. Clearly, the principles of privacy and data protection must be balanced against additional societal values such as public health, national security and law enforcement, environmental protection, and economic efficiency. Despite the heated rhetoric,[65] this remains true regardless of whether privacy is viewed as a consumer protection issue, as is often the case in the United States, or as a fundamental human right, as in Europe. Even fundamental rights are seldom absolute and often need to accommodate competing rights and interests.[66] In this part, we lay out several potentially useful directions for progress, focusing on empowering individuals by enhancing transparency and accountability.

63. *See* EUROPEAN COMM'N, *supra* note 60, at 20 (imposing strict restrictions on "profiling"); COMM. ON CIVIL LIBERTIES, JUSTICE & HOME AFFAIRS, EUROPEAN PARLIAMENT, DRAFT REPORT ON THE PROPOSAL FOR A REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL ON THE PROTECTION OF INDIVIDUAL WITH REGARD TO THE PROCESSING OF PERSONAL DATA AND ON THE FREE MOVEMENT OF SUCH DATA (GENERAL DATA PROTECTION REGULATION) 65 (Jan Philipp Albrecht ed. 2009), *available at* http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/pr/922/922387/922387en.pdf (these restrictions would be further tightened according to the draft submitted by the European Parliament Rapporteur, which adds to Article 4 of the General Data Protection Regulation a definition of "profiling": "any form of automated processing of personal data intended to evaluate certain personal aspects relating to a natural person or to analyse or predict in particular that natural person's performance at work, economic situation, location, health, personal preferences, reliability or behaviour."); *Id.* at 32 (noting "a general ban is introduced on profiling as defined in Article 4 and it is only permissible where provided for by law, i.e., either by means of the data subject's consent or a statutory provision.").

64. Cohen, *supra* note 16, at 15 ("Some of the claims on behalf of Big Data, those framed in terms of a 'singularity' waiting in our soon-to-be-realized future, sound quasi-religious, conjuring up the image of throngs of dyed-in-the-wool rationalists awaiting digital rapture.").

65. *See, e.g.*, James Fontanella-Khan, *Brussels fights US data privacy push*, FIN. TIMES (Feb. 10, 2013, 8:30 PM), http://www.ft.com/intl/cms/s/0/903b3302-7398-11e2-bcbd-00144feabdc0.html#axzz2KmNKUWab (noting "Europe's most senior justice official is adamant she will fight US attempts to water down a proposed EU data protection and privacy law that would force global technology companies to obey European standards across the world. Viviane Reding, EU commissioner for justice, said that the EU was determined to respond decisively to any attempts by US lobbyists – many working for large tech groups such as Google and Facebook – to curb the EU data protection law").

66. S*ee* John Morijn, *Balancing Fundamental Rights and Common Market Freedoms in Union Law: Schmidberger and Omega in the Light of the European Constitution*, 12 EUR. L. J. 15, 24 (2006).

OBSCURITY – IN PRAISE OF FUZZINESS

One promising path is the concept of obscurity, allowing individuals to hide in plain sight. Individuals are far less troubled by data analysis processes that do not *single them out* from a group. Stutzman and Hartzog note that "for an individual to be obscure, an observer must not possess critical information that allows one to make sense *of* the individual."[67] In the context of big data, this can be achieved through various means of de-identification, preventing the metaphorical camera lenses from focusing on a particular individual. Indeed, this approach can be viewed as a reconceptualization of Warren and Brandeis' "right to be let alone."[68] One forceful technique is *differential privacy*, which allows researchers to draw lessons and derive valuable conclusions from a data set without being able to determine whether or not such conclusions are based on the personal data of any given individual.[69] Hence, differential privacy emphasizes not whether an individual can be directly *associated* with a particular revealed value; but rather the extent to which any revealed value *depends* on an individual's data. Another technique is *k-anonymity*, which requires that the data for each person contained in a data release cannot be distinguished from at least k-1 individuals whose information also appears in the dataset.[70] In a previous article, we have argued that there are limits to de-identification in the context of big data.[71] While we realize that de-identification is not a panacea, we recognize that there is a broad range of situations where it can be a mitigating precaution.

A more proactive approach, referred to by Stutzman and Hartzog as "obscurity by design" would mask personal information behind a veil of obscurity through means such as pseudonymization, restricted access policies and limited searchability.[72] This would allow information to be shared usefully while at the same time minimizing privacy risks. Similarly, privacy enhancing measures can be integrated into new

67. Woodrow Hartzog & Frederic Stutzman, *The Case for Online Obscurity*, 101 CAL. L. REV. (forthcoming 2013).

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

69. Cynthia Dwork, Differential Privacy (2006) (unpublished manuscript), *available at* http://www.dbis.informatik.hu-berlin.de/fileadmin/lectures/SS2011/VL_Privacy/Differential_Privacy.pdf.

70. Latanya Sweeney, *k-Anonymity: A Model For Protecting Privacy*, 10 INT'L J. UNCERTAINTY, FUZZINESS & KNOWLEDGE-BASED SYS. 557 (2002).

71. Tene & Polonetsky, *supra* note 3; *see also* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701 (2010) (the seminal article advocating de-identification skepticism); Felix Wu, *Privacy and Utility in Data Sets*, 84 U. COLO. L. REV. (forthcoming 2013).

72. Fred Stutzman & Woodrow Hartzog, Obscurity by Design: An Approach to Building Privacy into Social Media (2012) (unpublished manuscript) *available at* http://fredstutzman.com/papers/CSCW2012W_Stutzman.pdf.

technologies to minimize their privacy footmark. For example, Snapchat is a mobile application that enables users to share photos or videos that vanish several seconds after being viewed by recipients.[73] If a recipient manages to capture a screenshot of a flickering photo, the sender is promptly notified by the app. Hence, Snapchat manages to achieve by nimble design what the European legislators seek to impose by regulation, namely a "right to be forgotten" or to erase one's digital trail.[74]

ACCESS AND TRANSPARENCY

A second promising path entails empowering individuals by granting them access to their personal data in intelligible, machine-readable form. Individuals would thus become active participants in the big data economy, analyzing their own information to improve their health, finances, career prospects, traffic management and more. Through mechanisms such as personal clouds or data stores, individuals could contract with third parties who would get permission to selectively access certain categories of their data to provide analysis, value-added services and mash-ups. We have called this the "featurization" of big data,[75] making data analysis a consumer-side application and unleashing a wave of innovation in the market for personal data applications.[76] Indeed, the thriving market for mobile apps provides ample proof that user-side installs work in real life.[77] This "sharing the wealth" strategy is justified by both efficiency and fairness concerns. In addition, it will benefit not only individuals but also businesses, which will get access to higher quality data about individuals' expressed intentions as opposed to guessing such intentions by analyzing online clues.[78]

73. *See, e.g.*, Jenna Wortham, *A Growing App Lets You See It, Then You Don't*, N.Y. TIMES, Feb. 9, 2013, at A1.

74. *See* Jeffrey Rosen, *The Right to be Forgotten*, 64 STAN. L. REV. ONLINE 88 (2012); Peter Fleischer, *Foggy Thinking About the Right to Oblivion*, PETER FLEISCHER: PRIVACY…? BLOG (Mar. 9, 2011), http://peterfleischer.blogspot.co.il/2011/03/foggy-thinking-about-right-to-oblivion.html.

75. *Id.*

76. The pioneering work in this field is by Doc Searls. *See* DOC SEARLS, THE INTENTION ECONOMY: WHEN CUSTOMERS TAKE CHARGE (2012); RICK LEVINE, CHRISTOPHER LOCKE, DOC SEARLS & DAVID WEINBERGER, THE CLUETRAIN MANIFESTO: THE END OF BUSINESS AS USUAL (2000).

77. *See, e.g., iOS v Android: App Revenues, Downloads and Country Breakdowns*, GUARDIAN APPS BLOG (Dec. 4, 2012), http://www.guardian.co.uk/technology/appsblog/2012/dec/04/ios-android-revenues-downloads-country; Joel Rubinson, APPNATION & RUBINSON PARTNERS INC., HOW BIG IS THE US APP ECONOMY? ESTIMATES AND FORECASTS 2011-2015 (2011), *available at* http://www.slideshare.net/joelrubinson/an3-us-appeconomy20112015.

78. *See* Ira Rubinstein, *Big Data: The End of Privacy or a New Beginning*?, 3 INT'L DATA PRIVACY L. (forthcoming 2013).

A third path involves a different type of transparency—with respect to businesses' data analysis processes. Danielle Citron set the stage for the discussion of "technological due process," pointing-out that "automated systems jeopardize due process norms."[79] It is hard to audit, challenge or amend processes that are concealed in a black box. We propose that businesses be required to reveal the *criteria* used in their decision-making processes, if not the actual algorithms that may be subject to protection of trade secrets and other intellectual property rights.[80] As Louis Brandeis once wrote, "[s]unlight is said to be the best of disinfectants."[81] We trust that if the existence and uses of databases were visible to the public, businesses would more likely avoid unethical or socially unacceptable (albeit legal) uses of data. In certain cases, such as micro-targeting election campaigns, simply shining the light to expose different communications made to specific audiences may provide the necessary check on concerns of inappropriate pandering to constituencies. In other contexts, where the machine makes binding determinations as to individuals' legal rights, due process requires that the subjects of such decisions are able to challenge them.

CLASSIFICATION OF HARMS

In order to tailor appropriate responses to big data problems, policymakers need to better define the risk of harm model. The regulatory toolbox to address privacy problems (e.g., notice and choice; data retention limitations) does not necessarily answer, and in fact may exacerbate, other harms such as fairness and discrimination.[82] Given the blurry edges of the concept of privacy, privacy harms are notoriously difficult to categorize.[83] Yet without such categorization, privacy policy

79. Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1249 (2008).

80. *See, e.g.*, European Data Protection Directive, *supra* note 41, art. 12(a) (requiring organizations to provide an individual with "knowledge of the logic involved in any automatic processing of data concerning him at least in the case of the automated decisions. . . ."); *see also* Dodd-Frank Wall Street Reform and Consumer Protection Act, 2010, § 1100F Pub. L. No. 111-203 (2010) (codified as amended at 15 U.S.C. § 1681(m) (2012)) (requiring lenders to disclose to borrowers information used to in risk-based pricing decisions, including any numerical credit score used; the range of possible scores; and key factors that adversely affected the borrower's credit score).

81. Louis Brandeis, *What Publicity Can Do*, HARPER'S WKLY., Dec. 20, 1913, at 10, *available at* http://3197d6d14b5f19f2f440-5e13d29c4c016cf96cbbfd197c579b45.r81.cf1.rackcdn.com/collection/papers/1910/1913_12_20_What_Publicity_Ca.pdf.

82. For example, in order to comply with rules on affirmative action, certain organizations are compelled to collect and retain information about individuals' gender or race. In these cases, data deletion, while privacy protective, would be counter-productive.

83. *Contra* M. Ryan Calo, *The Boundaries of Privacy Harm*, 86 IND. L. J. 1131 (2011); Daniel Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477 (2006).

can become muddled with peripheral or even conflicting considerations. For example, as currently framed the European "right to be forgotten" may be viewed as affording protection for one's reputation rather than privacy; and the right to "data portability" arguably belongs in the sphere of competition – not privacy law.

A harms-based approach to privacy need not be limited to pecuniary or tangible harms. A better understanding of the effect of data analysis on fairness, discrimination, siloization and narrowcasting can expand the scope of privacy harms that are subject to legal protection. Cynthia Dwork and Deirdre Mulligan refer to fairness concerns heavily weighted by issues of discrimination, including price discrimination based on location (redlining) or on knowledge of the consumer's state of mind.[84] Jules Polonetsky and I point out that processing of personal data increasingly affects fairness, equality, and other values, which are no less important than—even if theoretically distinct from—core privacy interests.[85]

PUTTING DATA IN CONTEXT

A final response involves the concept of context, which is based on Helen Nissenbaum's "contextual integrity" analysis of privacy.[86] Privacy, according to Nissenbaum, is "a function of several variables, including the nature of the situation or context; the nature of information in relation to that context; the roles of agents receiving information, their relationships to information subjects; on what terms the information is shared by the subject and the terms of further dissemination."[87] This approach may require, for example, that certain categories of sensitive data (e.g., genetic data) be segregated from the decision-making process in certain contexts (e.g., employment applications). Where to draw the contextual line becomes a weighty policy question where considerations of national security or public health are involved. In these cases, involving, for example, harvesting of social networking information to detect potential terrorist threats[88] or analyzing search engine logs to

84. Cynthia Dwork & Deirdre Mulligan, Aligning Classification Systems with Social Values through Design (June 8, 2012) (unpublished manuscript) (manuscript on file with authors).

85. Tene & Polonetsky, *supra* note 3.

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

87. Helen Nissenbaum, *Privacy as Contextual Integrity*, 79 WASH. L. REV. 119, 155 (2004); *see also* FED. TRADE COMM'N., PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS (2012), *available at* http://ftc.gov/os/2012/03/120326privacyreport.pdf.

88. *See, e.g.*, Ryan Gallagher, *Software That Tracks People on Social Media Created by Defense Firm*, GUARDIAN (Feb. 10, 2013), http://www.guardian.co.uk/world/2013/feb/10/software-tracks-social-media-defence.

analyze harmful drug interactions,[89] individuals' privacy interests may be outweighed by public policy concerns.

Moreover, as Nissenbaum recognizes, relationships and therefore context can change over time. Some argue that soliciting express consent should be a prerequisite to *any* shift in existing boundaries. In reality, however, shifting contexts are not always readily negotiated. Rather, organizations should assess the effects of any prospective change on data subject expectations; convey their policies clearly and conspicuously; and in certain cases provide data subjects with an opportunity to opt out. When a change in context is radical and transparency measures inadequate to support it, express consent can be relied upon to ensure that data subjects are willing to accept a new data use.[90]

CONCLUSION

As we recognize the immense benefits of big data, we should avoid technological determinism that allows the machine to surge forward with disregard for evolving social norms. Instead of asking "what technology wants,"[91] we should explore what it is that *we want* to achieve with technology and what price we are, or are not, willing to pay in privacy, social cohesion, and individual rights. The lack of agreement in the effort to standardize a "Do Not Track" protocol demonstrates the challenge in seeking a technological solution when the value of the activity to be proscribed remains widely disputed.[92] Hence, we must first address the ethics and morality of the decisions that confront us. Practically, we need to devise agreed-upon guidelines for ethical data analysis and profiling, addressing such issues as obscurity by design; empowerment through useful access; transparency of decisional criteria; and categorization of potential harms. Technology innovators and data scientists will lead the way to new big data frontiers, but it is philosophers seeking "a new digital humanism"[93] who must closely follow in their footsteps.

89. *See, e.g.*, Nicholas Tatonetti, Guy Haskin Fernald & Russ Altman, *A Novel Signal Detection Algorithm for Identifying Hidden Drug-Drug Interactions in Adverse Event Reports*, 19 J. AM. MED. INFORMATICS ASS'N. 79 (2012).

90. Jules Polonetsky & Omer Tene, *It's Not How Much Data You Have, But How You Use It: Assessing Privacy in the Context of Consumer Data Integration*, *in* FUTURE OF PRIVACY FORUM, (2012), *available at* http://www.scribd.com/doc/115516310/It-s-Not-How-Much-Data-You-Have-But-How-You-Use-It-Assessing-Privacy-in-the-Context-of-Consumer-Data-Integration; *see also* Paul Ohm, *Branding Privacy*, 97 MINN. L. REV. (forthcoming 2013).

91. Kevin Kelly, WHAT TECHNOLOGY WANTS (2010).

92. Omer Tene & Jules Polonetsky, *To Track or 'Do Not Track': Advancing Transparency and Individual Control in Online Behavioral Advertising*, 13 MINN. J. L. SCI. & TECH. 281 (2012).

93. JARON LANIER, YOU ARE NOT A GADGET: A MANIFESTO (2010).