I/S: A JOURNAL OF LAW AND POLICY FOR THE INFORMATION SOCIETY

Putting Data Benefits in Context: A Response to Kift and Nissenbaum

GABE MALDOFF & OMER TENE*

CONTENTS

I. INTRODUCTION

This essay examines big data analytics practices in light of Helen Nissenbaum's contextual integrity paradigm. It critiques Paula Kift and Nissenbaum's paper, *Metadata in Context – An Ontological and Normative Analysis of the NSA's Bulk Telephony Metadata Collection Program*.[1] It argues that while asserting that the NSA's bulk surveillance program imposes a privacy violation – something that few if any commentators, even in the intelligence community dispute – the paper glosses over the critical next step in the analysis, assessing privacy costs against data benefits. Such cost benefit analysis

* Gabe Maldoff is an associate at Bird & Bird and formerly Westin Fellow at the International Association of Privacy Professionals (IAPP); Omer Tene is Associate Professor, College of Management School of Law, Rishon Lezion, Israel, and Vice President of Research of Education at the IAPP.

[1] Paula H. Kift & Helen F. Nissenbaum, *Metadata in Context – An Ontological and Normative Analysis of the NSA's Bulk Telephony Metadata Collection Program*, 13 I/S J.L. & POL'Y FOR INFO. SOC'Y 333(2017).

underlies existing privacy frameworks, including: the FTC's unfairness doctrine; the European "legitimate interest of the controller" test; and Nissenbaum's own contextual integrity analysis. This essay suggests that while previous scholarship and standard frameworks have focused on developing taxonomies and analyses of privacy harms, they have paid cursory attention to categorizing and weighing data rewards. An assessment of benefits is essential for a complete analysis of controversial data practices in both the government and business contexts.

II. BIG DATA – AND "METADATA IN CONTEXT"

Big data continues to grate against existing techno-social norms. While even its most forceful advocates simultaneously caution against its risks,[2] the expansion of big data collection and use remains relentless. Fueled both by speculative projections and by proven results, it has become commonplace to find data analysts in almost every organization, from retailers, manufacturers and hospitals to municipal and national governments, political and religious organizations, and policing and security services. Data scientists uncover trends and correlations that could not have been identified without advances in computing power and the accumulation of previously unfathomable quantities of data.

Big data threatens to undo the Fair Information Principles, which have long defined privacy law and best practices.[3] While in the past, privacy protections revolved around providing individuals with notice of a data practice and the choice to opt in or out, it has now become increasingly difficult to predict and disclose how personal information will be used. Not only do many data analytic techniques eschew the hypothesis testing model of the analog world, but the rise of artificial intelligence and machine learning has made it virtually impossible to predefine the trajectory of data analysis.[4] Big data, therefore, calls for

---

[2] ALEC ROSS, THE INDUSTRIES OF THE FUTURE 152 (2016).

[3] 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); Ira S. Rubinstein, *Big Data: The End of Privacy or a New Beginning?*, 3 INT'L DATA PRIVACY L. 74 (2013); EXEC. OFF. OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES (2014), https://permanent.access.gpo.gov/gpo64868/big_data_privacy_report_may_1_2014.pdf [https://perma.cc/8C5N-P65U].

a new framework for analyzing the privacy implications of data practices.

Helen Nissenbaum's seminal book, *Privacy in Context*,[5] took on this challenge. In it, Nissenbaum argued that information privacy is fundamentally contextual. Privacy depends on an individual's reasonable assumptions about the social context in which information will be used. An individual will feel aggrieved if her information is used in a manner that violates her reasonable expectations. As shorthand, Nissenbaum defined this as a right to "contextual integrity."

This essay presents big data from the point of view of contextual integrity. As Paula Kift and Nissenbaum demonstrate in *Metadata in Context*, contextual integrity is extremely effective at identifying when a new practice will impact privacy. The framework, however, provides little guidance on when a practice may nonetheless be justified in spite of its impact on contextual integrity. By its very nature, big data often presumes an impact on the contextual integrity of an information flow to gain new data insights. The legitimacy of data analytics will thus depend on whether the benefits of a new practice outweigh its risks. Alas, without the tools for assessing data benefits, privacy protections may be eroded by false data promises just as technological progress may be squandered by undervaluing potential gains.

III. UNPACKING CONTEXTUAL VIOLATIONS

Contextual integrity is a two-step analysis. First, a reviewer must assess the effect of a new practice on the *actors*, *attributes*, and *transmission principles* involved in an information flow. If these factors are not altered, then the information practice passes the test. If, however, the practice alters any of these factors, then there is a "*prima facie* violation" and the analysis proceeds to the second step. With this second step, contextual analysis "requires an evaluation of the moral and political factors affected by the [new practice] and whether or not the benefits of altering information flows in this way

---

[4] NAT'L SCI. & TECH. COUNCIL COMM. ON TECH., EXEC. OFF. OF THE PRESIDENT, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE (2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf [https://perma.cc/7TFB-DBMV].

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

justifies potential costs in light of contextually specific goals and ends."[6] In other words, contextual integrity calls for *cost-benefit analysis* where a new practice imposes a privacy risk.

In *Metadata in Context*, Kift and Nissenbaum deploy this framework to examine the NSA's telephony metadata collection program. Using contextual integrity, the authors challenge existing legal doctrines that have shielded metadata collection from constitutional scrutiny. Namely, the authors question the distinctions between content and non-content, private information and business records, and concealed versus plainly visible information, arguing that these tropes fail to account for material technological changes that have changed the context of metadata collection over the past two decades as information technology has become ubiquitous.

Under their analysis, the metadata program altered the *transmission principles* of information flows because subscribers did not knowingly or voluntarily share metadata with the government. It also altered the *attributes* of metadata because, by aggregating metadata and applying data analytics to it, the NSA could glean from the resulting mosaic insights that could not be identified from any single tile. Finally, while subscribers may have expected telephone service providers to have access to metadata, by sharing the data with the government, these companies expanded the range of *actors* involved in the information flow.

To be sure, not all big data analytics will violate this first step of the contextual integrity test. But, many practices will because, at its zenith, big data analytics relies on acquiring unforeseen insights from an ever expanding pool of data sources. The transmission principles, attributes and actors involved in an information flow are susceptible to change without notice as information is collected, repackaged and shared in novel ways. Big data analytics breaks contextual integrity almost by definition.

By relying on social expectations, this first step of contextual integrity is vulnerable to evolving attitudes and understandings about data protection. As big data becomes the new norm, there is a risk that privacy protections are eroded with acquiescence of contextual integrity. In other words, if individuals come to expect unexpected data practices, privacy protections will be lost. A similar vulnerability afflicts constitutional Fourth Amendment protection under the *Katz v.*

---

[6] Kift & Nissenbaum, *supra* note 1, at 367-68.

United States decision.[7] *Katz* established a two-part test to measure whether a person has a "reasonable expectation of privacy," including a subjective prong, checking whether "a person [has] exhibited an actual (subjective) expectation of privacy[,]" and an objective prong, verifying whether "the expectation [is] one that society is prepared to recognize as 'reasonable.'"[8] In setting the dials of constitutional protection to the tune of societal expectations, *Katz* created a self-fulfilling prophecy where the less privacy (or more surveillance) individuals expect, the less constitutional protection they are entitled to.[9]

In *Metadata in Context*, Kift and Nissenbaum deftly overcome this challenge by focusing on the voluntariness of one's participation in a data practice, rather than knowledge of it. In the case of NSA metadata collection, since the social costs of opting out of electronic communications are so high – subscribers have neither a reasonable alternative to using their devices nor any control over the metadata that those devices create in the course of their operation – they do not *voluntarily* share the metadata. And even to the extent that subscribers are aware of the metadata they share, they cannot reasonably predict the type of inferences that the government may be able to draw from the data using secretive analytic techniques. Moreover, the government's vast capabilities of aggregation and analysis fundamentally alter the nature of the information, from discrete points of metadata to a rich tapestry from which the government can draw consequential inferences.

The same holds true for many private sector data practices. The fact that an individual might expect that data will be collected and used in any number of ways does not mean that the individual has a deep enough understanding of the practice to fully accept it. Indeed, as a study by Turow, Hennessy and Draper revealed, the more individuals know about what marketers do with their personal information, the more likely they were to feel resigned about sharing it and to resist rational decision-making with regard to privacy trade-

---

[7] Katz v. United States, 389 U.S. 347 (1967).

[8] *Id.* at 361 (Harlan, J., concurring).

[9] *See, e.g.*, Omer Tene, *What Google Knows: Privacy and Internet Search Engines*, 2008 UTAH L. REV. 1433 (2008).

offs.[10] Thus, just like for the NSA's metadata program, the legitimacy of data analytics frequently will turn on the second step of contextual integrity: cost-benefit analysis.

IV. THE CRITICAL SECOND STEP

The contextual integrity test "presumptively favors protecting the integrity of entrenched informational norms."[11] The result is that the second step of the analysis plays the role of gatekeeper for new data practices, legitimizing only those practices that can demonstrate an overriding benefit. This represents a marked departure from Fourth Amendment doctrine, which helps explain why contextual integrity results in a more rigorous analysis of metadata collection than that seen in some legal challenges.

Under the Fourth Amendment, a court must first ask whether a search intrudes upon a constitutionally protected "reasonable expectation of privacy" before deciding whether a search was reasonable. For litigants, that first step – proving that there was a reasonable expectation of privacy – often is the critical juncture. The first step of the analysis can excuse surveillance practices from constitutional scrutiny altogether. Most notably, as a result of the "third party doctrine," whole classes of digital surveillance have been excluded from Fourth Amendment analysis. Once a practice is found to constitute a "search," however, clear parameters exist for determining whether the practice is nonetheless "reasonable." In most situations, reasonableness is defined by the presence of warrant, based on probable cause, to validate search. Only in rare circumstances, where, for example, there is an exigency or some other overriding concern, will a search be valid in the absence of a warrant.

With contextual integrity, more practices will proceed to the second level of analysis, but Kift and Nissenbaum offer few clues as to how reasonableness would then be determined. Instead, they note merely that a presumptive contextual violation can be overcome only if new practices are demonstrably 'more effective at achieving

---

[10] JOSEPH TUROW, MICHAEL HENNESSY & NORA DRAPER, THE TRADEOFF FALLACY: HOW MARKETERS ARE MISREPRESENTING AMERICAN CONSUMERS AND OPENING THEM UP TO EXPLOITATION (2015), https://www.asc.upenn.edu/sites/default/files/TradeoffFallacy_1.pdf [https://perma.cc/TN3G-FKYQ].

[11] Kift & Nissenbaum, *supra* note 1, at 367.

[contextual] values, ends, and purposes[,]" or the equivalent.[12] The absence of defined terms for the analysis becomes evident when the authors move to assess the costs and benefits of the NSA's metadata program. According to the authors, the metadata program fails this second step. They point to the high financial costs of the program and the fact that it diverts resources from "traditional and targeted surveillance [techniques]."[13] More importantly, the NSA program inflicted material impacts on "civil liberties such as privacy, freedom of speech and association, transparency, due process and the balance of power between the government and its citizens."[14]

Against these impacts, Kift and Nissenbaum suggest that the NSA program's benefits were flimsy. While they note that intelligence representatives claimed that the metadata program thwarted more than fifty different terrorist attacks, they cite a report from the Privacy and Civil Liberties Oversight Board finding "only [one] case in which the bulk collection of telephony metadata played a significant role in the containment of terrorist activity."[15] Even this one case, the authors argue, is unconvincing.

Yet even as they criticize the intelligence community, Kift and Nissenbaum also weaken their own analysis. For it is obvious that under *any* cost-benefit analysis, zero benefits are outweighed by legitimate privacy concerns. But how does this arithmetic work if the benefits are non-trivial? For example, had the NSA's claim to have thwarted fifty attacks stood up, would that have justified the widespread privacy intrusion? And what if the NSA's program successfully prevented a single high-magnitude terrorist event? What about just one death of a child? Or a single death of an old man? *Metadata in Context* gives no guidance.

The authors also note that while the September 11 attacks provided the impetus for the program, the intelligence community failed to prevent the attacks "not because of insufficient information *collection* but because the FBI and NSA had an insufficient understanding of the rules that governed information *sharing* between intelligence agencies – information they already had thanks

---

[12] NISSENBAUM, *supra* note 5, at 180.

[13] Kift & Nissenbaum, *supra* note 1, at 370.

[14] *Id.*

[15] *Id.* at 369.

to conventional law enforcement techniques."[16] Ostensibly, by sharing existing data, the authors argue the intelligence community could have better achieved its goals with less intrusion upon privacy.

Yet it is not clear how this would square with Kift and Nissenbaum's model for contextual integrity. The number of individuals affected by an information practice is not a factor in contextual analysis. One of the theory's central insights, in fact, is that sharing or repurposing information may impact privacy as much as collection in the first place. Indeed, in some cases, mass, automated monitoring may have *less* impact on privacy than specific, targeted surveillance.[17] For example, when an intelligence agency zeroes in on an individual target, sharing that person's information among other agencies could lead to tangible harms, like being placed on no-fly list, being audited by the IRS, or, in an extreme scenario, being targeted by a drone strike.[18]

This may be a slip, for the authors' analysis appears to fit more closely with traditional privacy metrics that emphasize collection than it does with contextual integrity's focus on information use. It reveals, however, the absence of tools at the authors' disposal for rigorously assessing costs, and in particular, benefits, at the analytical step requiring cost-benefit analysis. In a contextual analysis of big data, too, we are bound to reach an equivalent stumbling block.

V. THE NEED FOR DATA BENEFIT ANALYSIS

Cost benefit analysis in privacy law is not unique to the second step of contextual integrity. Existing legal frameworks and organizational practices recognize the need to balance privacy risks against data benefits. These frameworks already provide detailed guidance on the many flavors of privacy risk. However, risks are only one part of the cost-benefit equation. To understand whether a data

---

[16] *Id.* at 368.

[17] Omer Tene, *A New Harm Matrix for Cybersecurity Surveillance*, 12 COLO. TECH. L.J. 391, 401 (2014).

[18] *See* Christian Grothoff & J. M. Porup, *The NSA's SKYNET Program May Be Killing Thousands of Innocent People*, ARS TECHNICA UK (Feb. 16, 2016), http://arstechnica.co.uk/security/2016/02/the-nsas-skynet-program-may-be-killing-thousands-of-innocent-people/ [https://perma.cc/LLA7-38JF].

practice is valid in light of contextual violations, decision-makers must also dissect, prioritize and quantify data benefits.

The need for data benefits analysis is evident in the FTC's use of its unfairness authority under Section 5 of the FTC Act. To find that an act or practice is unfair, the FTC must prove that "the act or practice causes or is likely to cause substantial injury to consumers which is not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition." Under this balancing test, the FTC has tried to articulate the contours of privacy harms, dissecting both the probability of privacy risks and the magnitude of potential harms in a line of enforcement cases.[19] But aside from broadly recognizing that "big data analytics can provide numerous opportunities for improvements in society,"[20] the agency has produced little guidance on how to quantify data benefits and assess them against corresponding risks.

This may explain why the FTC has scarcely used its unfairness authority (independently of a deception claim) to pursue violations other than a company's failure to implement appropriate data security practices.[21] In data security cases, the balancing test is simplified because inadequate security typically provides consumers with *no* discernible benefits. Any harm to consumers will easily outweigh the nonexistent benefits. Without some way of quantifying benefits, however, the agency may find it more difficult to act on practices that provide more than trivial benefits – a smartphone flashlight application[22] or more targeted ads.[23] This hamstrings the FTC's ability

---

[19] *See, e.g.*, Opinion of the Commission, LabMD, Inc., F.T.C. Docket No. 9357 (2016), https://www.ftc.gov/system/files/documents/cases/160729labmd-opinion.pdf [https://perma.cc/7NX2-TL8Q] ("This reading is supported by prior Commission cases applying the unfairness standard, which also teach that the likelihood that harm will occur must be evaluated together with the severity or magnitude of the harm involved.").

[20] FED. TRADE COMM'N, BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION? (2015), https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf [http://perma.cc/G4WF-BFGP].

[21] Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 594 (2014).

[22] Decision and Order, Goldenshores Technologies, LLC, F.T.C. Docket No. C-4446 (2014), https://www.ftc.gov/system/files/documents/cases/140409goldenshoresdo.pdf [https://perma.cc/H5W5-4F2F].

to use its unfairness authority in privacy cases, even as the complexity of data flows makes it more difficult to rely purely on consumer deception.[24]

The FTC's reliance on cost benefit analysis is likely to become more pronounced under the Trump administration. FTC Commissioner Maureen K. Ohlhausen called for renewed emphasis on cost benefit and economic analysis in one of her first speeches after taking on the role of acting chairman in 2017.[25] Under this view of Section 5, the FTC would have to analyze "more rigorously what constitutes 'substantial injury' in the context of information about consumers."[26] This, in turn, will give more importance to the FTC's Bureau of Economics ("BE"), which is tasked with analyzing the economic impact of consumer protection and competition investigations and rulemakings. One study found that BE's analysis in privacy cases has been limited by the difficulty of "assigning a dollar value to a privacy violation," on the one hand, and assessing the value of privacy regulation on the other.[27] The challenges of effective cost benefit have pushed the Commission to pursue enforcement under its deception authority, even where unfairness would fit more closely with the perceived violation.

Likewise, current privacy impact assessment (PIA) frameworks provide little insight into balancing impacts with the promises of new technologies. This led the National Institute for Standards and

---

[23] Decision and Order, Sears Holdings Management Corp., F.T.C. Docket No. C-4264 (2009), https://www.ftc.gov/sites/default/files/documents/cases/2009/09/090604searsdo.pdf [https://perma.cc/79DX-G862].

[24] CHRIS JAY HOOFNAGLE, FEDERAL TRADE COMMISSION PRIVACY LAW AND POLICY (2016).

[25] Maureen K. Ohlhausen, Opening Keynote at the ABA 2017 Consumer Protection Conference (Feb. 2, 2017), https://www.ftc.gov/system/files/documents/public_statements/1069803/mko_aba_consumer_protection_conference.pdf [https://perma.cc/8AR8-BL5M].

[26] Concurring Statement of Acting Chairman Maureen K. Ohlhausen, Vizio, Inc., F.T.C. Docket No. 1623024 (2017), https://www.ftc.gov/system/files/documents/public_statements/1070773/vizio_concurring_statement_of_chairman_ohlhausen_2-6-17.pdf [https://perma.cc/3SBG-P32K].

[27] Chris Jay Hoofnagle, *The Federal Trade Commission's Inner Privacy Struggle*, THE CAMBRIDGE HANDBOOK OF CONSUMER PRIVACY WELFARE (2017), https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2901526 [https://perma.cc/CY2H-6CF7].

Technology (NIST) to call for "additional tools that facilitate repeatable and measurable methods for identifying, prioritizing, and mitigating privacy problems."[28] In its final report on privacy engineering and risk management, NIST made significant strides at identifying privacy risks "that extend beyond unauthorized access to PII."[29] For example, in the context of big data, the NIST framework highlighted the risk of "unintended bias or discrimination in systems that determine eligibility for goods, services, and employment opportunities," as well as the chilling effect that "unanticipated revelations about individuals and their online connections and communities" could have on free speech.[30] Critically, however, while NIST found that "[c]ontext . . . is the foundation for the interpretative analysis necessary to understanding when a privacy boundary line has been crossed," the NIST framework provides no clues as to how to build on this foundation.[31] There is no discussion of data benefits or to cost-benefit analysis, outside of vague references to an "acceptable level of risk."[32]

This leaves organizations without the necessary guidance at the crucial stage, after the PIA, in which they must determine whether and how to proceed. For example, while discussing low-likelihood/high-impact activities and high-likelihood/low-impact activities, the draft NIST framework recommended "mitigation" and "controls" respectively to reduce the risk.[33] But not all risks can be mitigated or controlled. And, in some cases, even known and unavoidable harms may be justified. For example, Facebook's decision to implement the "News Feed" feature led to a "storm of protest"

---

[28] NAT'L INST. OF STANDARDS AND TECH., DEPT. OF COMMERCE, PRIVACY RISK MANAGEMENT FOR FEDERAL INFORMATION SYSTEMS, INTERNAL REPORT DRAFT 8062 (2015), http://csrc.nist.gov/publications/drafts/nistir-8062/nistir_8062_draft.pdf [https://perma.cc/EJ6Z-7RAP].

[29] NAT'L INST. OF STANDARDS AND TECH., DEPT. OF COMMERCE, AN INTRODUCTION TO PRIVACY ENGINEERING AND RISK MANAGEMENT IN FEDERAL SYSTEMS, NISTIR 8062 (2017), http://nvlpubs.nist.gov/nistpubs/ir/2017/NIST.IR.8062.pdf [https://perma.cc/6D6N-58KA].

[30] *Id.* at 39-40.

[31] *Id.* at 23.

[32] *Id.* at 13.

[33] NAT'L INST. OF STANDARDS AND TECH., *supra* note 28, at 23-24.

because it altered the context of social interactions, but ultimately it became the company's most popular feature.[34] Or, more importantly, big data analysis could conceivably help find a cure for a terminal disease or expedite disaster recovery efforts in an area struck by an epidemic, justifying a higher degree of privacy risk than data practices geared simply at improving ad targeting return on investment.

European data protection law provides a helpful lens for examining cost-benefit questions through the "legitimate interests" clause,[35] which will remain part of the General Data Protection Regulation (GDPR) when it comes into effect in 2018.[36] In a nod to contextual integrity, analyzing a controller's legitimate interests under the GDPR requires "consideration [of] the reasonable expectations of data subjects based on their relationship with the controller," including "whether a data subject can reasonably expect at the time and in the context of the collection of the personal data that processing for that purpose may take place."[37]

Unlike contextual integrity, however, legitimate interests *start* with an analysis of organizational interest, and only if there is a sufficient interest does the analysis proceed to examine whether that interest is "overridden by the interests or fundamental rights and freedoms of the data subject."[38] In exploring the application of this test, the Article 29 Working Party found that the concept of "interest" is broader than a "purpose," encompassing benefits not just derived by the controller, but also by society at large. Nonetheless, "interests that are too vague or speculative will not be sufficient."[39]

The fact that benefits are uncertain, however, should not block a project in *all* cases. Look at President Obama's billion-dollar pledge to cure cancer, which appears likely to survive into the Trump presidency

---

[34] NISSENBAUM, *supra* note 5, at 62.

[35] Council Directive 95/46/EC, art. 7(f), 1995 O.J. (L 281) 40, http://eur-lex.europa.eu/LexUriServ/ LexUriServ.do?uri=OJ:L:1995:281:0031:0050:EN:PDF [https://perma.cc/FX4C-DDSU].

[36] Council Regulation 2016/679, 2016 O.J. (L 119) 36.

[37] *Id.* at (L119) 9.

[38] *Id.* at (L 199) 36.

[39] Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC ["Legitimate Interests Opinion"], ARTICLE 29 DATA PROTECTION WORKING PARTY (2014).

– a goal at once improbable and yet so noble, it was named the "Cancer Moonshot."[40] Innovation – indeed, business more generally – inevitably requires taking chances. Any assessment of data benefits should allow for some uncertain benefits, while also weeding out those that are too improbable. Just as privacy costs are measured in *risk* of harm, so too must data benefit analysis include consideration of the *chance* of a data benefit. Polonetsky, Tene, and Jerome account for uncertain benefits in their framework for data benefits analysis by discounting the magnitude of a potential benefit against its likelihood of occurring, as risk analysis does for the risk of harms.[41]

Ultimately, however, the costs and benefits of any new practice fall differently upon individuals, communities, organizations, and society at large. The difference between justifiable risks and irresponsible risks depends not only on the odds of success and the magnitude of the wager, but also on who stands to win and who stands to lose. While it seems logical to ensure that the risk-bearer and beneficiary are one and the same, such a result is neither always possible nor desirable. For example, it may be necessary, and indeed courts have agreed, to mandate vaccination programs that benefit society at large, even while they impose costs on the few who object.[42]

As in many other legal contexts, these complex ethical issues necessarily raise the question of the burden of proof and who the decision maker should be.[43] The European system places its thumb on the scale in favor of individuals and legislatures. Individuals because an organization's legitimate interests may be overridden by those of the data subject, even if they are not necessarily legitimate.[44] Only

---

[40] Kim Smuga-Otto, *Will Biden's Cancer Moonshot Survive the Trump Administration?*, DISCOVER MAGAZINE D-BRIEF BLOG (Feb. 2, 2017), http://blogs.discovermagazine.com/d-brief/2017/02/02/cancer-moonshot-trump-administration/#.WMGoW9LyhaQ [https://perma.cc/W9B7-HWNV].

[41] JULES POLONETSKY, OMER TENE & JOSEPH JEROME, FUTURE OF PRIVACY FORUM, BENEFITS-RISK ANALYSIS FOR BIG DATA PROJECTS (2014), https://fpf.org/wp-content/uploads/FPF_DataBenefitAnalysis_FINAL.pdf [https://perma.cc/B97R-H9XS].

[42] Jared P. Cole & Kathleen S. Swendiman, *Mandatory Vaccinations: Precedent and Current Laws*, CONG. RES. SERV. (May 21, 2014), https://www.fas.org/sgp/crs/misc/RS21414.pdf [https://perma.cc/QCA9-6PTY].

[43] POLONETSKY ET AL., *supra* note 41, at 12.

[44] *Id.* at 6 ("Even individuals engaged in illegal activities should not be subject to disproportionate interference with their rights and interests.").

"compelling" legitimate interests will overcome an individual's objection.[45] This highlights the privileged role of legislators in this system as one way, and in practice, perhaps the only way, to demonstrate an interest is compelling is to point to a legislated policy objective. Thus, while legitimate interest analysis makes strides at evaluating privacy costs and benefits, ultimately it shies away at the critical moment, leaving it to the data subject or the legislature to make this assessment.

And yet the emerging dataverse, with its seamless, constant data flows, requires a broader circle of ethical decision-makers. Some decisions should be made only by those who will be affected. In other cases, it will be impossible to capture user consent in all the myriad ways that data is being collected and used, both online and in physical spaces such as smart cities or homes. While legislation may accommodate all the diverse interests of stakeholders affected by data decisions, legislatures simply cannot move quickly enough to meet changing technology, nor should they be expected to guide every data decision that occurs in any organization.

By identifying when a practice mismatches those who bear the risk and those who stand to benefit, data benefit analysis will facilitate selecting a decision-maker appropriate to the task. For some decisions, the ethical considerations and value judgments will require input from experts and ethical review committees.[46] Other data decisions, however, may be justified only by those who bear the risk.

VI. CONCLUSION

In *Metadata in Context*, Kift and Nissenbaum examined the NSA's metadata program in light of contextual integrity. This essay employed the framework to analyze big data practices, finding that new data practices, by definition, will alter informational context. Thus, for big data, the question is not whether there is a contextual violation, but rather, when is such a violation justified. Contextual integrity leaves open this question – the "second step" – to cost-benefit analysis, but it offers little guidance on how to tally data benefits. This problem arises not just with contextual integrity, but

---

[45] Council Regulation 2016/679, *supra* note 36, at (L 119) 45.

[46] Jules Polonetsky, Omer Tene & Joseph Jerome, *Beyond the Common Rule: Ethical Structures for Data Research in Non-Academic Settings*, 13 COLO. TECH. L.J. 333 (2015).

also with other frameworks, like the FTC's unfairness doctrine or Europe's legitimate interests test, that promote the assessment of inconsistently defined privacy risks against vaguely articulated data benefits.

Information privacy has reached a critical juncture. Aided by the work of Nissenbaum and others, the public is increasingly attuned to privacy risks inherent in new technologies. With the context of information flows constantly changing, it is time to focus on the next step, cost benefit analysis.[47]

[47] *See* FUTURE OF PRIVACY F. & PROGRAM ON ECON. & PRIVACY AT GEORGE MASON U. ANTONIN SCALIA L. SCH., CALL FOR PAPERS – DEVELOPING A BENEFIT-COST FRAMEWORK FOR DATA POLICY (2017), https://fpf.org/wp-content/uploads/2017/03/FPF_GMU-Call-for-Papers-BCF-FINAL.pdf [https://perma.cc/34UR-JQU8] (calling for submissions to address the absence of literature "focused on establishing a firm foundation for the type of benefit-cost analysis that seems baked into privacy's regulatory framework").