

A Theoretical Framework for Ethical Reflection in Big Data Research

Michael Steinmann, Sorin Adam Matei and Jeff Collmann

1 Introduction

Scientific progress has increasingly become reliant on large-scale data collection and analysis methodologies. The same is true for the advanced use of computing in business, government, and other areas. Utilizing massive computational resources, such methodologies can automatically capture and analyze characteristics and processes of entire statistical populations. Sampling is ideally replaced by census-like, complete counting of cases and characteristics. Interconnections between individual elements are turned into graph edges. Complete graphs facilitate research that takes into account case dependencies. They are ideal for detecting diffusion processes in a variety of populations.

When applied to objects of study at micro-scale, such as bio-molecular research, big data research can take instantaneous snapshots of extremely complex systems (genes, proteins, etc.), categorize them, and detect patterns or anomalies. When applied at a macro-scale, big data can involve remote sensing networks, including radar, satellites, or telescopes to capture real time information about highly complex phenomena, such as weather patterns, climate change, or mapping the depths of the universe.

M. Steinmann (✉)

College of Arts & Letters, Stevens Institute of Technology, Castle Point on Hudson,
Hoboken, NJ 07030, USA

e-mail: msteinma@stevens.edu

S.A. Matei

Brian Lamb School of Communication and Cyber Center, Purdue University,
100 N University Drive, West Lafayette, IN 47907, USA

e-mail: smatei@purdue.edu

J. Collmann

Department of Microbiology and Immunology, Georgetown University,
3700 O St NW, Washington, DC 20057, USA

e-mail: collmanj@georgetown.edu

In general, big data is concerned with the exhaustive capture of information about complex systems and the subsequent exploration and explanation by means of an extensive investigation of their elements and characteristics. The depth and comprehensiveness of such an approach are impressive, providing the ability to find the proverbial needle in the haystack. Answers both general and specific can emerge from observing all situations in which even small changes occur. Protein expression in genes is now used as key for opening the door to explaining specific diseases; temperature upticks in the Arctic measured hourly can capture global warming trends; and the propensity of individuals to search for “cold medicine” through Google can be an indicator that the next flu epidemic is about to start. Although critics have pointed out that not all promises of big data research might be realized at the end, the very fact that such promises exist already changes the expectations and attitudes toward research and can lead to all sorts of new experiments with the analysis of data.

The complete, detailed account of phenomena promised by big data represents a boon for research and poses ethical challenges of different kinds. Genetic profiling of entire populations, to take one example, may lead to finding new relationships between genes and disease, and also pinpoint specific individuals who carry such genes. If this is the case, should their genetic potential be reduced or eliminated in the name of public health? If applied to humans this would represent a return to eugenics, a theory and practice long since rejected. If applied to other forms of life, this can lead to a reduction in biological diversity. A search for patterns in open source data collection can be used to inoculate those most vulnerable for a disease before an epidemic has fully developed. But what if searches suggest the emergence of a political movement or protest?

Big data methodologies, thus, have a double potential, both for sharpening our scientific insights and for potentially creating significant ethical dilemmas. At the time being, dilemmas D related to privacy are probably the most important, as they implicate a series of ethical values. However, in the current discussion it seems far from clear what privacy means in each case and how it matters. In addition, due both to its rapidly evolving nature and the hypothetical character that many applications still possess, big data presents a specific challenge to the reflection on ethical issues. A one-size-fits-all approach does not seem appropriate for it. For these reasons, ethical reflection first has to target the various aspects of big data and their impact on human privacy in a differentiated way. We will try to do so by articulating the various normative dimensions that privacy entails. Second, ethical reflection has to address the long-term goals involved in big data research on human subjects. According to the forward-looking character of big data methodologies, our reflection has to be able to anticipate and preempt ethical issues in human subject research by articulating desirable practices and overarching values, such as trust in the integrity of scientific research.

Our methodical approach to ethical reflection on big data can be seen as a well-defined pluralism. This means, on the one hand, that ethical reflection has to address a variety of values, or principles, that cannot be further subsumed under one over-arching value. On the other hand, this ethical pluralism does not mean that in each case conflicting judgments have to be made, which would render ethical

reflection practically ineffective. Instead, each situation needs to be judged contextually. Following Nissenbaum (2011), we suggest placing strong emphasis on the context of using big data. Different contexts raise different concerns, which then require analysis with respect to different, specifically chosen ethical principles. This makes it possible both to make meaningful ethical judgments and address the underlying real-world problems in a multifaceted way. In addition, we suggest that special consideration has to be given to the process of decision-making, which often follows a trade-off model, or cost-benefit analysis. Given the dynamic and promising nature of big data, the trade-offs that are made are likely to raise ethical concerns of their own. To minimize such concerns, we propose that each trade-off analysis should not start until a minimum ethical threshold is determined, one under which certain core values should not be traded off for any possible social or individual benefits.

To understand the potential ethical impact of big data, we need to begin by inventorying the essential modalities of big data manipulation that may impact privacy. These are put under what we call the 4R rubric: reuse, repurpose, recombine, or reanalyze, which is detailed below (Sect. 2). We will then start our discussion of ethical challenges with some key definitions related to privacy and its various normative dimensions. These dimensions, born out of core human values, include the principles of non-maleficence, beneficence, justice, autonomy, and trust (Sect. 3). We will continue with a discussion of the contextual nature of privacy and how it motivates and shapes the specific treatment of privacy in each case. To simplify this discussion we propose a “privacy matrix,” which helps matching privacy dimensions and given contexts with specific types of trade-offs (Sect. 4). We will propose a heuristic model that uses trade-off analysis, yet overcomes the difficulties typically implied by the utilitarian logic of cost-benefit analyses. We propose a model that starts with determining a minimum “concern threshold” (Sect. 5).

2 The 4R Approach

The ethical impact of big data analysis is born out of two main concerns. On the one hand, big data tends to be exhaustive and precise. Big data deals not with samples, but with populations. Big data is like a sieve, most of the solid matter (objects, people, behaviors, etc.) that goes into it is captured. Only the liquid part of the universe of observation passes through. Also, big data is relational. Being able to count everything, it can determine if the characteristics of the elements under observation are shared and whether the shared characteristics create networks of affiliation or interaction. Elements are understood not only as static objects with certain characteristics, but as generative, evolutionary nodes that can impact or be impacted by other nodes. Ethically, when studying individuals, this means that we can know not only if people are of a certain kind but how susceptible they are to change and from which direction (connection) this change can come.

The other big concern related to big data is born out its ability to be reused, repurposed, recombined, or reanalyzed. This is what we call the 4R challenge of big

data. Given the connectedness of big data, its elements can be easily imagined as lego pieces ready to be rearranged and connected to other pieces or collections of pieces (populations) to obtain new insights. The new insights can produce new knowledge but also unforeseen threats for the individuals or population under observation. Since the individuals released their data for a specific use and goal, any further analysis that reveals new processes in that population can lead to insights that were not envisaged or considered desirable by the individuals who contributed the data. For example, a study of Google searches on “pain medicine” to identify arthritic patients that are underserved in specific geographic areas identifies a cluster of searches coming from a college town, where the population is much younger than the typical arthritic sufferer. One can then suppose that the searches are related to pain medication abuse, rather than legitimate use.

Or, suppose that the entire population in a given region, inhabited by a homogeneous native population, is screened genetically to identify the possibility of a relationship between genetic makeup and a given inheritable disease (sickle cell anemia). After the study, the data is also used to identify the genetic haplotypes present in the population. The conclusion is that a significant number of individuals have a genetic inheritance that is not common with that which is considered traditional for that native population. This may affect the notions of cultural identity and nativity.

Such examples, hypothetical only in form, as we will see below, reveal the tremendous ethical difficulties created by big data. In what follows, we will investigate them in more detail, focusing especially on privacy concerns generated by the 4R challenges.

2.1 Reuse

Reuse refers to taking data originally collected for a specific scientific purpose and using them again for comparable purposes in comparable domains. The reuse activities may engage either the original investigators or other investigators. The possibility of reuse, particularly of data originally acquired in scientific activities covered by the Common Rule (Office for Human Research Protection 1993), raises the question of the responsibilities that investigators have for what happens to data once they become available to secondary investigators. Additionally, it poses the question of the responsibilities of secondary investigators for complying with, or reaffirming the conditions of a data set’s original collection.

A Case: Reusing genomic data from the Havasupai Indians

Scientists at Arizona State University conducted a series of investigations on blood samples obtained from the Havasupai Indians, a small tribe of people living at the bottom of the Grand Canyon. The studies began when the Havasupai approached ASU for help in understanding the high prevalence of diabetes among their people. In addition to conducting research on the possible existence of a genetic basis for diabetes, ASU scientists re-used blood samples drawn from individual members of

the Havasupai tribe to conduct and publish results on multiple studies of which the Havasupai had no knowledge on a range of other disorders and characteristics. Upon accidentally hearing of the secondary use of the blood samples, the Havasupai sued ASU who, after paying over \$1.7 M in legal fees, settled the case by paying \$700,000 in reparations and providing other benefits in kind (Harmon 2010; Jacobs et al. 2010).

Ethical implications: The Havasupai claimed harms of honor and cultural integrity, finding specific fault with papers purporting high level of inbreeding among the Havasupai and published claims about their origins at odds with their own traditions.

2.2 Repurposing

In contrast to reuse, repurposing refers to taking data originally collected for a specific purpose in a specific domain and analyzing them for unrelated purposes in a domain other than their domain of origin. In addition to the questions posed by reusing data, repurposing big data poses questions about the legitimacy of analyzing data acquired under one privacy context and employing it in a different privacy context.

A Case: Repurposing educational administrative data for scientific research

Social science investigators are finding great value in linking administrative records from multiple administrative data systems and entities to longitudinal datasets at different levels of analysis (individual, family, program, school, etc.). State and local agencies collect the data for multiple purposes, such as program accountability, client tracking, and service effectiveness. “Data” in this case refers to administrative records, established when a person or family applies for social, health or educational services (such as enrolling a child in school). Most states also routinely collect data on newborn babies, their parents (especially mothers), including prenatal care; any birth defects or signs of vulnerable health (e.g., hearing loss registry); and other important social indicators. Public health departments collect a range of data and routinely work with hospitals, clinics and other providers in tracking persons to ensure adequate care and provision of services as well as effective disease monitoring.

Ethical implications: This type of research highlights the importance of considering ethical provenance in employing big data. The scientific investigators have to consider the impact for their own analyses of the conditions under which the data was originally collected, specifically that sometimes the records are used differently from their original purpose. Data that were originally collected without consent for use in research can potentially be repurposed without the knowledge of individuals that the records concern. As the distance grows between the original data source and its eventual uses in research, the gap potentially also grows between what individuals initially expected to happen with their information and the research

that might actually be conducted. The NPRM on the Common Rule has recommended that a specific form of consent, Broad Consent, be developed to address these issues but has not yet offered any details. Depending on what actually gets developed, the Broad Consent mechanism might address the question of ethical provenance because it enables notice; but to be effective, it should offer an opportunity to shield one's records at any point in the process from initial data acquisition to incorporation in a research project.

2.3 *Recombining*

The term big data frequently evokes a process of combining and recombining data from various sources to achieve greater analytic yield. In addition to the questions posed by reusing and repurposing data, recombining data poses questions about the possibility of developing new information not available to the investigator simply from the constituent data sets. From a privacy perspective, recombining data potentially enables re-identification of individuals from data that contains no specific identifiers or has been intentionally stripped of identifiers. Indeed, a research project may posit such re-identification as an explicit goal, for example in attempting to track persons with a infectious disease across time and space. The possibility of re-identification through recombining data raises questions about privacy protection distinct from information protection.

A Case: Recombining data to forecast forced migration

Investigators at Georgetown University are recombining multiple sources of big data about forced migration to better forecast, respond to, and help alleviate the consequences of humanitarian crises. Because detailed local data is difficult to obtain in a timely manner, this project explores the effectiveness of using open-source, online data to help identify indirect indicators of displacement/forced migration. Indicators relevant to this project include: economic, political, social, demographic and environmental changes affecting movements; intervening factors such as government refugee policies; and community and household characteristics. Parsing irrelevant information from the true indicators, calibrating results, understanding how these indicators change through time, and identifying and removing potential bias, requires large-scale data analysis and potentially, new computational methods for developing meaningful descriptive and predictive models. To date, the big data Georgetown uses for this study include open-source media articles and Twitter data. The investigators have access to EOS, a vast unstructured archive of over 700 million publicly available open-source media articles that has been actively compiling since 2006. New articles are added at the rate of approximately 300,000 per day by automated scraping of over 22,000 Internet sources in 46 languages across the globe. The project also collects data from Twitter—hundreds of thousands of tweets per day for the last 6 months. When relevant, the investigators also draw data from the scholarly literature of history, anthropology,

economics and other social sciences as well as the gray literature of governmental and non-governmental organizations. Long term plans include adding data from the archives of collaborating international and non-governmental organizations.

Ethical implications: This case highlights the importance of protecting the results of big data analyses with the explicit intent of better describing and aiding specific individuals or communities; that is, potentially re-identifying people and, thus, creating privacy breaches for the purpose of humanitarian aid.

2.4 *Reanalysis*

Big data archives have been assembled, particularly in public health and healthcare, with comparative or longitudinal purposes in mind. Although investigators may identify some specific objectives at the time of the archives' creation, they also expect and hope that new uses may emerge as scientific knowledge grows, lines of inquiry develop and techniques for extracting new information from collected data sources become more sophisticated.

A Case: Reanalyzing Newborn Screening Data

State mandated programs provide screening and data collection for 4 million newborns in the U.S. each year. After newborn screening is completed, the residual dried blood spots (RDBS) and data can be stored for quality assurance and research purposes depending on state practice and statutes. Currently, fourteen states store RDBS for research purposes. Storage of RDBS and data can range from a few months to the entire life of the program depending on the state. For example, California has stored the RDBS and data for newborns born in California for the past 52 years. Programs in Minnesota and Texas lost law suits alleging use of their RDBS for purposes not included in the original parental consent, including but not limited to research. Indiana is currently involved in an active lawsuit.

The State of Minnesota was sued by 21 families who alleged the program's collection, use, storage, and dissemination of RDBS and test results without written, parental consent violated the Minnesota State Genetic Information Act of 2006. Some of the dissemination of the RDBS and data associated with newborns were for research purposes. First begun in 2009, the lawsuit was initially dismissed in district court and the dismissal was upheld on appeal. However, on Nov. 16, 2011, the Minnesota Supreme Court ruled that the use of the blood spots and test results for anything other than the initial screening was not explicitly authorized in statute. The State of Minnesota settled the lawsuit and destroyed 1.1 million RDBS and their associated data prior to the November 16, 2011 ruling. The Minnesota legislature revised the statutory language explicitly to authorize short-term storage and the use of blood spots and test results for program operations, and to require written, informed consent for long-term storage and use of blood spots or test results.

Ethical implications: This case highlights the importance of the concept of ethical horizon and the conditions for building trust in subjects who understand that

their data may be used in ways not yet imagined. Parents provided information about their children with no knowledge of its potential use outside that context, including scientific research. Like the Havasupai case, when public health officials or scientists take actions beyond the terms of original consent for data collection, they jeopardize the public's trust in the institution of science. In contrast to the Havasupai, the case of newborn screening potentially affects a majority of families in the United States and a major resource for public health and genetic research for years to come.

3 Pluralism of Principles

The ethical concerns that are raised by the cases mentioned above deal for the most part with privacy concerns. Given the complexity of the cases, privacy is more than the proverbial “right to be left alone” as defined by judge Brandeis (*Olmstead vs. United States* 1928). In fact, although there is no doubt that privacy can be seen as a value, or ethical principle, it is hard to define its normative dimension specifically, compared to other values, and to show what exactly its practical implications are. In the following, the terms “value” and “principle” can be used interchangeably, although we will mostly use the latter in referring to privacy, as it seems more directly action-related than the term “value.” Still, ideas or states of affairs are “values” not in and for themselves but because they entail specific actions. The meaning of the term “value” can be defined as a quality that makes the corresponding actions desirable or obligatory. If privacy is called a “value,” then, certain rules should also be called up to regulate certain actions so that specific human qualities are protected.

We start from the assumption that privacy alone cannot be defined as value. Although it is possible to say: “You should not do x because it violates privacy,” such a statement is incomplete. One can always ask: “But why does privacy matter?” This means that privacy is embedded in a set of other values. There are five values, or principles that we have identified in our previous work as being associated with privacy. These principles also have clear and direct utility in designating practical guidance for protecting privacy in research contexts. They are: nonmaleficence; beneficence; justice; autonomy; and trust. Taken together, the five principles define our approach as a well-defined pluralism. We understand the principles in the following way.

Do no harm or non-maleficence: harm has to be widely conceived in the case of big data. While physical harm is not likely to occur, or only as a remote consequence, a violation of privacy can have measurable effects, for example in the case of a financial loss. Psychological distress also has to be counted as harm and is measurable to a certain degree. The well-known formula that is used in medical ethics—“*primum non nocere*” (“first, no harm”)—can be applied to the use of data, too. The question whether individuals are harmed or not has primacy over other concerns. If real harm occurs, other normative principles, such as autonomy and

trust, become secondary. In a sense, all normative questions concerning big data have to do with some sort of harm, but only impacts that can be seen as a painful experience of some sort should be subsumed under the category of harm. For example, a certain violation of autonomy can be unwanted and undesirable without though producing a directly measurable negative effect. (This does not mean that in such cases violations of autonomy bear no normative concern, it just means that ‘harm’ might not be the most appropriate category to articulate these concerns.) It has to be noted that a certain amount of harm can be seen as ethically permissible. In medical ethics, the category of minimal risk is used (Office for Human Research Protection 1993). It can also be applied to research projects in big data, provided that minimal risk can be defined accurately and responsibly in this field.

Do good or beneficence: beneficence can be defined as concern for the well-being of others. Compared to non-maleficence, beneficence can often be seen as supererogatory (see Beauchamp 2013). Research projects can be permissible even if they do not maximize the well-being of participants or the public, at least not directly, and researchers have no obligation to contribute to the increase of the well-being of others. On the other hand, they are almost always obliged to avoid harm. It has to be noted that beneficence is not necessarily identical, or reducible, to an interest in promoting economic benefits but can be used in a wider sense (even if cost-saving and similar outcomes can certainly be seen as beneficial). Beneficence can play a crucial role in trade-off analyses (see below). However, especially with regard to beneficence the pluralism of our approach comes to bear. Having intentions of doing good can only be used as a legitimizing principle of actions if the other normative principles mentioned here are equally addressed. As a rule of thumb, it seems appropriate to use the principle primarily in a critical way, by asking, for example, whether a certain use of data does indeed yield any genuine benefits to participants or the public. In general, beneficence has to be applied with some caution, as big data professionals sometimes tend to exaggerate the social benefits of their innovations. In such cases, the burden of proof lies with those who claim to act according to an interest in beneficence, not with the users or recipients of the data applications that they create.

Justice: the principle is concerned whenever opportunities, rights, and goods are to be distributed among the individuals or groups who have been targeted by big data analyses. Violations of justice are social disadvantages of all kind, or acts of discrimination (see Executive Office of the President 2014). For example, an analysis of big data can lead to the result that certain socio-economic groups can be treated differently compared to others because they are less likely to benefit from opportunities that are given to them. Some evidence exists that big data are used to sort out less lucrative social groups (Marwick 2014). However, discrimination can also occur when data are *not* used to examine the existing disadvantages of groups. In such cases, the principle of justice can be critically applied to the design of research projects, not just to the use of their results.

Autonomy: the principle can be applied both in a concrete and a general sense. In the concrete sense of the term, autonomy is a well-established principle in other fields of applied ethics, where it refers to the freedom and capability of

decision-making in individuals. The procedure of informed consent, which is both an ethical and a legal requirement in research projects, is meant to ensure that all relevant information about a given study is disclosed to the participants and the latter have been afforded the opportunity to deny or modify their participation (for recent discussions, see White 2013). One of the questions that is addressed in other parts of this book (Collmann, FitzGerald, Wu and Kupersmith) asks to what degree the tool of seeking informed consent is realistic in big data studies. In the general sense of the term, autonomy refers to the social, political and economic practices of individuals that allow them to realize freedom (see Cohen 2012). For example, citizens have to be given the opportunity to articulate their political opinion freely, while customers are to be given the opportunity to choose goods according to their liking and establish contracts of all kinds with other economic subjects. In this sense, autonomy is an overall quality of social practices that is not reducible to isolated acts of decision-making. Democratic societies establish such practices as open opportunities for citizens to exert and cultivate autonomy, for example through equality before the law and electoral procedures. These opportunities can be seen as desirable goal within a society. Although their violation does not necessarily cause any direct harm to individuals (see above), it might restrict the overall autonomy they have, or perceive to have, at their disposal. Big data has a potential to undermine practices of autonomy, for example through the possibility of tracking and profiling individual behavior at all times, or through the use of data for the prediction of individual decision-making.

Trust: this principle refers to the informal agreements that have to exist among the members of society in order to allow individuals to pursue their personal good. It also has to exist between individual members of society and their institutions. Trust enables individuals to engage in innovative social projects and take risks. It eliminates the burden of securing the appropriate conditions each time that an individual acts. Trust is closely linked to autonomy insofar as individuals have to have confidence that their autonomy is both realizable and protected. Trust, however, has also to exist in cases where there is an asymmetry between the members of society. For example, parents have to trust schools to treat the data of their children confidentially, while schools have a right, and perhaps an obligation, to collect all kinds of sensitive data for the purpose of education. Like beneficence, trust should only be used as a legitimizing principle if the other normative principles are also addressed. For example, trust often has to be balanced with autonomy, and such trade-offs have to be made in a transparent and revisable way. Again like beneficence, it seems best to use the principle primarily in a critical way, for example, in order to assess whether the use of personal data might erode trust in the long run, or whether there is a discrepancy between the trust that is requested from the participants of studies and the way trustworthiness is established by the respective institutions and research personnel.

This overview of the five principle documents the need for a well-defined pluralism. No single principle is sufficient to address all concerns that are raised by the use of big data. At the same time, the principles are not introduced arbitrarily but complement each other by addressing aspects that each principle, taken in an isolated

manner, has to leave out. But the pluralism also allows us to show the importance of trade-off analyses, as there can be cases in which the principles do not so much complement each other but compete. When this happens, the relevance of the respective normative points of view has to be addressed, which then makes it necessary to address the threshold conditions that exist for each principle's point of view. In the following, we will say more about this point. It can also be noted that similar attempts in the data science community to establish ethical principles have led to a comparable set of principles. See the Menlo Report (Dittrich and Kenneally 2012).

The pluralism of principles allows us to address in a nuanced way the practices that are involved in the use of big data. Non-maleficence, beneficence, and justice have in common that they can be seen as action-related principles. Each concerns the legitimacy of the actions that are performed, or of the consequences that directly follow actions. On the contrary, autonomy and trust are agent-related. They concern the attitudes and perceptions of individual agents insofar as they are dispositions for an infinite variety of actions. Action-related principles concern individuals as targets, or objects, of actions. Individuals are referred to insofar as they might be harmed, benefited, or disadvantaged. Agent-related principles, in turn, concern individuals as spontaneous agents, or subjects, by asking what is necessary for them in order to maintain their agency. In simpler terms, action-related principles are concerned with the protection of subjects, while agent-related principles aim at empowering them.

For a full assessment of the privacy concerns raised by big data, it is necessary to keep this distinction in mind. The individuals that are targeted by big data uses are regular members of society who pursue active practices on the various levels of social life. It seems easy to confuse data related to individual agents with mere data points and to assume that individuals are helped if only the flow of data is optimized. The idea that one can optimize human life by optimizing data technologies has rightfully been called "solutionism" (Morozov 2014). But insofar as big data analyses can interfere with the realm of individual agency and potentially pervade all aspects of social life, one has to consider all aspects that are relevant for the full realization of democratic practices. This makes it necessary to establish feedback loops that go beyond the realm of "pure" data analysis and involve the real-world concerns and needs of individual agents.

In more practical terms, the distinction between individuals as targets and as agents has been addressed by the difference between the "restricted access theory" and the "control theory" (see the overview in Tavani 2008). For these theories, privacy is realized either by restricting the access to personal data or by giving data sources control over the data they want to share. Both theories can be used legitimately, as individuals and groups both have to be protected passively and need to be involved actively in the data collection in case the possibility for such an involvement exists.

At the same time, the distinction between the principles outlined here allows us to distinguish immediate from long-term outcomes. For example, if predictive techniques can be implemented effectively they are likely to change the fundamental conditions of individual agency. In such cases, no immediately measurable outcomes have to be identified in order to raise concerns. A change in the basic conditions of social practices may call for ethical reflection even if no immediate harm can be identified at present. In addition, big data can concern both individuals

and communities. While harm, seen as measurable impact, is more likely to be identified in individuals, the interests of communities can also be addressed using the principle of justice.

4 The Variety of Contexts: The “Privacy Matrix”

Besides the variety of normative principles, it is pertinent to consider the practical context in which privacy matters. Given the fact that big data is generated in a broad variety of human affairs, sometimes in an automatic way and under some assumptions of publicity (e.g., geo-tagged photos posted on the web, or social media links and posts announcing political attitudes), while at other times with the expectation that the information will be strictly guarded (genetic profiling), it is imperative to consider the impact of context (social, scientific, economic, political) on the privacy regime of each situation.

A very similar approach has already been suggested by Nissenbaum (2011). According to her, “we must articulate a backdrop of context-specific substantive norms that constrain what information websites can collect, with whom they can share it, and under what conditions it can be shared” (32). In developing her approach, Nissenbaum warns that we need to take into account all possible contexts, not just commercial ones. Only by using a contextual approach can we navigate the complex decision-making landscape of privacy. Heuristically, she advises that we “locate contexts, explicate entrenched, informational norms, identify disruptive flows, and evaluate these flows against norms based on general ethical and political principles as well” (38).

Given the dynamic nature of big data, as it is evident in the uses described in Sect. 2, the relation to context seems to become ever more relevant. If big data are used according to their potential, they are very likely to switch contexts, for example between research and commercial applications, or between commercial applications and the government.

To simplify this discussion we have proposed elsewhere a “privacy matrix,” which helps matching privacy dimensions in given contexts that involve interactions between data collection agencies, scholars, commercial agents, political actors, and ordinary individuals (see Steinmann et al. 2015).

The privacy matrix				
Specifying principles	Privacy contexts			
	Social	Government	Commerce	Science
Nonmaleficence				
Beneficence				
Justice				
Autonomy				
Trust				

The matrix is structured in two dimensions: contexts and normative concerns. The normative concerns are based on the five principles mentioned above: non-maleficence, beneficence, justice, autonomy, and trust. The contexts are broadly defined, as social, government, science, and commercial. “Social” refers to open, public domains outside of business entities, research institution, or government bodies. While it is unlikely that genuine research of big data can be conducted in this realm, data still occur and research tools might get used, in one way or the other. The main idea of this privacy matrix is that the same ethical concerns can become more or less sensitive or more or less tractable according to each context, since in each context the amount of disclosure, the nature of disclosure, and the ultimate effects of disclosure vary in gravity. Also, the legal and moral expectations, especially as encapsulated in norms and legislations, allow transactions to be more or less permissive when it comes to privacy.

The matrix is to be used as a heuristic tool. Given a certain situation or research project that involves big data in one of the columns (that is, contexts), one walks down the value or concern list, considering at each step the specific nature of normative implications for the given context. Naturally, it is desirable that all ethical concerns should be considered important, since we cannot make capricious decisions as to what matters or not. Yet, not all principles are equally applicable. Ultimately, the heuristic process demands that we zoom in on the most relevant normative concern for a given context and try to answer the question: how will privacy be protected in such a way that the relevant concerns are met, while the data is still usable?

The matrix does not presuppose that data belong essentially only to one sector. By its nature, data cross borders and can be linked in a multitude of ways. As already said, data can “travel” from closed settings in, say, educational institutions to commercial institutions and the government. For example, data generated from individual practices, such as movement in space, can become relevant in commercial and governmental settings. Big data is protean in nature insofar as use and impact cannot be stated definitively by focusing only on one specific context or one specific way in which data are presently used. On the other hand, while data are likely to “travel” in these ways, their respective use is still always relevant in one specific context, for specific users and their purposes. The multi-contextuality of big data does not dispense the ethical reflection from considering each context specifically.

The rubrics in the matrix are thus not separate universes that endow data with an essence that prevents them from being used in other rubrics. They rather have to be seen as areas of use and relevance. The basic meaning of the matrix is in this sense dynamic, not static. The underlying normative idea is to determine specifically who controls the access and exploitation of data in each case. If data, for example, are used and stored within a governmental context, the task is to determine in which way the respective agencies can use the potentials of data sets. And, if data “travel” across the borders of contexts, how much of the control previously established will be lost, and which new purposes will be added?

Taking up Nissenbaum's (2011) use of the term "appropriateness," we can say that the purpose of the matrix is to define the appropriate concerns in each case of the use of data. On the side of subjects, or data sources, it is important to consider their rights and legitimate expectations. On the side of users, the legal and institutional responsibilities have to be established. In addition, all these considerations have to be qualified as revisable given the protean nature of big data.

It is worth noting that the contexts might also relate to different strategies of inclusion and exclusion. While there is a tendency to understand privacy predominantly as a way of restricting access to personal data, harm can also result from not considering some private data. If big data research leaves out, say, vulnerabilities due to race or social status, then the protection of privacy, paradoxically as it seems, becomes harmful. For this reason, it helps keeping both principles of non-maleficence and beneficence in mind and ask for each case: have we done as little harm as possible, and have we done as much good as we can?

5 Trade-off Analysis and Threshold Conditions

The matrix can also be employed in a heuristic methodology that uses a modified version of trade-off analysis, which may overcome some ethical difficulties typically implied by the utilitarian logic of trade-offs, such as those mentioned by Kelman (1981) and discussed in detail in the literature (Palm and Hansson 2006; Elgesem 2002).

We suggest a model that starts with determining a minimum "concern threshold" for each dimension. In other words, our model starts with the assumption that trade-off analysis needs to take into account that the results of a cost-benefits analysis cannot lead to reducing any of the dimensions (autonomy, trust, etc.) beyond a minimum acceptable value (threshold), which in no case can be 0. The "normal" minimum thresholds are to be determined as much as possible in absolute terms, regardless of context. Ideally, there should be a minimum of universally applicable level of beneficence, harm avoidance, trust-protective behavior, or autonomy-defensive procedures for all contexts. If this cannot be determined as an absolute value, it should, at the very least, be determined within a narrow band of variation.

For example, for all big data collection that involves harvesting information from social media, autonomy should be protected across contexts in such a way that in none of them the fundamental right to free expression is reduced to zero. In other words, in no context should the data collected be utilized in a manner that diminishes the right of the individuals observed to decide on what to believe or say or that leads to retaliatory measures against them by state or non-state actors. Of course, above this threshold some projects can disclose more and some less about what was said in what context by what type of user, according to the nature of the data collection process. If the data was collected, for example, from a large, government-sponsored organization using social media (e.g., a health peer-support

forum for former military personnel suffering of PTSD), anonymization measures need to be strict even if the communication was made in public, under a user's own name. On the other hand, if data was collected from a publicly available site, say, Twitter or Wikipedia, some information about the users can be used in the project, since such material is comparable, in terms of publication privacy, to letters to the editor or other publicly made statements. Yet, again, some publicly available sites or platforms, are public only in that anyone can sign up or apply to become a member. If upon joining the sites or platforms researchers enter spaces defined as private or "closed" the natural expectation of the members that the information is to a certain extent private should not be violated. Furthermore, any participation should be accompanied by appropriated disclaimers, announcing that researchers act in a research context and their goal is to collect data for research purposes only.

To make things more complicated, social media data collection can at times be automated. Specialized software can be instructed to "spider-walk" and harvest information from a variety of social media groups through tools and procedures provided by the platform and site administrators. Such tools, typically called APIs (Application Programming Interfaces) reveal and make available text, images, likes, or comments posted by site users, including those acting behind the "firewall" of "private groups." These applications should not be hidden behind the veil of technological automation. The control that the members imposed on access demands higher level of privacy protection, which demands informed consent from all members concerned.

Furthermore, trade-off analysis needs to include transparency as a core procedural assumption. In other words, the terms of the elements that are traded against each other, their measure and significance, and the exact cost incurred for each benefit should be visible and actively propagated. In addition, transparency needs to be prospective, not retrospective. The terms of the analysis and the presumptive results need to be announced before, not after, the trade-off process is completed.

Finally, transparency should aim at generating community participation in the trade-off process (Milne 2000; Hann et al. 2002). Transparency without active input from all sides is deceiving and in no way conducive to ethical behavior, quite to the contrary. Participation can be achieved in various ways and presents specific challenges within each community. These challenges can partly be practical, because community engagement can be difficult and cost-intensive, and partly arise from the context of data use. Commercial entities, for example, have a certain right to keep their practices secret if only to allow them to stay competitive, while in research it is often not possible to disclose all purposes of a study to its participants without distorting the results. Still, it is important to uphold the obligation that data research has to be transparent, or at least as transparent as possible given all legitimate considerations, to the subjects who provide the data in the first place. Feedback loops have to be created that involve the data sources, that is, individual agents, with the opportunity to exert their active agency.

Transparency is not the only moderating factor in determining the ethical impact of big data collection and analysis. Another factor that needs to be taken into account is specificity with respect to context. Context, as explained above, is to be

regulated in each case by a specific procedural approach that emphasizes the particular expectations of privacy of data providers (individuals) and data collectors (e.g., researchers). For any ethically responsible trade-off analysis, maximum attention has to be paid to these requirements. In a commercial context, for example, transparency requires the user to be informed fully and in detail how data is being collected, what is done or will be done with it, and whether any possible sunset policies exist. Methods of redress and opt-out need to be offered. However, since the transaction is conducted on a commercial basis, involving monetary exchanges or fiduciary interests, which require tracking down payments and material interests, as well as adjudication of ownership of content or other type of intellectual property, perfect anonymity cannot be enforced. Also, disclosure to third parties, such as governmental agencies, in cases involving criminal acts (drug dealing or sexual exploitation on an open social media site, material support to terrorist organizations through fundraising, etc., infringement of intellectual property, abuse or violence) should be considered legitimate types of disclosure. In a research context, on the other hand, transparency should include more than informing the user as to the methods to be used to protect identity but also reassure her that her identity will not be disclosed even in situations that are currently considered within the purview of criminal law. Researchers that operate under terms of use and privacy statements regarding data collection that emphasize the absolute anonymity of the respondents should make sure that they use all necessary means to guarantee it and to inform the users on the measures they will use to do so, even in situations that would typically force data holders to release personally identifiable information. Furthermore, researchers need to pro-actively enforce procedures of anonymization by data aggregation or reduction of data granularity that avoids disclosure of private data through re-analysis and recombination.

6 Conclusions

The elements that are mentioned in this article—the well-defined pluralism of normative principles, the matrix listing the various privacy contexts, and the challenges for any trade-off analysis to consider minimum threshold conditions—exemplify the specific challenges that arise through the new methods of big data analysis. Ethical reflection needs to be adaptive to the evolving nature of big data. At the same time, it has to develop conceptual tools that can be fine-tuned to the various cases that can arise. Some concerns mentioned in this article are still hypothetical and will perhaps need to be changed once the results, failures or successes, of big analysis become evident in the future. On the other hand, while this means that the ethical reflection is to certain parts also still hypothetical, it affords at least the possibility of engaging the community of researchers and data professionals from the onset of new developments.

References

- Beauchamp, T. (2013). The Principle of beneficence in applied ethics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/win2013/entries/principle-beneficence>, Accessed 13 November 2015.
- Cohen, J. (2012). Configuring the networked citizen. In A. Sarat, L. Douglas, & M. M. Umphrey (Eds.), *Imagining new legalities: Privacy and its possibilities in the 21st Century* (pp. 129–153). Stanford: Stanford University Press.
- Dittrich, D., & Kenneally, E. (2012). *The Menlo report: Ethical principles guiding information and communication technology research*. US Department of Homeland Security. <http://www.dhs.gov/sites/default/files/publications/CSD-MenloPrinciplesCORE-20120803.pdf>, Accessed 13 Nov 2015.
- Elgesem, D. (2002). What is special about the ethical issues in online research? *Ethics and Information Technology*, 4(3), 195–203.
- Executive Office of the President. (2014). *Big data: Seizing opportunities preserving values*. http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf, Accessed 13 Nov 2015.
- Hann, I.-H., Hui, K.-L., Lee, T., & Png, I. (2002). Online information privacy: Measuring the cost-benefit trade-off. *ICIS 2002 Proceedings*, 1.
- Harmon, A. (2010). Indian tribe wins fight to limit research of its DNA. *The New York Times*. Retrieved from <http://www.nytimes.com/2010/04/22/us/22dna.html>
- Jacobs, B., Roffenbender, J., Collmann, J. Cherry, K., LeManuel, L., & Bassett, K., et al. (2010). Bridging the gap between genomic scientists and indigenous peoples. *Journal of Law, Medicine and Ethics*, 38(3), 684–696. http://arep.med.harvard.edu/pdf/Jacobs-JLME_10.pdf
- Kelman, S. (1981). Cost-benefit analysis: An ethical critique. *Regulation*, 5, 33.
- Marwick, A. (2014). How your data are being deeply mined. *New York Review of Books*. January 9. <http://www.nybooks.com/articles/archives/2014/jan/09/how-your-data-are-being-deeply-mined/?pagination=false>, Accessed 13 Nov 2015.
- Milne, G. R. (2000). Privacy and ethical issues in database/interactive marketing and public policy: A research framework and overview of the special issue. *Journal of Public Policy & Marketing*, 19(1), 1–6.
- Morozov, E. (2014). *To save everything, click here: The folly of technological solutionism*. New York: Public Affairs.
- Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus—The Journal of the American Academy of Arts & Sciences*, (Fall), 32–48.
- Office for Human Research Protection. (1993). *Protecting human research subjects: Institutional review board guidebook*. Washington, DC, U.S. Government Printing Office. http://www.hhs.gov/ohrp/archive/irb/irb_guidebook.htm, Accessed 13 November 2015.
- Olmstead v. United States. (1928). 277 U.S. 438, 48 S. Ct. 564, 72 L. Ed. 944.
- Palm, E., & Hansson, S. O. (2006). The case for ethical technology assessment (eTA). *Technological Forecasting and Social Change*, 73(5), 543–558.
- Steinmann, M., Shuster, J., Collmann, J., Matei, S. A., Tractenberg, R. E., & FitzGerald, K., et al. (2015). Embedding privacy and ethical values in big data technology. In S. A. Matei, M. G. Russell, & E. Bertino (Eds.), *Transparency in Social Media* (pp. 277–301). Springer International Publishing. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-18552-1_15
- Tavani, H. (2008). Informational privacy: Concepts, theories, and controversies. In K. E. Himma & H. Tavani (Eds.), *The Handbook of Information and Computer Ethics* (pp. 131–164). Hoboken: Wiley.
- White, L. (2013). Understanding the relationship between autonomy and informed consent: A response to Taylor. *The Journal of Value Inquiry*, 47(4), 483–491.

Ethical Reasoning in Big Data

An Exploratory Analysis

Collmann, J.; Matei, S.A. (Eds.)

2016, X, 192 p. 14 illus. in color., Hardcover

ISBN: 978-3-319-28420-0