Journal of the Midwest Association for Information Systems

Volume 2020 | Issue 2 Article 3

**Date: 07-31-2020**

# Call Me BIG PAPA: An Extension of Mason's Information Ethics Framework to Big Data

**Jacob A. Young**
*Bradley University, jayoung@fsmail.bradley.edu*

**Tyler J. Smith**
*Bradley University, tjsmith3@fsmail.bradley.edu*

**Shawn H. Zheng**
*Bradley University, hzheng@fsmail.bradley.edu*

## Abstract

In 1986, Richard Mason proposed the PAPA framework to address four ethical issues society would likely face in the information age: privacy, accuracy, property, and accessibility. In this paper, we propose an extension to the PAPA framework by appending three additional issues relevant to information ethics in the big data era. First, we outline the four components of Mason's original PAPA. Second, we briefly review the major technological changes that have occurred since Mason proposed his framework. Third, we outline concepts relevant to the big data context. Fourth, we propose and discuss our extension by appending three ethical issues related to behavioral surveillance, interpretation, and governance to Mason's original PAPA framework, forming BIG PAPA. Lastly, we discuss how these issues impact practice and how they can inform future research.

**Keywords:** big data, ethics, privacy, security

DOI: 10.17705/3jmwa.000059

## 1. Introduction

Just as society entered the information age, Richard Mason proposed his PAPA framework, comprised of *privacy*, *accuracy*, *property*, and *accessibility* (Mason, 1986). Mason's seminal paper provided a much-needed warning that has remained relevant as the foundation of information ethics for decades (Peslak, 2006, p. 117). Unfortunately, it appears much of Mason's message was largely ignored as news of massive security breaches, targeted marketing, and invasive tracking technologies litter today's headlines.

Although Mason's original arguments are just as appropriate today, we offer an extension in the context of big data. First, we outline the four components of Mason's original framework. Second, we briefly review the major technological changes that have occurred since Mason proposed PAPA. Third, we outline concepts relevant to the big data context to establish a contextual foundation. Fourth, we propose and discuss our extension by appending three ethical issues related to *behavioral surveillance*, *interpretation*, and *governance* to Mason's original PAPA framework, forming BIG PAPA. Lastly, we discuss how these issues impact practice and how they can inform future research.

## 2. Mason's PAPA Framework

Before discussing our extension, we must first ground our paper by outlining the key elements of Mason's framework. Although Mason recognized that several ethical issues would become increasingly important in the information age, he focused his discussion on four primary areas of concern: privacy, accuracy, property, and accessibility. These issues were organized into the PAPA acronym. Mason posed several key questions as he introduced an issue, which we have summarized in this section.

### 2.1. Privacy

> *What information about one's self or one's associations must a person reveal to others, under what conditions and with what safeguards? What things can people keep to themselves and not be forced to reveal to others? (Mason, 1986, p. 5).*

Mason provided two examples in his discussion on privacy. First, he recounted an effort by the legislature of the state of Florida to assess whether building codes were resulting in underutilized buildings. One study tasked a researcher with observing and recording the usage of toilets, mirrors, and sinks in bathroom facilities at Tallahassee Community College. As one would expect, the school's students, faculty, and staff argued that the study invaded their privacy and violated their rights. Instead of apologizing, the state claimed that the data being collected was more valuable than the harm it caused. Amazingly, it was not until the American Civil Liberties Union got involved that the state finally suspended the study. Of course, by that time the state had already collected enough information.

Mason then illustrated what he coined the *threat of exposure by minute description*. He noted that users regularly provide information to certain parties, but without consenting to that information being shared with others or merged into a central database. In his second example, Mason recounted an unsanctioned investigation that occurred when curious programmers at the city of Chicago's computer center began cross-referencing several databases based upon employee name and I.D. At first, they identified employees with unpaid parking fines. Next, they discovered employees who owed various fees associated with the alcohol and drug abuse program. This was understandably met with outrage once news of their activity was leaked to the public. In response, the city then established new rules governing the computer center's operations to better protect employee privacy.

At that time, minute description was still largely based upon cross-sectional data captured at infrequent intervals. As technology evolved, however, collection frequency increased, and the level of detail was enhanced. Mason described these connections as *threads* that would ultimately result in the formation of all-knowing dossiers. These advancements have allowed for more threads to be woven within and among datasets, forming a permanent record for every individual. Such knowledge is primed for abuse. Not only is someone's privacy likely to be invaded, if embarrassing information falls into the wrong hands it could be leveraged for blackmail.

### 2.2. Accuracy

> *Who is responsible for the authenticity, fidelity and accuracy of information? Similarly, who is to be held accountable for errors in information and how is the injured party to be made whole? (Mason, 1986, p. 5).*

As large datasets become more interconnected, data integrity becomes paramount. Therefore, Mason stressed the importance of information accuracy by highlighting specific consequences caused by errors in computerized bank transactions and weather forecasts. The first example pertained to a bank's refusal to acknowledge receipt of a mortgage payment simply because a new computer system did not show it as paid. To prove his claim, the customer, Louis Marches, presented his coupon book to show that it had clearly been stamped as "paid" by the bank teller. Instead of acknowledging the error, the bank's employee's put their faith entirely in the computer system. Compounding the matter even further, the bank refused to accept subsequent payments until the "unpaid" payment was satisfied. This continued until the bank eventually foreclosed on the property. Making matters even worse, Marches' wife, who had been in bed recovering from a heart attack, suffered a stroke upon learning of the foreclosure from a debt collector.

The second example tells the tragic story of a man lost at sea in 1980. A weather forecast produced by the National Weather Service had stated that a nearby storm would not impact the ship's course, yet they soon found themselves in 80 knot winds and seas cresting at 60 feet. The erroneous forecast failed to predict the turbulent conditions due to a faulty buoy. Since the forecasting model was without the buoy's data, it miscalculated the storm's trajectory by several miles. Although both incidents resulted in large settlements to the injured parties, Mason recognized that we run the risk of repeating these mistakes if information systems are not carefully developed and tested.

### 2.3. Property

> *Who owns information? What are the just and fair prices for its exchange? Who owns the channels, especially the airways, through which information is transmitted? How should access to this scarce resource be allocated? (Mason, 1986, p. 5).*

Mason's focused his discussion on property around developments in artificial intelligence and the scarce resource of communication bandwidth. First, he worried about extraction of novel human thought and ingenuity and its subsequent implantation into machines. For example, will those who discover new insights be justly compensated if their ideas are replicated into billions of machines? When information is easily copied and transferred, controlling its exchange becomes increasingly difficult.

To illustrate his second concern with communication channels, Mason draws a parallel to Garrett Hardin's essay, "The Tragedy of the Commons." Since herdsmen directly benefited from each additional animal added to a pasture, but only indirectly experienced the costs of grazing, the pasture would eventually be destroyed by overuse. If the infrastructure of our data networks is also treated as a commons, we too run the risk of abusing it. On the other hand, when a communication medium is owned by a single entity, they wield considerable control over what conversations can take place. Therefore, we must find ways to ensure that a balance is found between ownership and access.

### 2.4. Accessibility

> *What information does a person or an organization have a right or privilege to obtain, under what conditions and with what safeguards? (Mason, 1986, p. 5).*

As Mason points out, literacy is critical to one's participation in, as well as the advancement of, any society. Mason explains that literacy goes much further than the ability to read. First, intellectual skills, including the ability to reason and calculate, must be developed through education. Second, one must have access to the necessary technology. Lastly, information must be accessible, else it cannot be consumed. Mason argues that one's own knowledge and economic status determines whether these three requirements have been met, and to what degree. Although technology has advanced rapidly, not all have benefited from it at the same rate. This has resulted in a large segment of the population becoming increasingly information poor. Mason continues by describing the necessary steps to access information stored in modern databases. Those who cannot complete the steps are likely to become what Mason refers to as "information drop outs" who will likely need greater assistance in the future.

### 2.5. New Social Contract

In his concluding remarks, Mason called for a new social contract that would preserve everyone's right to maximize their human potential. He made it clear that the dawn of the information age represented a critical junction for society, with the fate of future generations at stake. Mason stressed the importance of developing information systems that would "enhance the dignity of mankind" (Mason, 1986, p. 11). Mason hoped that keeping PAPA in the forefront of our minds would lead us to make wiser decisions. He encouraged developers to ensure that future information systems:

- would not "unduly invade a person's privacy;"
- produce and maintain accurate records;
- protect the available bandwidth to avoid repeating the "Tragedy of the Commons;"
- protect intellectual property; and,
- are widely accessible to foster information literacy.

In Mason's words, failing to abide by these guidelines would risk "information bankruptcy or desolation" (Mason, 1986, p. 11). Unfortunately, despite his best efforts, it seems these lessons were either largely ignored or forgotten over the following decades.

## 3. Evolution of Technology

> *"The real danger is the gradual erosion of individual liberties through automation, integration, and interconnection of many small, separate record-keeping systems, each of which alone may seem innocuous, even benevolent, and wholly justifiable" (U.S. Privacy Protection Study Commission, 1977).*

Much has changed since Mason's PAPA framework was published in 1986, especially as it pertains to data production. As Smolan & Erwitt (2012) point out, human civilization had only produced a cumulative total of 5 billion gigabytes of data from the beginning of recorded history until 2003. In 2011, we produced 5 billion gigabytes of new data every two days. By 2013, the interval was reduced to 10 minutes, and in 2015, just 10 seconds. Such rapid growth in data production propelled humanity into the big data era. Although it might sometimes feel like it, we did not get here overnight. The big data phenomenon was made possible due to revolutionary technological advancement over several decades.

In the 1980s, personal computing had truly arrived, putting the power of digital creation in the hands of the average consumer. The 1990s saw the birth of the Internet as the World Wide Web was made available to the public in 1991. Search engines, such as Yahoo! Search and Google, helped users find specific information across the vast sea of newly available data.

Although cellular devices emerged in the 1990s, it was not until the 2000s that mobile devices became readily available. Throughout his career as a science fiction writer, David Gerrold naturally made several predictions about future technology. Perhaps none rings as true today as Gerrold's (1999, p. 61) contribution to "Smart Reseller" magazine:

> *I've got a cell phone, a pocket organizer, a beeper, a calculator, a digital camera, a pocket tape recorder, a music player, and somewhere around here, I used to have a color television. Sometime in the next few years, all of those devices are going to meld into one. It will be a box less than an inch thick and smaller than a deck of cards…I call this device a Personal Information Telecommunications Agent, or Pita for short. The acronym also can stand for Pain In The Ass, which it is equally likely to be, because having all that connectivity is going to destroy what's left of everyone's privacy.*

Clearly, Gerrold's vision could not have been more prophetic. At the turn of the millennium, roughly 37 percent of Americans had a mobile phone, reaching 90 percent by the end of the decade (Kornstein, 2015). Smartphones, such as Apple's iPhone, significantly elevated mobile device capabilities, and ultimately realized Gerrold's Pita fears.

In 2009, electronic health records were mandated in the United States through the Health Information Technology for Economic and Clinical Health Act (HITECH Act) to streamline medical care. The 2010s saw wide adoption of social media platforms, such as Facebook and Twitter, promising to "connect you with the people around you," and the growth of artificial intelligence and predictive analytics (Power, 2016). Despite utopian ideals, technology was so disruptive during the 2010s that it has been considered by some as "The Decade Tech Lost Its Way" (Rahimian & Kelly, 2019).

## 4. Big Data Concepts

To deal with the emerging challenges created by these technology advancements, Doug Laney (2001) proposed a three-dimensional data challenge framework with fast increasing data volume, velocity and variety, which has become ubiquitous in definitions of big data. More recently, Van Rijmenam (2014) suggested including veracity, variability, value, and visualization. In this section, we briefly define and discuss these terms to form our contextual foundation.

Data **volume** is increasing exponentially. Gantz and Reinsel (2012) estimated that the total size of the digital universe in 2010 stood at 1.2 zettabytes (ZB). It was estimated to increase to 4.4 ZB by 2013 and was expected to grow 40% per year into next decade, almost doubling the size every two years (Gantz & Reinsel, 2012). By their estimates, the total size of digital universe in 2020 will likely reach 40 ZB (Gantz & Reinsel, 2012).

Van Rijmenam (2014, p. 5) referred to **velocity** as "the speed at which data is created, stored, analyzed, and visualized." The rapid growth and abundance of emerging technologies such as sensors, connected devices, smart personal/home/office devices, 5G networks, mobile devices, autonomous automobiles, cloud services, e-health, and virtual reality, collectively accelerated the need for real-time data accumulation (Ariyaluran Habeeb et al., 2019). Velocity is generally captured through two traits: (1) frequency of generation; (2) frequency of handing, recording, and publishing (Kitchin & McArdle, 2016). However, to holistically understand the velocity of big data, we must not only embrace technological advancement, but also simultaneously address potential privacy and security issues.

**Variety** refers to a wide range of data types, which include structured, unstructured, and semi-structured, as well as various data sources. Structured data usually resides in relational databases and is characterized by pre-defined and searchable fields. For instance, customer records such as name, phone numbers and Zip codes stored in a customer relationship management system or sensory device seeks to create a continuous, longitudinal record of usage. Data may be human or machine generated within the relational databases. Unstructured data on the other hand is not constrained by pre-defined data models or schema and resides in NoSQL databases. Common examples of unstructured data are text, audio, video files and images. Semi-structured data is a hybrid of structured and unstructured data. It usually has some defining characteristics but does not conform to a structure as rigid as what we see in relational databases (Kitchin & McArdle, 2016). For instance, pictured taken on a smart with mobile devices, are unstructured but are often geo tagged and time stamped.

**Veracity** describes the degree to which data is truthful and trusted. Veracity is increasingly being recognized as it is an essential component to extract value from big data. Veracity comprehensively describes data quality and can be translated to the reliability or consistency of the data (Gupta & Rani, 2019), or confidentiality, integrity, and availability of the data (Kepner et al., 2014). A wide range of factors can sacrifice data veracity depending on the type of data and the stage of data analysis, for example, inherent biases in data processing, untrustworthy data sources, and abnormalities. Common examples include user entry errors, duplication, and corruption, all of which pose threats to the potential value of big data.

**Variability** considers the inconsistencies of the data flow. Data loads become challenging to maintain at the same speed, especially with an increase in social media usage, which generally causes a peak in data loads when certain events occur (Katal, Wazid, & Goudar, 2013). Other variability issues relate to the multitude of data dimensions, types, and sources that eventually lead to the inconsistencies in the data.

Data **validity** in a statistical sense addresses the degree to which the tool measures what it claims to measure (Kelley, 1927). In the big data context, it refers to the correctness of data for its intended use (Gupta & Rani, 2019). For instance, sentiment analysis has recently become extremely popular in dealing with user-generated text data across social media platforms as it provides either positive, negative, or neutral sentiment predictions. While sentiment analysis can be a powerful tool in certain contexts, such as political campaigns, it has unknown validity for emotions of interest incapable to detect mood. For example, a customer with the verified purchase tag on Amazon writes a sarcastic review on a product. Although the review possesses veracity, it can easily be interpreted as a positive review. In this case, the sentiment analysis fails to capture the intended sarcasm and thus undermining the data validity.

Data's perceived **value** largely drives corporations to collect as much data as frequently as possible. While technology provided the spark, the insatiable appetite added the necessary fuel to spawn the big data inferno. Data's potential value is rooted in a desire to gain new insights, and subsequently, transform these insights into proper and timely actions that were not feasible before. When data is collected, sold, exchanged, data becomes a commodity that possesses value, the true value of big data, however, lies in the process of finding insights (Gupta & Rani, 2019). For example, the department store Macy's adjusts nearly 73 million items based on demand and supply in near-real-time for them to maximize profit. In a similar vein, retail giant Walmart utilizes semantic and text mining techniques to create better search results for its website that results in increased revenues (Desjardins, 2015).

**Visualization** refers to the use of visual elements such as charts, graphs, and maps to represent information and data (Tableau Software). The volume, veracity, and variety that characterizes big data make it particularly challenging to conduct visualization (Chen & Zhang, 2014). Firms not only use data visualization tools (e.g., Tableau, Chartist, Grafana, and D3.js) to transform large and complex data sets into simplistic, yet interactive pictures and dashboards to aid decision making, they also integrate user data with in-app visualizations. This is a common practice in social media platforms, such as the Snap Map from Snapchat, Find My Device service from Apple, as well as interactive filters from Snapchat and Instagram.

## 5. BIG PAPA

Twenty years after Mason's PAPA was published, Peslak (2006) reaffirmed the relative importance of the four original components. Privacy was perceived as the most important ethical issue, followed by accessibility, accuracy, and property. In this section, we suggest adding three issues (behavioral surveillance, interpretation, and governance) to the PAPA framework and propose an extension that we refer to as BIG PAPA.

Although Mason explored these interrelated ideas in his original work, we believe that our proposed extensions should be considered distinct issues in the context of big data. As shown in Figure 1, the behavior issue is primarily rooted in Mason's discussion on privacy, whereas interpretation is a subset of Mason's accuracy issue. Our standalone governance issue is formed from Mason's call for adequate safeguards throughout PAPA.

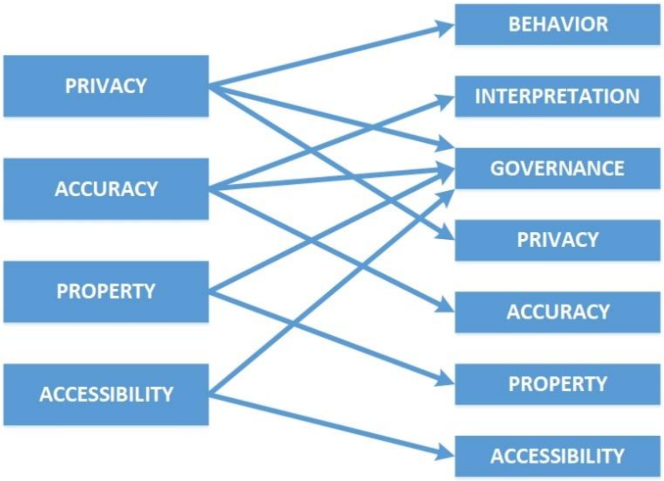

**Figure 1.**

While these issues are certainly not new, early concerns were focused on small silos of data. Therefore, any negative impact that could be caused by unethical behavior was naturally constrained. As technology and data consumption grew, these issues became more important. We illustrate this development by loosely overlapping each issue with data prior to the big data era (Figure 2), whereas each issue has become fully entrenched by big data in recent years (Figure 3).

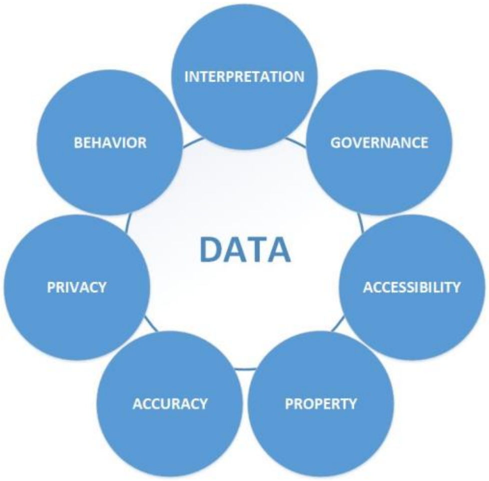

**Figure 2. Pre-Big Data Era**

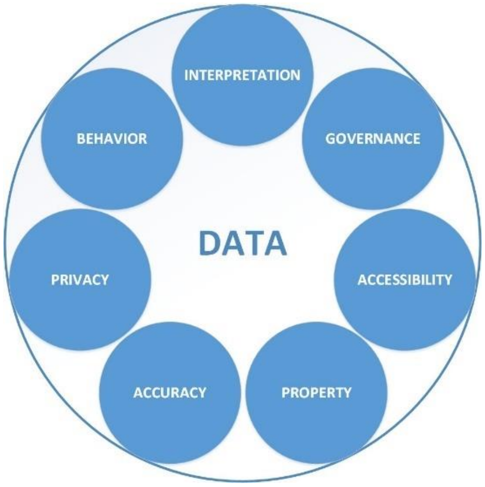

**Figure 3. Post-Big Data Era**

The seven issues that form BIG PAPA, as well as related questions, are provided in Table 1. We discuss each issue in the context of how big data has impacted society. We provide several examples to highlight the importance of our extension, as well as how Mason's original issues (privacy, accuracy, property, and accessibility) have evolved.

| **Issue** | | **Related Questions** |
|---|---|---|
| **B** | Behavioral Surveillance | • How can users be better informed before making an active choice to share data?<br>• How do we preserve individual liberty when behavior is constantly monitored? |
| **I** | Interpretation | • How can we avoid developing flawed models built upon erroneous or incomplete data?<br>• How can we reduce the likelihood of drawing erroneous conclusions?<br>• How can we educate others to recognize poor analyses? |
| **G** | Governance | • What controls are in place to handle ethical challenges related to big data?<br>• Who will watch the watchmen? |
| **P** | Privacy[1] | • What information about one's self or one's associations must a person reveal to others?<br>• Under what conditions and with what safeguards?<br>• What things can people keep to themselves and not be forced to reveal to others? |
| **A** | Accuracy[1] | • Who is responsible for the authenticity, fidelity, and accuracy of information?<br>• Similarly, who is to be held accountable for errors in information?<br>• How is the injured party to be made whole? |
| **P** | Property[1] | • Who owns information?<br>• What are the just and fair prices for its exchange?<br>• Who owns the channels, especially the airways, through which information is transmitted?<br>• How should access to this scarce resource be allocated? |
| **A** | Accessibility[1] | • What information does a person or an organization have a right or a privilege to obtain?<br>• Under what conditions and with what safeguards? |
| [1] Mason's original PAPA framework | | |

**Table 1. BIG PAPA Framework**

### 5.1. Behavioral Surveillance

> *How can users be better informed before making an active choice to share data? How do we preserve individual liberty when behavior is constantly monitored?*

When our every move, message, or even thought is being surveilled, it can have a chilling effect on our behavior (Lashmar, 2017; Richards, 2012; Stoycheff, 2016). Therefore, we focus this issue on any data that can reveal individual activity, such as location, communication, and digital interaction. Mason's discussion of privacy is certainly related to behavior, but technology at that time did not allow for the refined, targeted, and continuous data collection that is conducted today. Instead, most data were consciously provided by users themselves. Thus, we discuss behavior that is observed by virtue of active choices made by a user, as well as behavior data acquired through passive means of observation and collection. For example, a user might knowingly elect to share such data by making an active choice to opt in to obtain the benefits of a given product or service, or the behavior data could be passively collected without any clear user awareness or consent.

#### 5.1.1. Active Choice

> *"Historically, a conversation that you might have in the hallway is private by default, public through effort…Conversely, when you engage online in equally public settings such as on someone's Facebook Wall, the conversation is public by default, private through effort. You actually have to think about making something private because, by default, it is going to be accessible to a much broader audience"* (Boyd, 2010).

When users make conscious choices to engage in certain behavior, they must assume some level of risk, even if their assessment of the risk does not reflect reality. For example, many subscribe to the idea that they are doing nothing wrong, and therefore have "nothing to hide" to justify their oversharing behavior (D. J. Solove, 2007). Yet, when questioned on the issue, most will quickly concede that they wear clothes, put curtains on their windows, and lock the door when they go to the bathroom. Until users come to the realization that they should care about protecting digital behavior just as much as their physical activity, it is difficult to demonstrate how their actions can jeopardize their privacy and security.

Sadly, there are countless instances of technology use resulting in serious consequences that users never fully anticipated yet would never have been possible without their active involvement in adopting the technology. Nude photos have been stolen from cloud backups and distributed online (Peterson, Yahr, & Warrick, 2014). Residences have been broken into after occupants reveal on social media that they will not be home for an extended period (Sanchez-Garrido, 2016). Smartwatch data uploaded to fitness websites has exposed sensitive locations, including military bases (Hern, 2018). Hackers have spied on and harassed vulnerable children when insecure baby monitors and cameras are connected to the Internet (Chiu, 2019).

Of course, we are not blaming victims for the illegal actions of others, but we do advocate for the development of basic security and privacy competencies before choosing to adopt certain technologies. Unfortunately, it is usually only after personally experiencing negative effects that they are motivated to modify their behavior to reduce future risk. Ignorance may be bliss, but that does not change the fact that individual behavior regularly undermines their own privacy, security, and liberty. Obviously, much of this issue can be attributed to poor consumer awareness, which we will discuss later.

### 5.1.2. Passive Observation and Collection

> *"No one disputes that the deployment of cheap, ubiquitous video cameras has made an environment of near total surveillance technologically feasible. Whether that's a good thing or a bad thing, however, depends on how much you trust the cameraman" (Grossman, 2001).*

After it was discovered that the National Football League had secretly agreed to allow police to use facial recognition technology to search for criminals among attendees of Super Bowl XXXV, the game became scornfully known as the "Snooper Bowl" (Brey, 2004; Grossman, 2001). Today, companies such as Clearview AI are scraping image data from publicly available resources, such as social media, to improve the accuracy of their facial recognition system (Hill, 2020). These efforts have caused many to question whether participating in seemingly innocent social media fads, such as the Facebook's "10 Year Challenge," is only helping to train facial recognition algorithms to better predict aging effects (O'Neill, 2019).

While it is understood that individuals should not expect privacy in public spaces, the invasive surveillance technology that has developed in the big data era has taken 'Big Brother' to the extreme (Power, 2016). The precision and pervasive nature of the countless devices designed to track identifiable individuals today has only amplified its impact on society (Joh, 2016). For example, computer vision can be used to track vehicles through license plate readers (Lum, Hibdon, Cave, Koper, & Merola, 2011; National Law Enforcement and Corrections Technology Center, 2010; Ozer, 2010; Zmud, Wagner, Moran, & George, 2016) and law enforcement agencies have deployed International Mobile Subscriber Identity-catchers, also known as stingray devices (Bates, 2016; Boyne, 2016; Norman, 2016; Pell & Soghoian, 2014). Stingray devices pose as cell phone towers to eavesdrop on cellular communication, uniquely identify citizens, and track their movements. Fortunately, as Power (2016) points out, U.S. courts have found several innovative investigatory techniques to be unconstitutional, such as the use of thermal imaging to peer inside homes to detect marijuana plants or the placement of GPS tracking devices onto vehicles without warrants.

However, the government does not fund the bulk of these surveillance systems. Much of the tracking technology is developed, financed, and implemented by technology companies under a "surveillance as a service" business model (West, 2019). Because the data is already being collected and stored by the private sector, access can easily be granted to a variety of interested parties, such as law enforcement and data brokers—companies who buy, sell, and exchange data (Otto, Anton, & Baumer, 2007). For example, Amazon has been criticized for sharing home surveillance videos captured by its Ring products with police departments without obtaining warrants (Cox, 2019). Instead, investigators simply need to acquire consent from owners of the recording device, effectively creating a surveillance network that bypasses court oversight. Considering that Amazon incentivizes police departments to promote Ring products, consumers are essentially encouraged to fund a budding surveillance state out of their own pocket. Not only do consumers have to worry about how their data will be collected and shared by companies without their knowledge, they should also be concerned about how vulnerabilities in these devices can expose such data to other parties ("Ring under fire over weakness in video device security," 2020).

Another example is Nextdoor, a social media platform that claims to be "the best way to stay informed about what's going on in your neighborhood" and states the following in their privacy policy (Nextdoor, 2019):

> *"If you decide to invite new members to join Nextdoor, you can choose to share their residential or email address with us, or share your contacts with us, so we can send an invitation and follow-up reminders to potential new members on your behalf."*

Rather than solicit the information directly, Nextdoor actively encourages new and existing members to divulge the personal information of others, without their neighbors' knowledge or consent. The neighborhood monitoring made possible through NextDoor can be referred to as a form of lateral surveillance, where citizens take an active role in monitoring the behavior of others, which was traditionally left to professionals (Andrejevic, 2004). While many recognize the inherent privacy risks associated with using such services (Masden, Grevet, Grinter, Gilbert, & Edwards, 2014), those who use Nextdoor are unlikely to recognize how their actions have the potential to cause real harm to others, such as those who do not wish to have sensitive information shared when relocating to escape abusive relationships and/or domestic violence. Further, such services have the potential to perpetuate discrimination, effectively building a "digitally gated community" (Kurwa, 2019).

In Mason's call for a new social contract, he sought information systems that would not "unduly invade a person's privacy to avoid the indignities that the students in Tallahassee suffered" (Mason, 1986, p. 11). Regardless of whether an individual actively consents to their behavior being monitored, the level of detail captured by modern devices is not consistent with this noble goal. Is this the role that we want technology to play in our society? If not, we must make an active stand to reclaim our privacy.

### 5.2. Interpretation

> *How can we avoid developing flawed models built upon erroneous or incomplete data? How can we reduce the likelihood of drawing erroneous conclusions? How can we educate others to recognize poor analyses?*

Although obtaining a census is far more possible in the big data era (Kitchin & McArdle, 2016), most analyses are still dependent upon samples. Models are built to infer relationships and predict future events. Therefore, as analytical capability advances, we must first guard against developing flawed models built upon erroneous or incomplete data, then avoid acting upon biased or incorrect conclusions. While Mason's inclusion of accuracy might appear to satisfy this issue, we argue that one can have accurate data, yet still form incorrect conclusions due to poor modeling and/or flawed interpretations. Therefore, we have teased out this issue from Mason's broader discussion on accuracy.

#### 5.2.1. Garbage In, Garbage Out

In *Weapons of Math Destruction*, Cathy O'Neil (2016) explains how reliance on algorithms based upon flawed data can have disastrous consequences for society, such as reinforcing discriminatory practices, and at a massive scale. For example, facial recognition algorithms have been shown to have unacceptable accuracy rates and to further perpetuate inherent biases (Snow, 2018).

Further, because most algorithms are developed to gain a competitive advantage, and thus considered proprietary, we have limited visibility into how they were formed and are being employed. Efforts such as NIST's Facial Recognition Vendor Test (FRVT) have focused primarily on algorithm efficiency and effectiveness, not detecting flaws due to biased data (Introna & Wood, 2004). If algorithms are implemented before independent and comprehensive testing can be conducted, the impact of big data is more likely to result in harmful outcomes, even if the original motivations are well intended.

Given the constant creep towards predictive policing (Benbouzid, 2019; Joh, 2016), such as attempting to identify terrorists based upon whether they have purchased a life insurance policy (Van Rijmenam, 2014, p. 85), it is clear that this issue could not be more critical today. Therefore, it is essential for those dependent upon big data to ensure that analytic output meets information quality standards (Lee, Strong, Kahn, & Wang, 2002), especially before such results are applied to practice.

#### 5.2.2. Lies, Damned Lies, and Statistics

The rush to employ analytics can also place tremendous power in the hands of untrained individuals. As Wheelan (2013, p. 95) plainly put it, "statistics cannot be any smarter than the people who use them. And in some cases, they can make smart people do dumb things." One such example is referred to as the prosecutor's fallacy (Fenton & Neil, 2011).

Assume that an expert witness from a scientific laboratory correctly testifies that a given piece of biometric evidence, such as a fingerprint or DNA, can be attributed to one out of one million Americans. When a prosecutor later argues that the chances of the defendant not being the guilty party are 0.0001%, they are incorrectly interpreting the statistic. If there are 300 million people in the United States, the evidence could be attributed to a pool of 300 potential suspects. Therefore, the probability that the evidence does not belong to the defendant is 99.67% (299/300). These types of damning errors are simply not acceptable, especially when life and liberty is at stake.

In the legal arena, the onus is on the states and the federal government to train their prosecutors on how to interpret and use statistical data. Even so, the Innocence Project and the Innocence Network, a group of independent organizations that work to exonerate wrongfully convicted people often by using DNA to prove innocence. Since the project began, 367 people have been exonerated, including 21 that have served time on death row (Innocence Project, 2020). Approximately 44% of those exonerated are reported to have been wrongly convicted due to a misapplication of forensic science.

### 5.2.3. Calling Bullshit

As data becomes more accessible and as society grows increasingly reliant on data-driven products, we must not only guard against conducting poor statistical analyses, but also better educate the average citizen to recognize flawed claims. To combat this emerging issue, Carl Bergstrom and Jevin West developed a course at the University of Washington entitled "Calling Bullshit: Data Reasoning in a Digital World" (https://callingbullshit.org). Bergstrom and West (2020) West define such *bullshit* as "language, statistical figures, data graphics, and other forms of presentation intended to persuade by impressing and overwhelming a reader or listener, with a blatant disregard for truth and logical coherence." Regardless of the method of delivery, we must ensure that both producers and consumers of data understand the inherent risks associated with statistical analyses.

## 5.3. Governance

*What controls are in place to handle ethical challenges related to big data? Who will watch the watchmen?*

Governance involves ensuring that necessary controls are in place to handle future ethical challenges (Smith, Milberg, & Burke, 1996). While Mason (1986) did call for safeguards to protect individuals from unethical behavior, these arguments were embedded within each issue. We believe that governance, or the lack thereof, in the big data era is an issue in and of itself. We discuss governance with respect to the roles industry and government should play in mitigating unethical and illegal behavior (Richards & King, 2014).

### 5.3.1. Industry

Governance currently operates under an idealistic model of shared responsibility among industry, government, and consumer. Industry groups and professional associations are expected to adhere to codes of conduct, government regulators provide enforcement and ensure adequate safeguards are in place, and consumers must make informed choices. Self-regulation is certainly preferable, but when self-regulation fails, the consequences of illegal and unethical behavior must be severe enough to deter others. Unfortunately, as with all ideals, they are rarely achieved as envisioned.

Although Equifax agreed to pay up to $700 million as part of a settlement in response to their 2017 data breach, only $425 million would go towards compensating approximately 147 million affected consumers (Federal Trade Commission, 2019). Equifax exposed names, dates of birth, Social Security numbers, physical addresses, and other personal information. Considering the severity of the breach, averaging under $3.00 per victim is offensive, and certainly does not make anyone whole, as is a major purpose of bringing a tort lawsuit.

Perhaps more egregious is the fact that Google was only fined $7 million after its Street View mapping project was caught secretly capturing passwords, e-mail addresses, medical and financial records, and other personal information as Google's vehicles traveled nearly every road in the country (Streitfeld, 2013). Considering the value this data holds, such a small fine does not even register as a slap on Google's wrist.

Clearly, self-regulation under the shared responsibility approach has not been effective in discouraging unethical practices, particularly when penalties are not proportional to offenses. This reality has led to calls for much needed reform of the data broker industry (Kuempel, 2016). However, most companies in the era of big data could be classified as data brokers, especially when one considers the volume, velocity, and variety of the data they collect. Therefore, we argue that far more comprehensive reform is needed for society to have any hope of meaningfully curtailing unethical data practices.

### 5.3.2. Governments

*"So this is how liberty dies…with thunderous applause" – Queen Padmé Amidala in Star Wars: Episode III - Revenge of the Sith (Lucas, 2005).*

For as long as industry produces data, governments at all levels will be waiting to tap in. In the U.S. alone, the federal government has attempted to achieve key escrow by building backdoors into phone hardware with the Clipper chip

(Abelson et al., 2015), used the terrorist attacks of September 11, 2001 to justify global mass surveillance (Kirk, 2014), and pressured technology companies to break their own devices following the terrorist shooting in San Bernardino, California (Hack, 2016; Newkirk, 2018).

Despite passionate pleas in the name of national security, these programs are regularly found to be wasteful, ineffective, and unconstitutional once they are subject to critical review and oversight from independent bodies. For example, the government was forced to admit that the mass surveillance programs revealed by Edward Snowden failed to thwart any terrorist threats that were not already detected through traditional methods (Medine, Brand, Cook, Dempsey, & Wald, 2014a, 2014b).

As one might expect, Dinev, Hart, & Mullen (2008) found that users' willingness to provide personal information is negatively influenced by privacy and government intrusion concerns yet is positively influenced by users' perceived need for government surveillance. These concerns can be seen in the effort that ultimately led San Francisco, ironically in the heart of Silicon Valley, to ban the use of facial recognition (Barber, 2019). However, these restrictions only prevent city agencies from purchasing surveillance technology, leaving the door open for corporations to share data with government officials. In 2009, Eric Schmidt, then CEO of Google, had the following to say with respect to industry cooperation with government demands:

> *"I think judgment matters. If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place. If you really need that kind of privacy, the reality is that search engines -- including Google -- do retain this information for some time and it's important, for example, that we are all subject in the United States to the Patriot Act and it is possible that all that information could be made available to the authorities." (Schneier, 2009).*

This mindset is consistent with the "surveillance as a service" concept (West, 2019), where private firms are complicit in doing the government's bidding. However, given society's reliance on technology, if our only hope to maintain privacy is by "not doing" something, we would be prevented from essential activities, such as making financial transactions, receiving medical treatment, and merely existing in public places. Clearly, such a standard is not acceptable in a free society, yet similar statements are commonly repeated by government leaders and those responsible for the world's largest technology companies.

As Scott McNealy, cofounder of Sun Microsystems, famously proclaimed in 1999, "You have zero privacy anyway…Get over it" (Sprenger, 1999). Yet, in 2015, he stated that "It scares me to death when the NSA or the IRS know things about my personal life and how I vote. Every American ought to be very afraid of big government" (Noyes, 2015). Likewise, Mark Zuckerberg, founder of Facebook, who has built his social media empire on "making people more open and connected" with little regard to limiting the sharing of data to third parties, paid more than $30 million for four houses surrounding his own residence in an effort to protect his own personal privacy (Nicks, 2016).

Given that industry struggles to self-regulate and some of the most invasive practices have been employed by governments, perhaps we should look elsewhere for a solution to the governance problem? One intriguing option is self-sovereign identity (Mühle, Grüner, Gayvoronskaya, & Meinel, 2018; Tobin & Reed, 2017). Self-sovereign identity allows a user to share the minimum level of detail for a given situation by presenting verifiable claims.

For example, consider a situation where an individual must prove that they are of a certain age. Instead of being forced to produce a driver's license or other official document that unnecessarily reveals sensitive information, he or she could simply present a verifiable claim that has been digitally signed by a government agency. The claim could be as simple as stating that the individual is over 18 years old. Not only does this prove eligibility, it does so without sharing any other sensitive information typically present on other authoritative documents, such as birthdate or home address.

While this is an extremely simple example, self-sovereign identity can be used for just about any conceivable use case. Ultimately, self-sovereign identity puts control back into the hands of the individual rather than continue to populate centralized repositories maintained by increasingly powerful data brokers and governments. Widespread adoption of self-sovereign identity would dramatically curb the current appetite for consumer data by starving data brokers of the minute description that makes a dataset so valuable.

### 5.4. Privacy

Mason's discussion on privacy highlighted the early beginnings of what would later become known as data brokers. These companies are motivated to obtain as much data as possible due to both perceived and real value. While data brokers certainly present a significant threat to privacy, we approach the issue by focusing on consumer awareness, calls for the “right to be forgotten,” as well as how privacy rights have been viewed in both common law and statutes.

### 5.4.1. Consumer Awareness

Several measures have been developed to assess consumer privacy concerns: Smith, Milberg, & Burke (1996) developed their global information privacy concern (GIPC) scale; Stewart & Segars (2002) developed the Concern for Information Privacy Instrument; and, Malhotra, Kim, & Agarwal (2004) developed the Internet Users' Information Privacy Concerns (IUIPC) instrument. Studies have also focused on particular use cases. For example, Bélanger, Hiller, & Smith (2002) investigated how four trust indices impacted consumer perceptions of privacy in e-commerce and Awad & Krishnan (2006) found that consumers who valued information transparency were less likely to provide information to aid in personalized marketing.

While this stream of research has yielded insights regarding privacy concerns, improving consumer awareness in practice has been a challenge, especially when the primary method of delivery has been through terms of service and privacy policies. Milne, Culnan, & Greene (2006) evaluated over three hundred privacy notices and concluded that in just two years, notice readability had declined, while length had increased. McDonald & Cranor (2008) later estimated that it would take approximately 244 hours per year for an individual to read the privacy policies for each website visited. Further, most terms of service agreements must be accepted in their entirety, which prevents users from negotiating fairer exchanges (Walker, 2012). Although there have been attempts to simplify terms of service, such as “Terms of Service; Didn’t Read” (https://tosdr.org/), placing the privacy burden entirely on the consumer is not only unreasonable, but completely impractical for those who value their privacy, yet do not understand the risks associated with certain technology.

Boritz & No (2011) discovered that much of the privacy research had been conducted in the early 2000s, creating a void that was not accounting for rapidly developing technology. Several other studies have reviewed privacy literature in information systems (Bélanger & Crossler, 2011; Pavlou, 2011; Smith, Dinev, & Xu, 2011) in an effort to spur future privacy research. Lowry, Dinev, & Willison (2017) encouraged researchers to treat security and privacy as an IS artifact. However, despite greater attention to the privacy issue in academia, it is difficult to see any substantial progress being made in practice with respect to increasing consumer awareness.

### 5.4.2. Right to be Forgotten

One of the reasons it has been difficult for the average consumer to anticipate negative consequences of certain technologies is due to the impressive speed at which they have advanced. Without having enough time to observe and process how these developments might impact them, many participate, only to eventually find themselves regretting previous behavior, such as posting personal information on social media (Rosen, 2012).

Because of this, many have called for an online undo option, or Ctrl+Z (Jones, 2016), known as the “right to be forgotten,” that would allow users to remove information from the seemingly permanent record made possible through the Internet. In January 2012, the European Commission announced that they would recognize this right, which was formally adopted in 2016. Viviane Reding, European Commissioner for Justice, Fundamental Rights, and Citizenship, explained the basic premise behind the right to be forgotten: “If an individual no longer wants his personal data to be processed or stored by a data controller, and if there is no legitimate reason for keeping it, the data should be removed from their system” (Reding, 2012, p. 5).

Despite the obvious benefits, valid criticisms against the right to be forgotten have been raised regarding how it might conflict with the rights of others, namely rights to free speech and freedom of expression (Walker, 2012). However, proponents recognize that requests to remove information under the right to be forgotten should be limited to content that the subjects of the data uploaded themselves (Ausloos, 2012; Walker, 2012). This approach respects individual privacy yet prevents the right to be forgotten from becoming a magic eraser that can be used to censor others.

Complicating the issue from a practical standpoint, however, is that content submitted by users can propagate to other online services (Ausloos, 2012), whose right to share would then be protected by their own freedom of expression. Further, content marked for deletion could still exist on data backups for a considerable amount of time. Therefore, the unfortunate reality is that the only reliable way for users to prevent others from maintaining a permanent record is to avoid sharing or allowing the data to be collected in the first place.

### 5.4.3. Right to Privacy in Common Law

In their seminal law review article, Samuel Warren & Louis Brandeis (1890) recognized that U.S. law was beginning to distinguish physical and psychological harms. The common law up to this point only recognized physical harms, such as battery. As Warren and Brandeis pointed out, the law began to recognize psychological harm that can result from the reasonable apprehension of physical harm. Warren and Brandeis applied this extension to scenarios involving publishing psychological damaging information in tabloids. The traditional common law notions of property, contracts, and copyright were inadequate to remedy psychological harms. Common law remedies still may not be adequate today.

Nickelodeon Consumer Privacy Litig., 827 F.3d 262, 293 (3rd Cir., 2016), involved a class action filed on behalf of 13-year-old plaintiff's alleging that Google and Viacom unlawfully collected information on the websites and videos the children viewed. The Plaintiffs claimed Google and Viacom committed the common law tort of intrusion upon seclusion. One of the four types of invasion of privacy torts, a defendant is found liable for the tort of intrusion upon seclusion when the plaintiff can show (1) an intentional intrusion (2) upon the seclusion of another that is (3) highly offensive to a reasonable person. Though the plaintiff's other claims failed, the plaintiffs in Nickelodeon were successfully able to prove that Google and Viacom had committed the tort of intrusion upon seclusion. What is highly offensive in the context of this tort to a reasonable person is guided by court decisions and lawyer arguments.

### 5.4.4. Right to Privacy in Statutes

Though some sectors in the United States are regulated in terms of privacy, some are not. The "sectoral approach" to privacy regulation creates a patchwork of legislation that overlaps in multiple areas of the economy and does not at all regulate in others (D. Solove, 2015). Some members of Congress have supported legislation to create a federal privacy law, though the efforts have yet to go anywhere. Giving a private cause of action and the preemption state privacy laws are some continual sticking points (Feiner, 2019).

The other approach to privacy regulation is the "omnibus approach" taken by the European Union. This approach is designed to regulate privacy no matter the sector and impacts American business, or any business that does business with citizens of the EU. The General Data Protection Regulation (GDPR) has been in effect for over a year and is a significant step in improving privacy regulation (Burgess, 2019). The GDPR attempts to alter how businesses handle their information and give consumers more control over their information.

While Congress has been unable to pass comprehensive bi-partisan legislation, some states have filled the gap. The California Consumer Privacy Act is the most recent state level legislation to directly address information privacy. This law is designed to allow consumers the ability to ask companies that collect data on consumers to see the information. Consumers can see personal data to include smartphone data, physical locations, and biometric data (Cowan & Singer, 2020).

The right to privacy is right difficult to define in the law, yet should be central to discussion on data and its use. Many laws dealing with privacy are centered on the premise that consumers will understand how their data is being used simply by being told. This fails to account for the true understanding of the content in the disclosure and does not adequately protect consumers. The right to be forgotten, much like the general right to privacy, is not easily implemented. The approach taken in the United States, based primarily on informed consent, is a confusing patchwork that leaves gaps in some areas and overlaps in others. The omnibus approach taken by the EU with the GDRP, is a significant step forward in creating an effective framework for the protection of consumer data.

Given the speed at which technology advances, it is unlikely for any law to fully protect privacy. The right to be forgotten is a desirable option, but the only guaranteed defense is to never share sensitive information in the first place. Therefore, it is imperative for consumers to understand how to protect their own interests.

## 5.5. Accuracy

Mason's accuracy issue primarily related to misinformation; errors and omissions that undermine data authenticity, fidelity, and accuracy. However, when it comes to the big data era, privacy advocates and industry should also be interested in how countermeasures can undermine all three characteristics. For example, obfuscation and disinformation can be employed by users to protect their privacy by shielding activity or information from others.

### 5.5.1. Misinformation

Mason focused his discussion on misinformation, such as inaccurate or incomplete data, and this issue could not be more relevant today. For example, a common remedy to data breaches, as will be discussed below, is to offer credit monitoring services. However, merely being told about a potentially material error on one's credit report is not enough as it is often difficult to correct. Under the Fair Credit Reporting Act, firms must "adopt reasonable procedures for meeting the needs of commerce for consumer credit, personnel, insurance, and other information in a manner which is fair and equitable to the consumer, with regard to the confidentiality, accuracy, relevancy, and proper utilization of such information…" (15 U.S.C § 1681 (b)).

Those "reasonable procedures" however, often result in significant errors that can require consumers to expend great energy, often to no avail, to resolve grave errors on their credit reports. More than one in five consumers have a "potentially material error" on their credit report (Klein, 2017). The errors may not originate from the reporting agency since they merely collect what data they are given, as was the case when the Social Security Administration mistakenly lists 6,000 people a year as deceased. One consumer did not discover that her credit report indicated she was deceased until she went to apply for an auto loan (T. Anderson, 2017).

### 5.5.2. Obfuscation

Given the overwhelming number of ways that one's privacy and security can be violated, it is certainly understandable why many resign themselves to living in what can appear to be an inescapable panopticon. Yet, for those who seek out solutions, there are several tools available to increase user anonymity and obfuscate their activity (Bazzell, 2019).

For example, Howe (2015) surveyed several countermeasures that could be used to mitigate various contemporary threats to privacy in both the digital and physical worlds. TrackMeNot (http://trackmenot.io/) was designed to limit the effectiveness of collecting a user's search engine history by shrouding each query amongst additional searches, hiding the true search amongst useless noise. *I Like What I See* (https://github.com/sklise/i-like-what-i-see), was developed to automatically click every instance of the word "like" present on a given website, preventing others from knowing a Facebook user's true interests. *ScareMail* (https://bengrosser.com/projects/scaremail/), another Chrome extension, appends nonsensical text to emails that is expected to be flagged by surveillance monitoring systems, such as the National Security Agency's PRISM and XKeyscore programs.

The *Facial Weaponization Suite* (http://zachblas.info/works/facial-weaponization-suite/) attempts to render facial recognition useless by wearing masks constructed based upon aggregated facial data. The *Invisible* kit (http://biogenfutur.es/) is an open source project consisting of two sprays, one designed to remove as much DNA as possible, and the other to obfuscate any that remains. Each of these highlighted countermeasures not only limit the usable information that is shared with certain systems or people, but intentionally seek to undermine and sabotage the intended purpose of collecting such data.

### 5.5.3. Disinformation

In addition to obfuscation, users can also employ disinformation to increase privacy (Whang, 2012). Alexander & Smith (2011, p. 58) defined disinformation as "intentional deception in communications scenarios" and described two primary types: destructive and constructive. Destructive disinformation would involve removing key information through redaction, whereas constructive might involve adding false information to a document. A simple example of disinformation would include using a fictitious persona to create an online account or subscribing to magazines to populate databases with inaccurate information as to who resides at a physical address.

Although obfuscation and disinformation techniques can increase one's privacy, employing them could potentially cause algorithms to be less reliable, undermining the value and effectiveness of big data. Therefore, organizations who desire a high level of accuracy must be aware that a growing percentage of consumer data will no longer be trustworthy.

## 5.6. Property

Mason's conceptualization of the property issue focused primarily on questions related to data ownership, equitable exchange, and access to transmission channels. In the big data era, the reliance on cloud storage and Bring Your Own Device (BYOD) policies have further complicated the already difficult ownership issue. Secondary information use has resulted in data exchanges that are largely inequitable for the consumer. Debates over net neutrality also threaten to impede access to electronic information and resources.

### 5.6.1. Bring Your Own Device

Instead of issuing company-owned equipment, such as smartphones, laptops, and tablets to employees, organizations have largely adopted the BYOD approach (Shim, Mittleman, Welke, French, & Guo, 2013). While providing several benefits (e.g., accessibility, convenience, flexibility, employee satisfaction, productivity, innovation, and cost-savings), this phenomenon also leads to several privacy and security implications, especially as it pertains to data breaches stemming from lost and stolen devices (Garba, Armarego, Murray, & Kenworthy, 2015).

However, as Mason pointed out, there are even cloudier situations that have no obvious answers. For example, who owns the data stored on an employee's personal phone when it is used for business purposes? If the company owns the data, how can they access it if they do not own the device? Further, should organizations be able to compel employees to install applications on employee owned devices?

With these questions in mind, it is easy to see how the BYOD approach also leads to numerous legal implications. Which laws and regulations apply to company BYOD policies largely depend on the sector the business is in and type of data collected. Companies need to sift through the data breach notification laws at the state and federal levels, state and federal laws and regulations on data security, international data protection laws, legal procedures that relate to eDiscovery, confidentiality obligations, contractual obligations, trade secret protection, and employment law related issues (Privacy Rights Clearninghouse, 2014).

### 5.6.2. Right to Share

Secondary information use refers to "the use of personal information for other purposes subsequent to the original transaction between an individual and an organization when the information was collected" (Culnan, 1993, p. 342). However, since terms and conditions are difficult and time consuming to read (McDonald & Cranor, 2008; Milne et al., 2006), users are unlikely to be fully informed as to what the organization can do with their data.

Further complicating this issue is the third-party doctrine, which essentially eliminates a user data rights, while granting organizations the "right to share" consumer data as they please (President's Council of Advisors on Science and Technology, 2014). The third-party doctrine is the legal proposition that people are not entitled to an expectation of privacy for information they voluntarily give to third parties. Therefore, the third-party doctrine currently applies to any data stored or processed through cloud services. Considering the digitization of every aspect of society, coupled with the shift back to centralized computing, users are left with little to no choice in preserving their privacy rights.

For example, in *Katz v. United States*, the Supreme Court proclaimed that "what a person knowingly exposes to the public, even in his own home or office, is not a subject of Fourth Amendment protection" (Thompson, 2014). Numerous subsequent cases explained the relatively limited reach of the doctrine in the pre-digital age. The debate as to whether the third-party doctrine unnecessarily constricts the privacy rights of Americans has continued well into the digital age.

## 5.7. Accessibility

Mason's original concern with respect to accessibility was that poor and disenfranchised communities would fall further behind due to limited access to information. While this fear did materialize, some could argue that the ease in which users can access information about others carries grave threats to personal privacy and safety.

### 5.7.1. Access to Knowledge

The Internet has drastically increased the amount of information available, yet many areas still lack access to reliable broadband connectivity due to infrastructure and transmission costs. Several efforts, such as OneWeb and SpaceX's Starlink, are currently underway to reduce these costs and provide global access to broadband Internet by placing massive satellite constellations in much lower orbits to achieve lower latency than existing satellite Internet service providers (del Portillo, Cameron, & Crawley, 2018). As primary and secondary education continues to shift toward digital learning through laptops and tablets, the accessibility gap between the haves and have nots will only continue to grow. The accessibility issue has never been more evident than during the 2020 COVID-19 pandemic. With educational institutions and businesses suddenly forced to transition to online delivery and work-from-home conditions, those without a reliable Internet connection or access to necessary hardware and software suffered severe hardships.

### 5.7.2. Too Much Access?

> *"Just because something is publicly accessible does not mean that people want it to be publicized" (Boyd, 2010).*

The technology that spawned the big data phenomenon has also made it possible for personal and sensitive information to be easily collected and disclosed on the Internet. When done intentionally, this form of cyberbullying behavior is commonly referred to as "doxing" due to the release of documents or information without the victim's consent (Douglas, 2016; Li, 2018). Such information is commonly obtained through open-source intelligence (OSINT) gathering techniques. Hackers regularly use OSINT to research individuals associated with their target organizations to increase the effectiveness of cyberattacks (Hayes & Cappa, 2018).

These same techniques have also been used in "swatting" attacks where someone calls the police to make a false report, such as claiming that a hostage situation is taking place at the target's residence, to cause a heavy police response (Jaffe, 2016; Li, 2018). The response typically involves deploying heavily armed SWAT (special weapons and tactics) teams, hence the name. Swatting attacks have been made against several celebrities, politicians, and other public figures, such as cybersecurity journalist Brian Krebs. The attack against Krebs was in retaliation for his story that identified a group of cybercriminals responsible for doxing public officials (Vaas, 2013, 2016). While these incidents were resolved without major harm, others have not been so fortunate. In 2015, the police chief for Sentinel, Oklahoma was shot by a homeowner when responding to a bomb threat called in by someone else (Slipke, 2015). In 2017, a dispute between Casey Viner and Shane Gaskill over an online video game resulted in an innocent man, Andrew Finch, being shot and killed by police as he exited his home in Wichita, Kansas (Jaffe, 2020).

Other examples can be found in the documentary series *Don't F**k with Cats: Hunting An Internet Killer* (Lewis, 2019). Concerned citizens took it upon themselves to track down those responsible for producing and uploading several viral videos showing a man torturing and killing kittens. Members of a Facebook group conducted extensive OSINT activities and identified several potential suspects. A social media firestorm was waged against one of the early suspects after his information was doxed among the Facebook group. It was later determined that the suspect was not involved, but not before he tragically committed suicide under the weight of false accusations. While their efforts were underway, a group member who had used a pseudonym on her Facebook profile was sent a link to an unsettling video that showed someone walking through the Las Vegas casino where she worked. Although the group eventually identified the individual responsible for the videos, their efforts were not successful in preventing Luka Magnotta from escalating his crimes. He continued to post additional videos until he was ultimately arrested and convicted of murdering Jun Lin, an international student at Concordia University, in 2012.

As these instances demonstrate, an affected individual or organization might be completely unaware that anyone is targeting them or that the sensitive information was available online until it has already resulted in harm. Thus, enhanced accessibility to data can also lead to disturbing and heartbreaking outcomes.

### 5.7.3. Unauthorized Access

While the personal computer ushered in distributed computing in the 1980s, emerging technologies are reverting to centralized computing due to the processing power needed to leverage big data. This trend increases the need for adequate connectivity for those who wish to adopt such systems. However, centralized databases also introduce several issues. For example, the push to convert patient data into Electronic Medical Records (EMRs) and increase access through health information exchanges can lead to security and privacy issues (Angst, 2010; Angst & Agarwal, 2009; Burns, Young, Ellis, Courtney, & Roberts, 2015). As we continue to desegregate data in the name of accessibility, we amplify breach magnitude and severity.

Although there is a clear need for anonymity (Bellaby, 2018), anonymization methods and policies, such as statistical disclosure control, are routinely proven inadequate in protecting consumer privacy in the big data era (Riederer, Kim, Chaintreau, Korula, & Lattanzi, 2016). In 1997, when Massachusetts' Group Insurance Commission decided to release records of hospital visits to researchers, Massachusetts governor William Weld gave reassurances that the data was anonymized and would not violate patient privacy (N. Anderson, 2009; Ohm, 2010). Sensing a challenge, Latanya Sweeney leveraged her research on U.S. Census and voter registration data (Sweeney, 2000), managed to identify Weld's health information, and his diagnoses and prescriptions to him.

Narayanan and Shmatikov (2008) demonstrated how users could be identified when using two separate data sources. Netflix, as part of their Netflix Prize data mining contest, had released over 100 million of what they believed to be sufficiently anonymized movie ratings from approximately 500,000 subscribers. When Narayanan and Shmatikov analyzed the ratings with similar data from Internet Movie Database (IMDb), they were able to identify anonymized Netflix ratings based upon public IMDb ratings. In discussing the implications of their study, Narayanan and Shmatikov (2008, p. 123) explained how seemingly worthless information could impact an individual's privacy:

> *First, his political orientation may be revealed by his strong opinions about "Power and Terror: Noam Chomsky in Our Times" and "Fahrenheit 9/11," and his religious views by his ratings on "Jesus of Nazareth" and "The Gospel of John." Even though one should not make inferences solely from someone's movie preferences, in many workplaces and social settings opinions about movies with predominantly gay themes such as "Bent" and "Queer as folk" (both present and rated in this person's Netflix record) would*

> *be considered sensitive. In any case, it should be for the individual and not for Netflix to decide whether to reveal them publicly.*

Successful reidentification also occurred when researchers at the Whitehead Institute were interested in assessing the risk of sharing anonymized DNA data (Gymrek, McGuire, Golan, Halperin, & Erlich, 2013; Van Rijmenam, 2014). The team found that DNA data sets that have been stripped of identifiers can still be attributed to surnames by examining specific genetic data and then cross-referencing freely available data sources on the Internet. What is particularly dangerous is that individuals can jeopardize relatives by sharing their own DNA, all without informing or obtaining consent. Given the rise of genetic testing services that are marketed to the average consumer, such as 23andMe and Ancestry.com, these findings have serious implications for the privacy and security of genetic information.

#### 5.7.4. Data Breaches and Reporting

With more and more data stored, used, and transferred for business purposes, data breaches are increasingly common. Despite the staggering number and scope of these data breaches, significant challenges for consumers and corporations are created due to a lack of consistent data protection framework (Tschider, 2015). Much like other areas regarding data and privacy, federal laws related to cybersecurity are sector specific. These laws include provisions in Gramm-Leach-Bliley and the HIPPA Breach Notification Rule.

Unfortunately, a comprehensive federal data breach notification and protection law does not seem likely anytime soon. Numerous bills, such as the Data Accountability and Trust Act (Rush, 2019), have been introduced recently in Congress, but have failed to gain traction. This bill, H.R. 1282 would require the Federal Trade Commission (FTC) to require certain businesses and organizations to establish security practices for the treatment and protection of personal information and provide specified notice and offer credit-monitoring services in the event of a breach.

Adding to the complication of sector specific laws, all 50 states currently have their own form of data breach notification statutes. Some of the definitions and requirements of the laws that can vary greatly include: 1) what type personally identifiable information triggers a breach notification obligation to individuals, 2) what form of data that information is in, 3) when notice must be given to individuals, 4) what form of that notice is permitted, 5) what must information about the breach must be included in the notice, 6) what states require notification to state agencies, and 7) when notification to the credit reporting agencies is required (Millar & Marshall, 2019).

## 6. Discussion

### 6.1. Contributions to Literature

Although Mason's PAPA is just as relevant today as it was in 1986, we have highlighted several modern examples that we must consider as we usher in a new wave of technological advancement. Our extension also distinguished three issues that were deeply embedded within Mason's framework but deserve increased attention. First, we examined how behavioral surveillance has become a significant threat to individual privacy and liberty. Second, we illustrated how data errors and misinterpretations can lead to devastating outcomes, especially when applied in high stakes situations. Third, we also explored how both industry and government have failed to properly regulate unethical behavior. Lastly, we revisited the four original issues comprising Mason's PAPA by examining how technology has not only impacted our relationship with information, but more importantly one another.

### 6.2. Practical Implications

As we enter the 2020s, the mass proliferation of the Internet of Things (IoT) will likely be its most defining characteristic. Nearly every category of electronic device sold today has built-in connectivity, essentially eliminating any remaining firewall between the digital and physical worlds. We encourage professionals in all fields to assess the ethical implications of their work. Identifying potential issues during the infancy of any new endeavor allows for safeguards to be embedded during the development process. Employing an "ethical by design" approach to technology development would also better protect stakeholders by ensuring that adequate thought has been put into the ethical implications for each element of the system (Beard & Longstaff, 2018).

### 6.3. Legal Implications

With the confusing web of state and federal privacy laws, the onus remains primarily on individuals and corporations to reevaluate their behavior through the interpretation of statistical models and its application to governance. Corporations and industries have a direct impact on the behavior of individual consumers and should take steps to educate consumers beyond bare minimum informed consent requirements on how their data is collected and used. Individuals must be able to understand the value their data has to corporations and how their everyday behavior contributes to the enhancement of that value. Additionally, corporations and even consumers still must advocate for strengthened privacy laws that protect consumers from forms of exploitation using their data.

## 7. Conclusion

We have provided several contemporary examples that continue to demonstrate the importance of Mason's seminal work. Each issue that Mason identified in his PAPA framework over 30 years ago is just as critical, if not more so, in the big data era. We also proposed our BIG PAPA extension to address emerging issues with respect to behavioral surveillance, interpretation of analytical models, and much needed governance. It is our hope that greater attention will be paid to these areas in both research and practice. We simply cannot afford to continue to ignore Mason's warnings any longer.

## 8. References

Abelson, H., Anderson, R., Bellovin, S. M., Benaloh, J., Blaze, M., Diffie, W., … Weitzner, D. J. (2015). Keys under doormats: Mandating insecurity by requiring government access to all data and communications. *Journal of Cybersecurity*, *1*(1), 69–79. https://doi.org/10.1093/cybsec/tyv009

Alexander, J. M., & Smith, J. M. (2011). Disinformation: A Taxonomy. *IEEE Security & Privacy Magazine*, *9*(1), 58–63. https://doi.org/10.1109/MSP.2010.141

Anderson, N. (2009, September 8). “Anonymized” data really isn’t—and here’s why not. Retrieved from https://arstechnica.com/tech-policy/2009/09/your-secrets-live-online-in-databases-of-ruin/

Anderson, T. (2017, January 13). You may be dead: Every year, Social Security falsely lists 6,000 people as deceased. Retrieved from https://www.cnbc.com/2017/01/12/social-security-falsely-lists-6000-people-a-year-as-dead.html

Andrejevic, M. (2004). The work of watching one another: Lateral surveillance, risk, and governance. *Surveillance and Society*, *2*(4), 479–497. https://doi.org/10.24908/ss.v2i4.3359

Angst, C. M. (2010). Protect My Privacy or Support the Common-Good? Ethical Questions About Electronic Health Information Exchanges. *Journal of Business Ethics*, *90*(S2), 169–178. https://doi.org/10.1007/s10551-010-0385-5

Angst, C. M., & Agarwal, R. (2009). Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion. *MIS Quarterly*, *33*(2), 339–370.

Ariyaluran Habeeb, R. A., Nasaruddin, F., Gani, A., Amanullah, M. A., Abaker Targio Hashem, I., Ahmed, E., & Imran, M. (2019). Clustering-based real-time anomaly detection—A breakthrough in big data technologies. *Transactions on Emerging Telecommunications Technologies*, *e3647*(April), 1–27. https://doi.org/10.1002/ett.3647

Ausloos, J. (2012). The ‘Right to be Forgotten’ – Worth remembering? *Computer Law & Security Review*, *28*(2), 143–152. https://doi.org/10.1016/j.clsr.2012.01.006

Awad, N. F., & Krishnan, M. S. S. (2006). The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization. *MIS Quarterly*, *30*(1), 13. https://doi.org/10.2307/25148715

Barber, G. (2019, March 14). San Francisco Bans Agency Use of Facial-Recognition Tech. Retrieved from https://www.wired.com/story/san-francisco-bans-use-facial-recognition-tech/

Bates, A. (2016). *Stingray: A New Frontier in Police Surveillance*. Washington, D.C. Retrieved from https://www.cato.org/publications/policy-analysis/stingray-new-frontier-police-surveillance?

Bazzell, M. (2019). *Extreme Privacy: What It Takes to Disappear in America*. (A. Martin, M. S. Williams, & J. Engstrom, Eds.).

Beard, M., & Longstaff, S. (2018). *Ethical by Design: Principles for Good Technology*. Sydney, Australia.

Bélanger, F., & Crossler, R. E. (2011). Privacy in the digital age: A review of information privacy research in information systems. *MIS Quarterly: Management Information Systems*, *35*(4), 1017–1041. https://doi.org/10.2307/41409971

Bélanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. *The Journal of Strategic Information Systems*, *11*(3–4), 245–270. https://doi.org/10.1016/S0963-8687(02)00018-5

Bellaby, R. W. (2018). Going dark: anonymising technology in cyberspace. *Ethics and Information Technology*, *20*(3), 189–204. https://doi.org/10.1007/s10676-018-9458-4

Benbouzid, B. (2019). To predict and to manage. Predictive policing in the United States. *Big Data & Society*, *6*(1), 1–13. https://doi.org/10.1177/2053951719861703

Bergstrom, C. T., & West, J. D. (2020). Calling Bullshit: Frequently Asked Questions. Retrieved from https://callingbullshit.org/FAQ.html

Boritz, J. E., & No, W. G. (2011). E-Commerce and Privacy: Exploring What We Know and Opportunities for Future Discovery. *Journal of Information Systems*, *25*(2), 11–45. https://doi.org/10.2308/isys-10090

Boyd, D. (2010, March 13). Making Sense of Privacy and Publicity. Retrieved from http://www.danah.org/papers/talks/2010/SXSW2010.html

Boyne, S. (2016). Stingray Technology, the Exclusionary Rule, and the Future of Privacy: A Cautionary Tale. *West Virginia Law Review*, *119*(3), 915–939. https://doi.org/10.2139/ssrn.2911844

Brey, P. (2004). Ethical aspects of facial recognition systems in public places. *Journal of Information, Communication and Ethics in Society*, *2*(2), 97–109. https://doi.org/10.1108/14779960480000246

Burgess, M. (2019, January 21). What is GDPR? The summary guide to GDPR compliance in the UK. Retrieved from https://www.wired.co.uk/article/what-is-gdpr-uk-eu-legislation-compliance-summary-fines-2018

Burns, A., Young, J. A., Ellis, T. S., Courtney, J. F., & Roberts, T. L. (2015). Exploring the Role of Contextual

Integrity in Electronic Medical Record (EMR) System Workaround Decisions: An Information Security and Privacy Perspective. *AIS Transactions on Human-Computer Interaction*, *7*(3), 142–165.

Chen, C. L. P., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314–347. https://doi.org/10.1016/j.ins.2014.01.015

Chiu, A. (2019, December 12). She installed a Ring camera in her children's room for 'peace of mind.' A hacker accessed it and harassed her 8-year-old daughter. Retrieved from https://www.washingtonpost.com/nation/2019/12/12/she-installed-ring-camera-her-childrens-room-peace-mind-hacker-accessed-it-harassed-her-year-old-daughter/

Cowan, J., & Singer, N. (2020, January 3). How California's New Privacy Law Affects You. Retrieved from https://www.nytimes.com/2020/01/03/us/ccpa-california-privacy-law.html

Cox, K. (2019, August 6). Police can get your Ring doorbell footage without a warrant, report says. Retrieved from https://arstechnica.com/tech-policy/2019/08/police-can-get-your-ring-doorbell-footage-without-a-warrant-report-says/

Culnan, M. J. (1993). "How Did They Get My Name?": An Exploratory Investigation of Consumer Attitudes toward Secondary Information Use. *MIS Quarterly*, (September), 341–363.

del Portillo, I., Cameron, B. G., & Crawley, E. F. (2018). A Technical Comparison of Three Low Earth Orbit Satellite Constellation Systems to Provide Global Broadband. In *69th International Astronautical Congress* (pp. 1–15). Bremen, Germany.

Desjardins, J. (2015, July 29). Order From Chaos: How Big Data Will Change the World. Retrieved from https://www.visualcapitalist.com/order-from-chaos-how-big-data-will-change-the-world/

Dinev, T., Hart, P., & Mullen, M. R. (2008). Internet privacy concerns and beliefs about government surveillance – An empirical investigation. *The Journal of Strategic Information Systems*, *17*(3), 214–233. https://doi.org/10.1016/j.jsis.2007.09.002

Douglas, D. M. (2016). Doxing: a conceptual analysis. *Ethics and Information Technology*, *18*(3), 199–210. https://doi.org/10.1007/s10676-016-9406-0

Federal Trade Commission. (2019, July 22). Equifax to Pay $575 Million as Part of Settlement with FTC, CFPB, and States Related to 2017 Data Breach. Retrieved from https://www.ftc.gov/news-events/press-releases/2019/07/equifax-pay-575-million-part-settlement-ftc-cfpb-states-related

Feiner, L. (2019, December 4). A federal privacy law is starting to crystallize, but Democrats and Republicans can't agree on how to do it. Retrieved from https://www.cnbc.com/2019/12/04/a-federal-privacy-law-is-starting-to-crystallize-senators-remain-divided-over-details.html

Fenton, N., & Neil, M. (2011). Avoiding probabilistic reasoning fallacies in legal practice using Bayesian networks. *Australian Journal of Legal Philosophy*, *36*(2011), 114–150.

Gantz, J., & Reinsel, D. (2012). *THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East*. Framingham, Massachusetts.

Garba, A. B., Armarego, J., Murray, D., & Kenworthy, W. (2015). Review of the Information Security and Privacy Challenges in Bring Your Own Device (BYOD) Environments. *Journal of Information Privacy and Security*, *11*(1), 38–54. https://doi.org/10.1080/15536548.2015.1010985

Grossman, L. (2001, February). Welcome to the Snooper Bowl. *TIME*, *157*(6), 72.

Gupta, D., & Rani, R. (2019). A study of big data evolution and research challenges. *Journal of Information Science*, *45*(3), 322–340. https://doi.org/10.1177/0165551518789880

Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., & Erlich, Y. (2013). Identifying Personal Genomes by Surname Inference. *Science*, *339*(6117), 321–324. https://doi.org/10.1126/science.1229566

Hack, M. (2016). The implications of Apple's battle with the FBI. *Network Security*, *2016*(7), 8–10. https://doi.org/10.1016/S1353-4858(16)30068-X

Hayes, D. R., & Cappa, F. (2018). Open-source intelligence for risk assessment. *Business Horizons*, *61*(5), 689–697. https://doi.org/10.1016/j.bushor.2018.02.001

Hern, A. (2018, January 28). Fitness tracking app Strava gives away location of secret US army bases. *The Guardian*. Retrieved from https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases

Hill, K. (2020, January 18). The Secretive Company That Might End Privacy as We Know It. Retrieved from https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html

Howe, D. C. (2015). Surveillance Countermeasures: Expressive Privacy via Obfuscation. *A Peer-Reviewed Journal About*, *4*(1), 88–98. https://doi.org/10.7146/aprja.v4i1.116108

Innocence Project. (2020). DNA Exonerations in the United States. Retrieved from https://www.innocenceproject.org/dna-exonerations-in-the-united-states/

Introna, L. D., & Wood, D. (2004). Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems. *Surveillance & Society*, *2*(2/3), 177–198.

Jaffe, E. M. (2016). *Swatting: The New Cyberbullying Frontier after Elonis v. United States*. *Drake Law Review* (Vol. 64).

Jaffe, E. M. (2020). From Terrorists to Trolls: Expanding Web Host Liability for Live-Streaming, Swatting, and Cyberbullying. *Boston University Journal of Science and Technology Law*, *26*(2), 51–66.

Joh, E. E. (2016). The new surveillance discretion: Automated suspicion, big data, and policing. *Harvard Law & Policy Review*, *10*, 15–42.

Jones, M. L. (2016). *Ctrl + Z: The Right to Be Forgotten*. New York, New York: New York University Press.

Katal, A., Wazid, M., & Goudar, R. H. (2013). Big Data: Issues, Challenges, Tools and Good Practices. In *2013 Sixth International Conference on Contemporary Computing (IC3)* (pp. 404–409). IEEE. https://doi.org/10.1109/IC3.2013.6612229

Kelley, T. L. (1927). *Interpretation of Educational Measurements*. Yonkers-on-Hudson, NY: World Book Company.

Kepner, J., Gadepally, V., Michaleas, P., Schear, N., Varia, M., Yerukhimovich, A., & Cunningham, R. K. (2014). Computing on masked data: a high performance method for improving big data veracity. In *2014 IEEE High Performance Extreme Computing Conference (HPEC)* (pp. 1–6). IEEE. https://doi.org/10.1109/HPEC.2014.7040946

Kirk, M. (2014). *United States of Secrets*. United States: PBS. Retrieved from https://www.pbs.org/wgbh/frontline/film/united-states-of-secrets/

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data and Society*, *3*(1), 1–10. https://doi.org/10.1177/2053951716631130

Klein, A. (2017, September 28). The real problem with credit reports is the astounding number of errors. Retrieved from https://www.brookings.edu/research/the-real-problem-with-credit-reports-is-the-astounding-number-of-errors/

Kornstein, S. (2015, June 29). The Rise of Mobile Phones: 20 Years of Global Adoption. Retrieved from https://blog.cartesian.com/the-rise-of-mobile-phones-20-years-of-global-adoption

Krämer, J., Wiewiorra, L., & Weinhardt, C. (2013). Net neutrality: A progress report. *Telecommunications Policy*, *37*(9), 794–813. https://doi.org/10.1016/j.telpol.2012.08.005

Kuempel, A. (2016). The invisible middlemen: A critique and call for reform of the data broker industry. *Northwestern Journal of International Law and Business*, *36*(1), 207–234.

Kurwa, R. (2019). Building the Digitally Gated Community: The Case of Nextdoor. *Surveillance & Society*, *17*(1/2), 111–117. https://doi.org/10.24908/ss.v17i1/2.12927

Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. *META Group Research Note*, *6*(70).

Lashmar, P. (2017). No More Sources?: The impact of Snowden's revelations on journalists and their confidential sources. *Journalism Practice*, *11*(6), 665–688. https://doi.org/10.1080/17512786.2016.1179587

Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: a methodology for information quality assessment. *Information & Management*, *40*(2), 133–146. https://doi.org/10.1016/S0378-7206(02)00043-5

Lewis, M. (2019). *Don't F**k With Cats: Hunting An Internet Killer*. United States: Netflix.

Li, L. B. (2018). Data Privacy in the Cyber Age: Recommendations for Regulating Doxing and Swatting. *Federal Communications Law Journal*, *70*, 317.

Lowry, P. B., Dinev, T., & Willison, R. (2017). Why security and privacy research lies at the centre of the information systems (IS) artefact: proposing a bold research agenda. *European Journal of Information Systems*, *26*(6), 546–563. https://doi.org/10.1057/s41303-017-0066-x

Lucas, G. (2005). *Star Wars: Episode III - Revenge of the Sith*. United States: Twentieth Century Fox.

Lum, C., Hibdon, J., Cave, B., Koper, C. S., & Merola, L. (2011). License plate reader (LPR) police patrols in crime hot spots: an experimental evaluation in two adjacent jurisdictions. *Journal of Experimental Criminology*, *7*(4), 321–345. https://doi.org/10.1007/s11292-011-9133-9

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, *15*(4), 336–355. https://doi.org/10.1287/isre.1040.0032

Masden, C. A., Grevet, C., Grinter, R. E., Gilbert, E., & Edwards, W. K. (2014). Tensions in Scaling-up Community Social Media: A Multi-Neighborhood Study of Nextdoor. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14* (pp. 3239–3248). New York, New York, USA: ACM Press. https://doi.org/10.1145/2556288.2557319

Mason, R. O. (1986). Four ethical issues of the information age. *MIS Quarterly*, *10*(1), 5–12.

McDonald, A. M., & Cranor, L. F. (2008). The cost of reading privacy policies. *I/S: A Journal of Law and Policy for* *the Information Society*, *4*(3), 543–568.

McLeod, A., Savage, A., & Simkin, M. G. (2018). The Ethics of Predatory Journals. *Journal of Business Ethics*, *153*(1), 121–131. https://doi.org/10.1007/s10551-016-3419-9

Medine, D., Brand, R., Cook, E. C., Dempsey, J., & Wald, P. (2014a). *Report on the Surveillance Program Operated Pursuant to Section 702 of the Foreign Intelligence Surveillance Act*. Washington, D.C. Retrieved from https://www.pclob.gov/library/702-Report.pdf

Medine, D., Brand, R., Cook, E. C., Dempsey, J., & Wald, P. (2014b). *Report on the Telephone Records Program Conducted under Section 215 of the USA PATRIOT and on the Operations of the Foreign Intelligence Surveillance Court*. Washington, D.C.

Millar, S. A., & Marshall, T. P. (2019, April 24). State Data Breach Notification Laws – Overview of Requirements for Responding to a Data Breach – Updated April 2019. Retrieved from https://www.natlawreview.com/article/state-data-breach-notification-laws-overview-requirements-responding-to-data-3

Milne, G. R., Culnan, M. J., & Greene, H. (2006). A Longitudinal Assessment of Online Privacy Notice Readability. *Journal of Public Policy & Marketing*, *25*(2), 238–249. https://doi.org/10.1509/jppm.25.2.238

Mühle, A., Grüner, A., Gayvoronskaya, T., & Meinel, C. (2018). A survey on essential components of a self-sovereign identity. *Computer Science Review*, *30*, 80–86. https://doi.org/10.1016/j.cosrev.2018.10.002

Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. *Proceedings - IEEE Symposium on Security and Privacy*, 111–125. https://doi.org/10.1109/SP.2008.33

National Law Enforcement and Corrections Technology Center. (2010). *The Results Are In: Automatic License Plate Reader Technology Leads to Success*. Research Triangle Park, NC.

Newkirk, D. (2018). “Apple: Good Business, Poor Citizen”: A Practitioner’s Response. *Journal of Business Ethics*, *151*(1), 13–16. https://doi.org/10.1007/s10551-016-3397-y

Nextdoor. (2019, December 5). Privacy Policy. Retrieved from https://legal.nextdoor.com/us-privacy-policy-2020/

Nickelodeon Consumer Privacy Litigation, 827 F.3d 262, 293 (3rd Cir., 2016)

Nicks, D. (2016, May 24). Mark Zuckerberg Bought Four Houses Just to Tear Them Down. Retrieved from https://money.com/mark-zuckerberg-houses/

Norman, J. (2016). Taking the Sting Out of the Stingray: The Dangers of Cell-Site Simulator Use and the Role of the Federal Communications Commission in Protecting Privacy & Security. *Federal Communications Law Journal*, *68*(1), 139-0_8.

Noyes, K. (2015, June 25). Scott McNealy on privacy: You still don’t have any. Retrieved from https://www.pcworld.com/article/2941052/scott-mcnealy-on-privacy-you-still-dont-have-any.html

O’Neil, C. (2016). *Weapons of Math Destruction*. New York, New York: Crown.

O’Neill, K. (2019, January 15). Facebook’s “10 Year Challenge” Is Just a Harmless Meme—Right? Retrieved from https://www.wired.com/story/facebook-10-year-meme-challenge/

Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review*, *57*(6), 1701–1777.

Otto, P. N., Anton, A. I., & Baumer, D. L. (2007). The ChoicePoint Dilemma: How Data Brokers Should Handle the Privacy of Personal Information. *IEEE Security & Privacy Magazine*, *5*(5), 15–23. https://doi.org/10.1109/msp.2007.126

Ozer, M. M. (2010). *Assessing the Effectiveness of the Cincinnati Police Department’s Automatic License Plate Reader System within the Framework of Intelligence-Led Policing and Crime Prevention Theory*. University of Cincinnati.

Pavlou, P. A. (2011). State of the information privacy literature: where are we now and where should we go. *MIS Quarterly*, *35*(4), 977–988.

Pell, S. K., & Soghoian, C. (2014). Your secret stingray’s no secret anymore: The vanishing government monopoly over cell phone surveillance and its impact on national security and consumer privacy. *Harvard Journal of Law & Technology*, *28*(1), 1–76.

Peslak, A. R. (2006). PAPA revisited: A current empirical study of the Mason framework. *Journal of Computer Information Systems*, *46*(3), 117–123. https://doi.org/10.1080/08874417.2006.11645905

Peterson, A., Yahr, E., & Warrick, J. (2014, September 1). Leaks of nude celebrity photos raise concerns about security of the cloud. Retrieved from https://www.washingtonpost.com/politics/leaks-of-nude-celebrity-photos-raise-concerns-about-security-of-the-cloud/2014/09/01/59dcd37e-3219-11e4-8f02-03c644b2d7d0_story.html

Power, D. J. (2016). “Big Brother” can watch us. *Journal of Decision Systems*, *25*(sup1), 578–588. https://doi.org/10.1080/12460125.2016.1187420

President’s Council of Advisors on Science and Technology. (2014). *Big Data and Privacy: A Technological Perspective*. Washington, D.C.

Privacy Rights Clearninghouse. (2014, October 1). Brind Your Own Device (BYOD)…at Your Own Risk. Retrieved from https://privacyrights.org/consumer-guides/bring-your-own-device-byod-your-own-risk

Rahimian, R., & Kelly, A. (2019, December 15). The Decade Tech Lost Its Way. Retrieved from https://www.nytimes.com/interactive/2019/12/15/technology/decade-in-tech.html

Reding, V. (2012). The EU Data Protection Reform 2012 : Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age. In *Innovation Conference Digital, Life, Design* (pp. 1–6). Munich, Germany. Retrieved from http://europa.eu/rapid/press-release_SPEECH-12-26_en.htm

Richards, N. M. (2012). The Dangers of Surveillance. *Harvard Law Review*, *126*(7), 1934–1965.

Richards, N. M., & King, J. H. (2014). Big Data Ethics. *Wake Forest Law Review*, *49*(2), 393–432.

Riederer, C., Kim, Y., Chaintreau, A., Korula, N., & Lattanzi, S. (2016). Linking Users Across Domains with Location Data. In *Proceedings of the 25th International Conference on World Wide Web - WWW '16* (pp. 707–719). New York, New York, USA: ACM Press. https://doi.org/10.1145/2872427.2883002

Ring under fire over weakness in video device security. (2020). *Network Security*, *2020*(1), 1–2. https://doi.org/10.1016/S1353-4858(20)30001-5

Rosen, J. (2012). The Right to be Forgotten. *Stanford Law Review*, *64*, 88–92.

Rush, B. L. Data Accountability and Trust Act (2019). Washington, D.C.: U.S. Congress.

Sanchez-Garrido, B. (2016). Social media's criminal element. *Risk Management*, *63*(1), 8–10.

Schneier, B. (2009, December 9). My Reaction to Eric Schmidt. Retrieved from https://www.schneier.com/blog/archives/2009/12/my_reaction_to.html

Shim, J. P., Mittleman, D., Welke, R., French, A. M., & Guo, J. C. (2013). Bring Your Own Device (BYOD): Current status, issues, and future directions. In *19th Americas Conference on Information Systems, AMCIS 2013 - Hyperconnected World: Anything, Anywhere, Anytime* (Vol. 1, pp. 595–596).

Slipke, D. (2015). Court document reveals more about Sentinel, OK, bomb threat. Retrieved from https://oklahoman.com/article/5386857/court-document-reveals-more-about-sentinel-ok-bomb-threat

Smith, H. J., Dinev, T., & Xu, H. (2011). Information Privacy Research: An Interdisciplinary Review. *MIS Quarterly*, *35*(4), 989–1015. https://doi.org/10.2307/41409970

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information Privacy: Measuring Individuals' Concerns about Organizational Practices. *MIS Quarterly*, *20*(2), 167–196. https://doi.org/10.2307/249477

Smolan, R., & Erwitt, J. (2012). *The Human Face of Big Data*. Sausalito, California: Against All Odds Productions.

Snow, J. (2018). Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots. Retrieved May 23, 2019, from https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28

Solove, D. (2015, November 13). The Growing Problems with the Sectoral Approach to Privacy Law. Retrieved from https://teachprivacy.com/problems-sectoral-approach-privacy-law

Solove, D. J. (2007). "I've Got Nothing to Hide" and Other Misunderstandings of Privacy. *San Diego Law Review*, *44*(May), 1–23. https://doi.org/10.2139/ssrn.998565

Sprenger, P. (1999, January 26). Sun on Privacy: "Get Over It." Retrieved from https://www.wired.com/1999/01/sun-on-privacy-get-over-it/

Stewart, K. a., & Segars, A. H. (2002). An Empirical Examination of the Concern for Information Privacy Instrument. *Information Systems Research*, *13*(1), 36–49. https://doi.org/10.1287/isre.13.1.36.97

Stoycheff, E. (2016). Under surveillance: Examining Facebook's spiral of silence effects in the wake of NSA internet monitoring. *Journalism & Mass Communication Quarterly*, *93*(2), 296–311. https://doi.org/10.1177/1077699016630255

Streitfeld, D. (2013, March 12). Google Concedes That Drive-By Prying Violated Privacy. Retrieved from https://www.nytimes.com/2013/03/13/technology/google-pays-fine-over-street-view-privacy-breach.html

Sweeney, L. (2000). *Simple demographics often identify people uniquely* (Data Privacy No. 3). Pittsburg, PA. Retrieved from http://dataprivacylab.org/projects/identifiability/paper1.pdf

Thompson, R. M. (2014). *The Fourth Amendment Third-Party Doctrine*. Washington, D.C.

Tobin, A., & Reed, D. (2017). *The Inevitable Rise of Self-Sovereign Identity: A white paper from the Sovrin Foundation*. Retrieved from https://sovrin.org/wp-content/uploads/2017/06/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf

Tschider, C. A. (2015). Experimenting with Privacy: Driving Efficiency Through a State-Informed Federal Data Breach Notification and Data Protection Law. *Tulane Journal of Technology and Intellectual Property*, *18*, 45–81.

U.S. Privacy Protection Study Commission. (1977). *Personal Privacy in an Information Society*.

Vaas, L. (2013, March 17). Hackers launch DDoS attack on security blogger's site, send SWAT team to his home. Retrieved from https://nakedsecurity.sophos.com/2013/03/17/swat-ddos-brian-krebs/

Vaas, L. (2016, July 15). Serial swatter, stalker and doxer Mir Islam given 2 years prison. Retrieved from https://nakedsecurity.sophos.com/2016/07/15/serial-swatter-stalker-and-doxer-mir-islam-given-2-years-prison/

Van Rijmenam, M. (2014). *Think Bigger: Developing a Successful Big Data Strategy for Your Business*. New York, New York: AMACOM.

Walker, R. K. (2012). The Right to Be Forgotten. *Hastings Law Journal*, *64*(2), 257–286.

Warren, S. D., & Brandeis, L. D. (1890). The Right to Privacy. *Harvard Law Review*, *4*(5), 193–220.

West, E. (2019). Amazon: Surveillance as a service. *Surveillance and Society*, *17*(1–2), 27–33. https://doi.org/10.24908/ss.v17i1/2.13008

Whang, S. E. (2012). *Data Analytics: Integration and Privacy*. Stanford University.

Wheelan, C. (2013). *Naked Statistics: Stripping the Dread from the Data*. New York, New York: WW Norton & Company.

Zmud, J., Wagner, J., Moran, M., & George, J. P. (2016). *License Plate Reader Technology: Transportation Uses and Privacy Risks*. College Station, Texas. Retrieved from https://scholarship.law.tamu.edu/facscholar%0Ahttps://scholarship.law.tamu.edu/facscholar/923

## Author Biographies

**Jacob A. Young** is an assistant professor of management information systems in the Foster College of Business and the Director of the Center for Cybersecurity at Bradley University. He earned his D.B.A. in Computer Information Systems from Louisiana Tech University. Dr. Young conducts research on privacy, security, and anonymity issues related to information systems with a primary focus on anonymous whistleblowing systems. He serves as the Senior Advisor on Cybersecurity at the National Whistleblower Center in Washington, D.C. His work has been published in *AIS Transactions on Human-Computer Interaction*, *Communications of the Association for Information Systems*, *Journal of Information Security Education*, the *Journal of the Midwest Association for Information Systems*, the *DePaul Business & Commercial Law Journal*, and other journals and conference proceedings.

**Tyler J. Smith** is an assistant professor of business law in the Foster College of Business at Bradley University. He earned his J.D. from Indiana University Robert H. McKinney School of Law and his LL.M. from Notre Dame Law School. He conducts legal research on information privacy, constitutional law, and labor-management relations. His work has been published in several journals, such as the *Fordham International Law Journal*, the *New York International Law Review*, and the *Indiana International & Comparative Law Review*.

**Haoran (Shawn) Zheng** is an assistant professor of management information systems in the Foster College of Business at Bradley University. He earned his Ph.D. in Information Systems and Business Analytics from Chapman Graduate School at Florida International University. Dr. Zheng's conducts research in adoption, integration, and assimilation of e-health systems, organizational changes and modern privacy issues related to big data and analytics.

This page intentionally left blank