Technology in Society 49 (2017) 31–36

Contents lists available at ScienceDirect

Technology in Society

journal homepage: www.elsevier.com/locate/techsoc

# Ethics & Big Data

Richard Herschel*, Virginia M. Miori

*Decision & System Sciences Department, Ervian K. Haub School of Business, Saint Joseph's University, Philadelphia, 19131, USA*

ARTICLE INFO

*Article history:*
Received 24 August 2016
Received in revised form
3 March 2017
Accepted 7 March 2017
Available online 10 March 2017

*Keywords:*
Ethics
Big Data
Kantianism
Utilitarianism
Social Contract Theory
Virtue Theory

ABSTRACT

Big Data is a digital phenomenon that enables the collection and use of massive amounts of data derived from both man and machine. This data is characterized in terms of its volume, variety, velocity, veracity, variability, and its complexity. While Big Data allows firms to rapidly capture, analyze, and exploit information, it can also enable access to data that compromises an individual's privacy. And this can happen either deliberately or inadvertently. Either way, Big Data fosters a discussion of ethical issues relative to the sharing and usage of data. Ethical debates are typically articulated within the context of ethical theories. These theories help to frame our understanding of moral issues. Their use affords insight into the context and the logic of the moral arguments being presented, thereby providing us with a rational mechanism by which to better evaluate whether an intended action or actual outcome is morally right or wrong. Four ethical theories are briefly reviewed in this paper: Kantianism, Utilitarianism, Social Contract Theory, and Virtue Theory. Each theory is than examined to show how it might be employed to examine Big Data issues.



## 1. Introduction

Big Data is a phenomenon that is fundamentally changing what we know and do. Big Data is all about capturing, storing, sharing, evaluating, and acting upon information that humans and devices create and distribute using computer-based technologies and networks. Data comes from a multitude of sources, including sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, RFID devices, and cell phone GPS signals to name a few. Today we are now generating 2.5 quintillion bytes of data — so much that 90% of the data in the world today has been created in the last two years alone [1]. Collectively this plethora of data is called Big Data.

Big Data is described by IBM in terms of four dimensions: volume, variety, velocity, and veracity. Where once organizational computer applications were the primary source of data generation, new social, personal, and device-to-device digital communications have accelerated data volume growth exponentially. The volume of data is now projected to be 40 zettabytes by 2020, an increase of 300 times from 2005. In addition, the variety of data is also increasing due to new sources and forms of data creation. Video sharing, social media, location services and other innovative forms of data generation and exchange means that digital media has become increasingly more data intensive and media rich. Improvements to the telecommunications infrastructure combined with the rapid deployment of high-speed wireless technologies worldwide have enabled greater bandwidth for transferring this data as well as the ability to share it globally. As a result, the velocity of data streaming has become so intensive that in 2016, it projected that there will be 18.9 billion network connections – almost 2.5. connections per person on earth. However, it must also be noted that the increase in the volume, variety, and velocity of data is indiscriminate relative to the quality of data being captured and exchanged. Hence Big Data is agitated by its veracity, because oftentimes the data being collected and distributed is incomplete and/or inaccurate [2].

SAS has identified two additional dimensions of Big Data – variability and complexity. Big Data variability is evidenced by the fact that data flows can be highly inconsistent with periodic peaks. Complexity is manifested in the nature of the data itself. It is both structured and unstructured and coming from multiple sources, which make the data difficult to link, match, cleanse, and transform across systems [3].

Big Data provides new opportunities for companies enhance their performance. The Big Data market has a strong momentum as businesses accelerate their transformation into data-driven companies. This momentum is driving strong growth in Big Data-related infrastructure, software, and services. A new forecast from

* Corresponding author.
*E-mail address:* herschel@sju.edu (R. Herschel).

http://dx.doi.org/10.1016/j.techsoc.2017.03.003
0160-791X/

International Data Corporation (IDC) sees the Big Data technology and services market growing at a compound annual growth rate (CAGR) of 23.1% over the 2014–2019 forecast period with annual spending reaching $48.6 billion in 2019 [4].

When combined with analytics and data mining, Big Data provides new opportunities for understanding and predicting consumer behavior … and more. Firms are using Big Data to enhance their relationships with existing customers and to exploit opportunities to attract new customers. In addition, Big Data is being analyzed to better manage supply chains, health care, to monitor equipment and facilities, and to create new products and services or to enhance existing ones.

However, this relatively new ability to capture, share, analyze, and act upon a wealth of new data is not without potential risk for firms and their customers. As noted above, oftentimes Big Data is difficult to manage and it is often incomplete or even inaccurate. Yet it is also rich and easily and continuously available in huge volumes for analysis. Because the nature of Big Data is so indiscriminate, firms may be privy to information that they never intentionally intended to collect. In other words, Big Data may incorporate information that infringes upon people's privacy. Because of this, Gartner asserts that when Big Data is subjected to sophisticated advanced analytics capabilities, new risks become inherently associated with Big Data. In fact, they predict that by 2018, 50% of business ethics violations will occur through improper use of Big Data analytics [5].

Herold [6] has identified a number of important privacy risk associated with Big Data and Big Data analytics. For example, with so much data and with powerful analytics, it may be impossible to remove the ability to identity individuals, if there are no rules established for the use of anonymized data files. For example, if one anonymized data set was combined with another completely separate data base, without first determining if any other data items should be removed prior to combining to protect anonymity, it is possible that individuals could be re-identified. The important and necessary key step she says that is usually missing is establishing the rules and policies for how anonymized data files can be combined and used together. She also notes that if data masking is not used appropriately, Big Data analysis could easily reveal the actual individuals whose data has been masked. Herold warns that Big Data can be used to try to influence and drive behaviors. What is implied here is that Big Data can be used by organizations to make a much wider variety of decisions that do not take into account the privacy of the individuals whose personal data is being exploited. This problem is further aggravated by the fact that that Big Data can be used to fill in gaps in information about individuals. This can occur because the collection of Big Data from online transactions and the Internet of Things oftentimes affords firms the opportunity to expand their knowledge about an individual without their knowledge or consent. In effect this means that decision making is oftentimes being transferred away from individual decisions that have knowable outcomes and replaced by actions derived from Big Data analytics which may have unintended consequences for many.

Zoldan [7] argues that many times the Big Data being utilized for decision making is not always correct as it is oftentimes incomplete, biased and/or missing context. Despite this, organizations frequently have a false sense of confidence in the data, since there is so much Big Data available. This is especially problematic if Big Data algorithms are drawing inaccurate conclusions about customer identities and behavior based on flawed data. Other potential problems associated with Big Data analysis are signal error and confirmation bias. Signal error occurs when large gaps of data have been overlooked. Confirmation bias is the phenomena that data is selectively used to confirm a preexisting viewpoint, while disregarding the data that refutes it. The point being made by Zoldan is that the use of Big Data necessarily requires skepticism and caution to avoid statistical false positives and incorrect findings that may lead to bad decisions and unintended risk for both the organizations and its customers.

Zwitter [8] argues that Big Data has the effect of shifting the focus of ethics away from the individual's ability to make moral judgements on some notion of right or wrong as well as their accountability. Instead, Big Data requires an examination of those that have control over Big Data, because Big Data can be used to target and manipulate people to consume or behave in a certain way. Big Data stakeholders wield a significant amount of power because they control the collection and the utility of Big Data, employing data derived knowingly or unknowingly from others. Big Data can have the effect of reducing knowable outcomes of actions, while increasing unintended consequences. Therefore, Zwitter contends that Big Data fundamentally changes the nature of ethical debates by redefining what power is and where it lies and the extent to which free will in fact guides one's actions.

The research approach taken in this paper is taxonomy based. The authors have examined a broad range of ethical approaches, building a complete taxonomy. From this collection, ethical approaches that most closely represented the current societal ethics were chosen for further examination. The complete taxonomy and ethical implications of other ethical approaches within the taxonomy are available upon request.

## 2. Big Data and ethical issues

Richards and King [9] note that large datasets are being mined for important predictions that often yield surprising insights. They assert that because of Big Data and the analytics used to examine it, all kinds of human activities and decisions are beginning to be influenced by Big Data predictions, including dating, shopping, medicine, education, voting, law enforcement, terrorism prevention, and cybersecurity. Yet while this is occurring, individuals have little idea concerning what data is being collected, let alone shared with third parties. Hence, Richards and King assert that existing privacy protections focused on managing personally identifying information are not enough when secondary uses of Big Data sets can reverse engineer past, present, and even future breaches of privacy, confidentiality, and identity. They note that Big Data efforts find many of the most revealing personal data sets such as call history, location history, social network connections, search history, purchase history, and facial recognition and much of this this information is are already in the hands of governments and corporations. And the collection of these and other data sets is only accelerating. Richards and King conclude that Big Data is producing increased institutional awareness and power that requires the development of Big Data ethics to protect individual rights.

Culnan and Williams [10] have noted the potential for abuses of informational reuse and unauthorized data that could result in privacy problems. They state that information reuse involves organizations making legal decisions about new uses for the personal information they have collected, while unauthorized access involves employees viewing personal information they are not authorized to view. Both activities, information reuse and unauthorized access, can potentially threaten an individual's ability to maintain a condition of limited access to his/her personal information, harm individuals, and subsequently threaten the organization's legitimacy in its interactions with consumers, shareholders, and regulators. Privacy harms resulting from unauthorized access can include a breach of confidentiality and trust, or the financial harm to individuals from identity theft or identity fraud. Unfortunately, Big Data has only enhanced the potential for such issues.

When examining the implications of Big Data on market research, Nunan and Di Domenico [11] identified privacy challenges created by the use of Big Data. The first issue they documented arises from different sets of data that would not previously have been considered as having privacy implications concerns being combined in ways that then threaten privacy. They call this the unintended use paradox. One example cited is the discovery by researchers who used publicly available information and photographs from Facebook and, through application of facial recognition software, matched this information to identify previously anonymous individuals on a major dating site. In another example, anonymous 'de-identified' health information distributed between US health providers was found to be traceable back to individuals when modern analytical tools were applied. With Big Data comes the possibility of significantly changing the relationship that individuals have with the data collected about them. Moreover, because Big Data and data mining findings are derived using correlations among data, there is a higher likelihood for finding random connectedness on the basis of random commonalities. The result of this is that Big Data analyses may yield information that not only compromises privacy, but suggests random connections based on incidental occurrences.

The second privacy challenge Nunan and Di Domenico ascertained revolves around the fact that data is increasingly being collected autonomously, independent of human activity. The authors note that with the emergence of network-enabled sensors on everything from electricity and water supplies to airplanes, the volume of data created by these devices, and the speed with which the data must be analyzed means that data collected is automatically analyzed without any consideration for individual consent.

Mandinach, Parton, Gummer, and Anderson [12] state that the ethical use of data involves knowing how to use data and how to protect privacy and maintain the confidentiality of data. Such knowledge includes how to remove identifying information from a data record and knowing who has access to data and when and how, and the process by which to release data or results. Unfortunately, oftentimes the processing of Big Data is automated, being processed by devices that using analytic algorithms that are insensitive to these issues. And even when humans are involved in the process, oftentimes the sheer volume of Big Data make such efforts impractical.

Governments have tried to mitigate the potential privacy issues inherent in the collection and use of Big Data. Nicolas Terry [13] notes that recently the European Court of Justice asserted a right of erasure that requires the data controller to take all reasonable steps to have individuals' data erased, including data provided by third parties without delay, for the personal data that the controller has made public without legal justification. In contrast, Terry finds the current views of the US government on Big Data regulation and how to deal with the threat to health privacy to be either incoherent or, at best, coalescing around inadequate downstream data protection models such as transparency and 'use point' regulation. He supports the US Federal Trade Commission [FTC] recommendation that affirmative [opt-in] express consent should precede the collection and sharing of information with data brokers.

King and Richards [14] recommend that barring governmental intervention, organizations should at least engage in conversations about Big Data ethics. For example, they argue that firms should define and enforce rules about data use and retention. In accord with the FTC recommendation, King and Richard believe people should have the ability to manage the flow of their private information across massive, third-party analytical systems. As already noted, Big Data is powerful because secondary uses of data sets produce new predictions and inferences. This leads to data being a business, with people such as data brokers, collecting massive amounts of data about consumers, often without their knowledge or consent, and shared in ways that people don't want or expect. For Big Data to work in ethical terms, the data owners (the people whose data is being used) should at a minimum be provided with a transparent view of how their data is being used – or sold. That said, the authors also note that it is not realistic to think of all information as being either secret or shared, completely public or completely private. Oftentimes data is shared or generated by design by trusted services for important services for consumers. Their point is that in these circumstances, there is no implied consent that the data can be used for any other purpose.

## 3. Ethical perspectives

A discussion of ethics and Big Data is dependent upon how one defines ethics. In general, ethics involves the analysis of conduct that can cause benefit or harm to other people. However, ethics is a topic that has been studied for at least 2400 years and in that time there have been a number of formulations of ethical principles.

Sound ethical theories share a common property. They enable the individual to make persuasive, logical, and reasoned arguments based on the principles stated by the ethical theory. To illustrate this, four ethical theories will be briefly examined: Kantianism, Utilitarianism, Social Contract Theory, and Virtue Ethics.

### 3.1. Kantianism

Kantian ethics, originating with the German philosopher Immanuel Kant (1724–1804), is an ethical theory concerned about not about what we do, but what we ought to do. What we ought to do reflects our dutifulness. Dutifulness reflects good will – the desire to do things right based upon rules that everyone ought to follow. That is, a dutiful person acts the way they do because of a morale rule. These rules are imperatives that are either hypothetical or categorical and they are the means by which reason commands our will and our actions. Hypothetical imperatives equate basically to conditional if, then statements relative to what you are trying to accomplish. Categorical imperatives command unconditionally as they are unequivocal. For example, one categorical imperative states that you should only follow moral rules that you would expect everyone else to follow. Another states that you should never treat people as means to an end, but as an end unto themselves since all persons have moral worth and should be treated with dignity. For Kant, rules are paramount. Everyone is held to the same standard and there are clear guidelines for appropriate behavior. Hence, in Kantianism it is not the outcome of a behavior that matters, rather it is the rule behind the action that is most critical [15].

### 3.2. Utilitarianism

Unlike Kantianism, Utilitarianism [originating from Jeremy Bentham (1748–1832) and John Stuart Mill (1806–1873)] examines right or wrong based on the consequences of an act or a rule. The act utilitarian perspective applies the principle of utility to individual moral actions and the rule utilitarian applies the principle of utility to moral rules. The right act is one that produces the greatest happiness for a community or society. A wrong act decreases the total happiness of the affected parties. The right moral rule of conduct is one where if it is adopted by everyone, will lead to the greatest net increase in happiness for all involved. Hence, in the utilitarianism ethical perspective, one must calculate what action or rule achieves the best results. That is, one must literally account for and the weigh the good and the bad elements affecting a situation to determine the net consequences of the action or rule.

Hence, unlike the Kantian perspective where the focus is upon examining the will that motivates an action, in Utilitarianism it is the "happiness" or the maximum well-being outcome that is most critical [16].

### 3.3. Social Contract Theory

Social Contract Theory [based on the arguments made by philosophers Thomas Hobbs (1588–1679), John Locke (1632–1704), Jean-Jacques Rousseau (1712–1778), and John Rawls (1921–2002)] is an ethical perspective that states that a person's moral and/or political obligations are dependent upon a contract or agreement that people have made to form the society in which they live. In this theory people are seen as rationale beings who understand that in order to create and maintain a society, people must cooperate and agree to follow certain guidelines in order to gain the benefits of social living. To do this, that people must choose rationality over their natural selfish instincts. That is, they must be willing to submit to a government and its laws in order to live in a civil society, rather than live in a "natural' state of anarchy and chaos. The social contract provides the justification for the establishment of moral rules to govern relationships among citizens as well as the mechanism capable of enforcing these rules – government [17].

### 3.4. Virtue Ethics

Virtue ethics emphasizes virtues, or moral character, rather than duties, rules, or the consequences of actions. Rooted in the arguments of Aristotle (384 BCE–322 BCE) and Plato (428 BCE–347 BCE), this theory defines a virtue as a character trait or disposition that is well entrenched in its possessor which makes that person good. There are two types of virtues, intellectual and moral virtues. Intellectual virtues are those derived from reasoning and truth. Moral virtues are deep-seated habits or dispositions formed through the repetition of virtuous actions over time. Morally good people realize happiness by consistently acting out their virtues, doing what any virtuous person would know to be right. For example, honesty, justice, generosity, and loyalty may be seen to be core virtues [18].

## 4. Applying ethical theories to Big Data issues

These ethical perspectives are useful for understanding how ethics informs Big Data-related issues. Frequently, there are articles in the press that discuss ethical concerns, though the underpinnings of the ethical viewpoint are left unclear. However, by employing the ethical perspectives described above, it is possible to better understand how and why ethics helps to inform an issue such as Big Data privacy concerns.

### 4.1. Kantianism and Big Data

Kantian analysis argues that one should always respect the autonomy of other people, treating them as ends in themselves and never only as means to an end. With Big Data, this would be a difficult case to make. Since data is routinely collected and analyzed to assess individuals without their consent, organizations employing Big Data are not respecting the autonomy of people and they are in fact using personal data as a means to an end to further the organization's self-interest. The nature of Big Data is that in general, people typically do not opt-in to their data collection and exploitation, demonstrating their consent and hence shared responsibility. This means that by default, their privacy is compromised for the gain of another.

Organizations utilizing Big Data may argue that they post information online informing consumers that the data they capture from a user's online behavior patterns may be used to offer new products or services. Some even provide their customers the option to opt-out of the firm's ability to share their information with organization's business partners. The fact is, however, that practically speaking, no one has the ability to determine how their data is actually shared and used, because the Big Data space is too big and there is no mechanism affording the individual the ability to actively monitor and control their private information. Hence to a great extent, individuals are blind to the sharing of their digitized data. While individuals may employ digital services to warn them of identity theft, to monitor credit issues, or to inform them when they are mentioned in postings, they are typically forced to be reactive rather than proactive posture in responding to information that may affect their privacy and secutity.

Big Data compromises the old adage that state "Treat people how you want to be treated". This phrase speaks to the Kantian notion that one should act only on the moral rules that you can imagine everyone else following. However, with Big Data individuals are frequently represented simply as data points that are then used to manipulate what the person will view in the future. That is, information is presented to individuals online that Big Data calculations determine best reflects their projected preferences based upon their previous search and online page view history. This algorithmic manipulation presumes the will of the individual without their explicit consent.

Using a Kantian viewpoint, one might ask whether everyone should assent to a rule that states that everyone's information can be shared with or without their permission, regardless whether it is accurate or inaccurate, complete or incomplete, current or dated, and this information can be used to influence and represent peoples' behavior and interests with or without their consent. This is probably unlikely, otherwise there would not be so many concerns expressed about Big Data and privacy rights and protection.

The point here is that Kantianism provides a relatively straightforward means for discussing the ethics of Big Data. It asserts that all people are rational, autonomous beings having moral worth and everyone is held to the same universal moral guideless. Because of this, Big Data is problematic for Kantian belief's because the actions associated with Big Data challenge the rights and fair treatment of the individual.

### 4.2. Utilitarianism and Big Data

Unlike Kantianism, assessing the ethics of Big Data from a Utilitarian perspective is fraught with complications. It requires that acts and rules be assessed using a utilitarian calculus where the god and bad of Big Data are weighed on a scale. From an Act Utilitarian perspective, for example, one would have to quantify the plusses and minuses of Big Data consequences relative to such factors as the intensity of the experience, its duration, the probability that something would occur, how close the experiences are in space and time, its ability to produce more experiences of the same kind, the extent to which pleasure is not diluted by pain or vice versa, and the number of people affected. To make a decision as to whether a use of Big Data is right or wrong, one would total the positive and negative consequences to all being affected, total up the positives and the negatives and choose the alternative with the highest amount. Rule Utilitarianism is more simplistic than Act Utilitarianism. It argues that we should follow a moral rule because its adoption would result in the greatest net increase in happiness. Big Data would be assessed relative to the weighing of its harms and benefits to society [19].

The major drawback of this approach is the ambiguity and biases inherent in trying to identify and quantify both the pros and

cons. Trying to reach consensus as to the intent and impact of Big Data would be problematic since the costs and benefits in the analysis would have to be quantified to a common economic unit of analysis. Moreover, the analysis may have inherent bias as certain issues may be afforded more weight than others by those performing the analysis. While the Utilitarian idea of balancing the pros against the cons is familiar to most as a system for judgement, in this case the process of using it to assess the value of Big Data is too complex to perform, inherently flawed by imprecise measurement, and thwarted by societies general lack of understanding about what Big Data is and the depth to which it is being used to affect their lives.

### 4.3. Social Contract Theory and Big Data

Social Contract theory emphasizes the creation of regulations and rules that rational people would agree to accept because they are to everyone's mutual benefit – as long as everyone else follows the rules. However, there are often differences between societies relative to the rules they adopt to govern their lives. For example, Europe and the United States have had confrontations about data privacy rights. In the United States, a variety of laws apply to different sectors, like health and credit. In the European Union, however, data protection is considered a fundamental right which they attempt to strictly enforce. Recently, the United States and Europe reached an agreement over their differences about what level of privacy individuals can expect when data is shared between them. This pact provides constraints on the free flow of data between Europe and the United States [20].

The key point being made here is that Social Contract Theory affords different societies the ability to envision, articulate and enforce the same moral right differently. Sometimes these differences create issues that require negotiation between societies to enable compromise mechanisms that will allow each party to protect the rules that have been established on behalf of their citizens.

Employing Social Contract Theory, one can say that an individual has the right to privacy, but also the duty not to invade the privacy of others. That said, Big Data clearly poses a challenge to both. Big data compromises moral rules and duties because in many ways it has rapidly become too powerful, too pervasive, and too essential to day-to-day life. It creates a moral challenge for societies, because people want to use the very technologies that create Big Data, yet they also want to try to control how it affects them. It may be quite difficult for societies to resolve the moral dilemmas that Big Data imposes, but their attempt to do so will inevitably be expressed in the rules they create to do just that. Inevitably, rational people will collectively determine what aspects of Big Data are morally right because of the resulting benefits they perceive as being afforded to their society.

### 4.4. Virtue Ethics and Big Data

Virtual Ethics concerns itself with the qualities that people need to flourish and be truly happy. It cares about the agent who performs an action and the appropriateness of their actions. A virtuous person does the right thing at the right time for the right reason. In this ethical perspective, moral decisions cannot be reduced to a set of rules, so instead one examines character.

Since Big Data is not a person, one must necessarily examine the feelings, character and actions of the people who deploy and use Big Data while also considering the intended and unintended effects of their actions on others. That is, we must assess those individuals who employ Big Data and determine whether their intentions and use of it are consistent with the actions of a virtuous person. Hence to assess the ethics of Big Data one could, for example, assess the moral wisdom of those who are using Big Data knowing that their actions may compromise the privacy of another. If one determines that these individuals are ignoring the rights of others by their actions, then one might reasonably argue that they are exhibiting a vice (e.g., dishonesty, greed) that reflects a deficiency in their moral character. On the other hand, if the use Big Data by medical authorities helps to prevent or manage disease, then we may conclude that this reflects the virtuous nature of those who do so. Hence, it all depends on what we conclude about the character and intent of those who employ Big Data. This means that Virtue Ethics can be arduous, because it requires that one be scrupulous in examining the action taken by someone to determine if that action is characteristic of a virtuous person. That is, after all, the approach that a virtuous person would take.

## 5. Conclusion

Ethical issues concerning Big Data are oftentimes discussed in the media. However, to more clearly understand how ethics applies to Big Data it is important to understand the tenants of the theories that inform these views. The ethical frameworks described herein each examine ethical behavior from a different perspective.

What makes ethics so valuable is that it helps us to frame our arguments about what is right or wrong using logical, rational arguments. One can use them to understand and evaluate whether they think the use of Big Data is morally right. As Quinn [19] notes, workable ethical theories all take people other than the decision maker into consideration, assume that moral good and moral principles are objective, and rely upon reasoning from facts and commonly held values. Using workable ethical theories therefore helps us to better articulate our issues with Big Data based on a clearly articulated set of moral values.

Big Data is becoming a major force in our daily lives. It affects what we know about others, what they know about us, and oftentimes how we act because of what the information it shares with us. Not only do we contribute to it, but so do the devices that we use and those that surround us.

By examining ethical theories, we can better recognize differing perspectives on Big Data-related moral situations, better understand the context and the logic of the arguments being presented, and in so doing better evaluate how the intended course of action is or should be justified.

The collection and use of Big Data has little to recommend it from an ethical perspective. This overall conclusion does indeed cast a negative light on the use of Big Data, but it also opens the door to finding ways to mitigate any ethical shortcomings. Big Data analysis is here to stay, with results facilitating advances in medicine, sustainability, behavioral analysis, and globalization to name a few. Positive outcomes provide the balance point that supports the use of Big Data; employing ethical theories helps us to better understand and manage how it affects our lives.

## 6. Future research

After mapping big data to ethical theories, an important next step is the collection of qualitative data examining the state of the users of big data. The authors have embarked on the development of two surveys to examine issues in detail. The first survey addresses “corporate” users of Big Data, including collection, analysis and sharing of data. The second survey addresses personal providers of Big Data, specifically addressing the trade-offs associated with sharing data via loyalty cards, internet usage, smart phones/tablets, social media and other smart devices.

Examination of these survey results will allow the creation of a

"Big Data Ethic" founded on the voice of the process (corporate voice) and the voice of the user (personal voice). The Big Data Ethic will provide the pathway for implementing necessary and sufficient security measures for data collection and use.

## References


[1] IBM, What Is Big Data?, 2016. https://www-01.ibm.com/software/data/bigdata/what-is-big-data.html (Accessed 12 May 16).

[2] IBM, The Four V's of Big Data, 2016. http://www.ibmbigdathub.com/infographic/four-vs-big-data (Accessed 12 May 16).

[3] SAS, Big Data: What it Is and Why it Matters, 2016. http://www.sas.com/en_th/insights/big-data/what-is-big-data.html (Accessed 8 May 16).

[4] IDC, New IDC forecast sees worldwide big data technology and services market growing to $48.6 billion in 2019, driven by wide adoption across industries. http://www.idc.com/getdoc.jsp?containerId=prUS4056011, (Accessed 9 May 16).

[5] Whatsthebigdata.com, IIA, Forrester, IDC, and Gartner on the future of Big Data analytics and cognitive computing. https://whatsthebigdata.com/2016/01/09/iia-forrester-idc-and-gartner-on-the-future-of-big-data-analytics-and-cognitive-computing/, 2016 (Accessed 13 May 16).

[6] R. Herald, 10 Big Data Privacy Problems, 2016. https://www.secureworldexpo.com/10-big-data-analytics-privacy-problems (Accessed 26 May 16).

[7] A. Zoldan, More Data, More Problems: Is Big Data Always Right?, 2016. http://www.wired.com/insights/2013/05/more-data-more-problems-is-big-data-always-right/ (Accessed 1 June 16).

[8] A. Zwitter, Big data ethics, Big Data & Soc. (July-December) (2014) 1–6.

[9] N.M. Richards, J.H. King, Big data ethics, Wake For. Law Rev. 49 (2014) 393–432.

[10] M.J. Culnan, C.C. Williams, How ethics can enhance organizational privacy: lessons from the Choicepoint and TJX data breaches, MIS Q. 33 (4) (2009) 673–687.

[11] D. Nunan, M. Di Domenico, Market research and the ethics of big data, Int. J. Mark. Res. 55 (4) (2013) 505–520.

[12] E.B. Mandinach, B.M. Parton, E.S. Gummer, R. Anderson, Ethical and appropriate data use requires data literacy, Phi Delta Kappan 96 (5) (2015) 25–28.

[13] N. Terry, in: Navigating the Incoherence of Big Data Reform Proposals, 2014 Public Health Law Conference: Intersection of Law, Policy and Prevention, Spring, 2015, pp. 44–47.

[14] J.H. King, N.M. Richards, What's up with Big Data ethics?, 2014. http://www.forbes.com/sites/oreillymedia/2014/03/28/whats-up-with-big-data-ethics/#2f6f8e012964, 2014 (Accessed 13 May 16).

[15] California State University Sacramento, Kantian ethics, 2016. http://www.csus.edu/indiv/g/gaskilld/ethics/kantian%20ethics.htm (Accessed 27 May 16).

[16] R. Cavalier, Utilitarian theories, 2016. http://caae.phil.cmu.edu/Cavalier/80130/part2/sect9.html (Accessed 1 June 16).

[17] C. Friend, Social contract theory, 2016. http://www.iep.utm.edu/soc-cont/ (Accessed 1 June 16).

[18] Stanford University, Virtue ethics, 2016. http://plato.stanford.edu/entries/ethics-virtue/ (Accessed 14 June 16).

[19] M.J. Quinn, Ethics for the Information Age, Pearson, Hoboken, New Jersey, 2017.

[20] M. Scott, U.S. and Europe in 'Safe Harbor' data deal, but legal fight may await, February 2, 2016. http://www.nytimes.com/2016/02/03/technology/us-europe-safe-harbor-data-deal.html?_r=0 (Accessed 15 June 16).