Technology Architectures in Regulatory Environments: The Case of Big Data Applications in the Banking Services Sector

Erik Paolo Capistrano, PhD.
Associate Professor
Cesar E.A. Virata School of Business
University of the Philippines Diliman, Quezon City, Philippines

Abstract

This research is dedicated to analyzing some external issues in the banking services sector's efforts to employ Big Data analytics. Majority of previous studies involving the phenomenal rise of Big Data analytics have focused on determining its feasibility and potential value-adding impacts on firms, but leaving a dearth of discussion in the context of heavily regulated and controversial environments, such as the banking services sector. While there are obvious benefits of Big Data towards this sector's efforts to consolidate client and industry data, predict risk and investment trends, and design and cross-sell products, issues on privacy and security, controversial industry monitoring, and rigid government oversight challenge current views on the potential of Big Data. Through an extensive examination of the proposed Big Data architectures from past research vis-à-vis existing regulatory environments within the banking services sector, this research proposes a framework in an attempt to bridge these two entities together, identifying certain strategic, tactical, and operational advantages and pitfalls of adopting Big Data analytics to enhance financial information systems. It is expected that this framework can shed light on some clouded issues regarding issues on the fit between Big Data technologies and industry regulations to aid managers on how to better integrate Big Data into their product and service offerings.

Keywords: Big data, financial information systems, financial services, portfolio management, customer relationship management, security

1. Introduction

1.1 Rationale

The Information Age has essentially become easier to exploit over the last few years. Several new technologies have made data and information exchanges easier and more efficient, spurring an increasing emphasis on big data, business analytics, and smart living and work environments (George, Haas, & Pentland, 2014). This has driven a significant increase in data collection activities (Chang, Kauffman, & Kwon, 2014), both in scale and scope and in the desire for more data-driven recommendations, solutions, and decisions (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). Through Web browsing, social media, Internet of Things (Hashem et al., 2015; He, Wu, Yan, Akula, & Shen, 2015; Kemp, 2014; Mansour, 2016; Tambe, 2014), crowdsourcing (Kim, Trimi, & Chung, 2014), cloud (Kemp, 2014), and mobile (H. Chen, Chiang, & Storey, 2012; Kemp, 2014) technologies, data has grown more ubiquitous and cheaper. The resulting scenario is that organizations now have more data than they could handle (Chang et al., 2014). Recent reports from the United States have shown that big data use has resulted in a 5-6% increase in productivity (Barton & Court, 2012) and as much as 10% in business growth (Akter & Wamba, 2016), hereby requiring 140,000 to 190,000 more workers with the appropriate analytical expertise and 1.5 million more data-literate managers (Tene & Polonetsky, 2012). In fact, some 4.4 million jobs were created due to the demand for big data technologies alone (Forest, Foo, Rose, & Berenzon, 2014). Some specific industries and applications have also yielded significant business results. Medical professionals, working with big data firms, are continuously experimenting with prediction models for occurrences of certain diseases vis-à-vis visits to physicians (Lazer, Kennedy, King, & Vespignani, 2014). Mainstream adoption is expected to boost global manufacturing and retail by US$325 billion (EY, 2014). Big data-driven pricing monitoring systems resulted in as much as 35% increases in online sales while big data-powered anti-fraud detection systems enabled credit card firms to save US$2 billion in potential losses annually (Akter & Wamba, 2016). This, coupled with the digitization of economies, has spurred the demand for better and more efficient data analytics, with the market value of business data being estimated at US$25 billion (Kemp, 2014), growing to US$53.4 billion by 2017 (Connors, Courbe, & Waishampayan, 2013; Gozman, Currie, & Seddon, 2015) and an estimated overall economic contribution of US$10-$15 trillion through 2030 (Kiron, Ferguson, & Prentice, 2013). Additionally, the software market for big data is estimated at US$16 billion, growing by 8% per year (Kemp, 2014). Digitization has enabled both firms and consumers to seek out and access a wider array of information pertaining to all sorts of product and service information, from production to distribution (Weill & Woerner, 2015). This has made data and information primary assets for many organizations, leading to increased efforts to collect and process them as much as possible (Demirkan & Delen, 2013), mainly because of the significant economic potential that can be derived from diverse sets of data (Ebner, Bühnen, & Urbach, 2014). Investments in big data were recorded at US$30 billion, and are expected to rise to US$76 billion by 2020 (Forest et al., 2014). However, despite all of the market and economic valuations estimated for big data and its applications, its real value has yet to be fully realized (Chang et al., 2014). Current studies looking at it from a value chain perspective are having a difficult time even estimating its real value (EY, 2014).

Big data has permeated almost every industry (Bedeley & Iyer, 2014), creating a new synthesis of real-time information and communication exchanges, and creating calls for a tighter and more consistent fit of functional areas, business objectives, and big data opportunities (Oracle, 2015; Turner, Schroeck, & Shockley, 2013). And currently, this has made business organizations more conscious about their efforts to acquire and process data and information using more sophisticated techniques and tools, and to produce higher quality outputs from such extensive and sophisticated processes. With this, big data has been argued to be essential to produce productivity and evolutionary breakthroughs (C. P. Chen & Zhang, 2014), enabling ecosystems to be fine-tuned for better and continuous flows of collaborative and informed knowledge of consumer behaviors (Weill & Woerner, 2015), aimed at finding innovative ways to differentiate (Demirkan & Delen, 2013). This is especially observable in high-impact environments such as e-commerce, market intelligence, e-government, healthcare, security (H. Chen et al., 2012; Demirkan & Delen, 2013; Frizzo-Barker, Chow-White, Mozafari, & Ha, 2016; Kim et al., 2014), traffic management (Tene & Polonetsky, 2012), services innovation and marketing (Akter & Wamba, 2016), personal healthcare and wellness (Crawford & Schultz, 2014), mechanics and engineering (George et al., 2014), and finance and banking (Bennett, 2013; Hu, Zhao, Hua, & Wong, 2012; Kemp, 2014; Kshetri, 2016; Madche, 2015; Srivastava & Gopalkrishnan, 2015). All of these high-impact environments have experienced improved efficiencies in managerial and market transactions costs and time-related costs (Akter & Wamba, 2016). And at the moment, there is no indication of it slowing down, simply because these high-impact environments not only can interact with one another, but their activities also spill over to other environments, creating various cascading effects and ripple effects through other aspects of business and society.

This simply means that data and information management practices have become a critical source of competitive advantage. While data and information can be easily acquired, the battleground for innovation and differentiation to attract more customers and attain greater market shares is defined by how well firms can make use of such acquisitions. Continuously managing, developing, and updating in an almost real-time pace the necessary and appropriate analytical and predictive models is growing to be an important component in making and executing strategies and tactics. In turn, strategy and tactics implementation has turned into a dynamic, rather than a static, action. This greatly and significantly addresses calls for more proactive actions given how quickly consumer behaviors and attitudes towards products and services change. Putting it in simpler terms, with so much raw data freely provided and exchanged, it is now a question of how this raw data can be effectively and efficiently harnessed for competitive organizational purposes.

However, despite its highly touted benefits, numerous technical and social issues still hound its full implementation. These include decisions on what pieces of data to combine from what sources (LaValle et al., 2011), difficulties in data capture, storage, analysis and visualization (C. P. Chen & Zhang, 2014), overwhelming scalability of data acquisition (Demirkan & Delen, 2013; Weill & Woerner, 2015) and processing vis-à-vis resource availability (H. Chen et al., 2012; Hashem et al., 2015), constantly growing and evolving pressures to develop up-to-date analytical (H. Chen et al., 2012; Lazer et al., 2014; Tambe, 2014) and managerial (LaValle et al., 2011) tools and techniques and human capabilities (Frizzo-Barker et al., 2016; Kim et al., 2014; Kshetri, 2016), concerns regarding standardizations on format and logical data models (Bennett, 2013; Kemp, 2014), and issues pertaining to its actual pervasiveness (Boyd & Crawford, 2012; H. Chen et al., 2012; Hu et al., 2012; Srivastava & Gopalkrishnan, 2015) which leads to risks of dire loss of privacy (Akter & Wamba, 2016; Boyd & Crawford, 2012; Che, Safran, & Peng, 2013; Frizzo-Barker et al., 2016; Katal, Wazid, & Goudar, 2013; Mansour, 2016). It must be noted that these problems and issues are both independent and interacting with each other, further complicating the current efforts to develop research on this topic.

1.2 Research Gaps and Objectives

It is glaringly obvious that big data contributes tremendously and favorably. However, because it is a disruptive piece of technology, many aspects regarding the pace of its development and adoption have been of issue, especially in the social and legal aspects of its use. The chain of events leading to its current levels of popularity, spreading through different business and social contexts, has also stimulated a wide range of research agendas over the last few years. Bulk of the current stream of research focused on the technical aspects, evaluating current levels of technologies and human skill sets vis-à-vis big data applications' requirements. There is no question of its potential regarding its technological and technical prowess. It is more of a question of the other aspects that influence its actual usability and feasibility that are currently in question.

Big data has generated a wave of excitement and interest that significantly challenges present thinking and perspectives about organizational and institutional structures and practices (Fuschi & Tvaronavičienė, 2011; LaValle et al., 2011), data infrastructures, business intelligence and analytics, and information technology strategies (Akter & Wamba, 2016; Ebner et al., 2014; Frizzo-Barker et al., 2016). There is also a growing interest in developing a means to standardize the approach towards such technologies (Kemp, 2014), given how widely varied and dispersed present requirements (Ebner et al., 2014) and practices (Bedeley & Iyer, 2014) are. This, coupled with that growing trends in information technology-based services and service-based information technologies that make extensive use of the big data phenomenon, also presents the challenge to rethink the roles and contributions information technologies and information systems from new organizational viewpoints (Demirkan & Delen, 2013).

Previous researches (C. P. Chen & Zhang, 2014; H. Chen et al., 2012; Hashem et al., 2015; Kim et al., 2014; Tambe, 2014) have agreed that the demand for appropriate data analysis tools and techniques has significantly outpaced the current and existing technological capabilities and technical know-how. In fact, there is a risk where the thinking has been directed to big data practices replacing traditional means, which is definitely not the case (Connors et al., 2013; Lazer et al., 2014). This has resulted in several gaps and challenges barring the exploitation of big data's full potential. Because of its near-ubiquity and wide coverage, the examination and discussion on big data can take many different points of view (Bedeley & Iyer, 2014). Unfortunately, most concerns about big data fall under the technological and technical realms of consideration (Connors et al., 2013). While most studies have focused on developing architectures and strategies to exploit big data, one aspect that is currently overlooked, and therefore understudied, is the regulatory environment. This is evidenced by the lack of even a basic agreement on the acceptable definition of what big data is (Akter & Wamba, 2016; Gandomi & Haider, 2015; Kemp, 2014), what it involves (Boyd & Crawford, 2012), and the standards in which big data is supposed to be used, especially when its value-adding contributions must be framed within certain contexts governed by language, laws, measures, and models (Demirkan & Delen, 2013). As an example, existing regulatory frameworks on privacy and the use of personal and personally identifying information appear incapable of keeping pace with big data-driven business norms and practices (Crawford & Schultz, 2014). In other words, it is a question involving if there are any big data capabilities that can potentially encroach on some legal or ethical concerns.

With such, this research focuses on the dearth of research and discussion towards the regulatory environment that governs big data processes, techniques, and technologies. This gap is caused by several overarching issues inherent to big data, such as its obvious potential for extreme intrusion and pervasiveness (H. Chen et al., 2012; Srivastava & Gopalkrishnan, 2015; Weill & Woerner, 2015) and the fact that there is no one-size-fits-all solution (C. P. Chen & Zhang, 2014). Additionally, the fact that currently, big data is considered to be a disruptive piece of technology that is developing at a significantly rapid pace, it is also therefore subject to similar issues such as privacy, security, and surveillance concerns, disruptions of conventional methods, labor practices, and legal approaches, and questions regarding on how to best utilize it (Frizzo-Barker et al., 2016). This further reinforces the need to discuss the development of the appropriate regulatory framework in which big data can safely and appropriately operate. Furthermore, an examination of the regulatory environment is important for several reasons. One is that it is still a major challenge, both for researchers and practitioners, to design the appropriate platform for proper data management (Hashem et al., 2015), which includes issues such as finding a consensus for standardizing and reconciling different types of data and information acquired from different sources (Bennett, 2013), without falling foul of legislation in intellectual property, property and data privacy, and other pertinent laws (Gozman et al., 2015). But as a cautionary note, it is presently difficult enough to comply with and enforce existing regulations, and the current framework may well be unmanageable if it extends to every piece of information (Tene & Polonetsky, 2012) that big data technologies can potentially touch, acquire, process, and analyze.

Specific for this research is the application of big data technologies in finance and banking activities (Hu et al., 2012), highlighting the fact that most of these activities are subject to government interventions through heavy regulations (Bholat, 2015; Crawford & Schultz, 2014; EY, 2014; Gozman et al., 2015; IFC, 2015; Kshetri, 2016; Munar, Chiner, & Sales, 2014; Tene & Polonetsky, 2012). Therefore, this research poses the following objectives for consideration:

1) To explore the possible components and capabilities of big data that are appropriate to meet specific finance and banking needs and requirements.
2) To identify possible gaps between the available big data technologies and the finance and banking business processes.
3) To determine if the available big data technologies vis-à-vis finance and banking practices adhere to the appropriate governance and regulatory environments.

The rest of this research is structured as follows: Section 2 presents the literature review, covering what has been discussed so far with regards to big data technologies. The first part presents a broad picture of what big data is and its applications in some high impact areas. The second part narrows down the scope to big data applications in the finance and banking industry, illustrating some examples and highlighting the ethical and regulatory issues coupled with these applications. The third part explores the possible legal remedies and their requirements to apply these remedies to this issue. Lastly, some theoretical discussion to further explain the ethical and regulatory issues and to support the possible legal remedies are presented. Section 3 presents the integration of the issues and remedies as discussed in the previous section, culminating in a proposed research framework that would serve as a starting point for both academics and practitioners to further delve into the regulatory issues surrounding big data use. Section 4 concludes this research, giving some directions for future efforts.

2. Literature Review

2.1 Big data principles

Big data is all about how to use the copious, and oftentimes cumbersome, amounts of data to generate meaningful information in a timely manner, not just to analyze and report on historical data, but to also predict and respond to changing needs and opportunities with as much timeliness, reliability, and accuracy as possible (Connors et al., 2013). Big data started as the management of terabytes, petabytes, and even exabytes of data (Hu et al., 2012; Kim et al., 2014), which are obviously bigger than any specific data format currently in existence (Bennett, 2013), done frequently to improve a firm's relationship with its key stakeholders (Kiron et al., 2013). To put it in some perspective, it is estimated that there are 2.5 quintillion bytes of data produced daily (Gozman et al., 2015). This makes for data with far more dimensions and far higher complexities (Chang et al., 2014). But importantly, its analytics capabilities feature highly reliable predictive and recommendation functions which are heavily based on a complex combination of statistics, descriptive and predictive data, and simulation and optimization models (Coppolino, D'Antonio, Romano, Campanile, & de Carvalho, 2015; Gandomi & Haider, 2015). This is based on customer and public data and information which, when combined with an organization's proprietary data and information, can create networks of extensive and effective analyses (Frizzo-Barker et al., 2016; Kiron et al., 2013), which can also be context-specific (Demirkan & Delen, 2013). Exchanges in data sources, index providers, data renderers, and data users are ideally held together by a structure governed by limitations on risks around data use (Kemp, 2014) and the context governed by the decision making environment (Işık, Jones, & Sidorova, 2013; LaValle et al., 2011).

Big data technologies have specifically targeted business processes, specifically the processes involving organizations' data management practices (Davenport, 2006). The principles of designing big data systems require a good theoretical framework and practical architecture to develop sound strategies and tactics (Kim et al., 2014) and to support a variety of analytical methods (C. P. Chen & Zhang, 2014). Ideally, big data should cover the spectrum of management, from the strategic to the day-to-day operations of an organization (LaValle et al., 2011). Big data strategies and tactics revolve around its main characteristics: volume, velocity, variety, variability, and complexity (Akter & Wamba, 2016; Bholat, 2015; Chang et al., 2014; Işık et al., 2013; Kemp, 2014; Kshetri, 2016), and even value (Gandomi & Haider, 2015; Katal et al., 2013), all interacting together to produce significantly better metrics for performance monitoring, and more importantly, generate significantly better prediction models (Frizzo-Barker et al., 2016), rather than merely reporting and giving feedback on queries (Coppolino et al., 2015). This further requires tools and techniques that enable more virtual collaborations that are highly adaptive, synchronous, and agile (Demirkan & Delen, 2013), especially to cope with the exponential growth of available data.

2.2 Big data capabilities and general applications

Big data has numerous capabilities. Historically, analytics were focused on integrating and creating greater value from data on an organizations' internal business operations, acquiring data from its sales and marketing activities, production inputs and outputs, human resource recruitments and developments, and resource allocation and distribution decisions (Davenport, 2006). Over the years, this practice has involved external sources, looking at consumer and customer behaviors for additional data to support internally acquired data. And therefore, more sophisticated and complex capabilities were developed. These include techniques on optimization methods, statistics, data mining, machine learning, visualization approaches, text analytics, video analytics, audio analytics, social network analysis (C. P. Chen & Zhang, 2014; Gandomi & Haider, 2015), and capabilities that make for easier access (Hu et al., 2012) and analysis (Tambe, 2014) to diverse sets of data, maximizing every bit of data and information to extract some value (Kiron et al., 2013). This enables for more contextualized (Demirkan & Delen, 2013) and more detailed micro-segmentation of the market and more accurate customer life events analysis (Srivastava & Gopalkrishnan, 2015). Furthermore, the market intelligence gathered can be further subdivided into two (He et al., 2015; Işık et al., 2013): internal business intelligence that extract meaningful data from unstructured textual data that include opinions, sentiments, emotions, and subjectivities; and external competitive intelligence that include further data mining into texts and visuals.

Big data can also be sourced from different online and offline channels (George et al., 2014). In the online world, the significantly amount of rich but raw structured and unstructured data from social media, the Internet of Things (Boyd & Crawford, 2012; Ebner et al., 2014; Hashem et al., 2015; Lazer et al., 2014; Mansour, 2016), and mobile technologies (H. Chen et al., 2012) alone is worth considering. Crowdsourcing (Kim et al., 2014) and cloud computing (Kemp, 2014) further adds to the very rich pool of data, all ready and waiting to be appropriately harnessed and used.

Because of these powerful capabilities, big data can competently address some high-impact areas (Chang et al., 2014). The first is in business, in terms of employing geo-location-based customer demand surveying and profiling, product bundling and cross-selling, and sensing market sentiments within social media platforms. The second is in consumer behavior, in terms of creating behavior-based segmentation, tracing “friends of friends” networks within social media platforms, and facilitating churn management. The third is in social issues, in terms of cataloging social, political, and electoral sentiments, providing support for education and health care concerns, and tracing human network communication patterns. In effect, these technologies provide a very significant opportunity to enhance organizational flexibility and shared risks and responsibilities, made possible through shared platforms and databases established via well-defined technology architectures and standards (Işık et al., 2013; LaValle et al., 2011).

The discipline of analytics alone has made its mark over the last two decades, demonstrating how far and how deep data can be used to derive various significantly value-adding insights contributing to organizational decision making. The key to making analytics favorably work is for organizations to reliably and even commonsensically integrate several pieces of data from different sources via different channels, instead of treating data as disparate pieces of resources. This requires a breakaway attitude towards viewing the opportunities that data can provide (Davenport, 2006). The widespread use of modeling and data optimization, immensely supported by senior management advocates, has made all of this possible.

In its relatively young history, big data applications has been observed in many situations, such as in e-commerce platforms like Amazon and eBay for their highly scalable sales platforms and recommendation systems, in search engines like Google for web analytics and cloud computing capabilities (Akter & Wamba, 2016; H. Chen et al., 2012), in market research for analyzing customer personal information to deliver more accurate product and service recommendations (Crawford & Schultz, 2014; Tambe, 2014), in banking for building transaction networks, monitoring interbank and intermarket activities (Hu et al., 2012), and better assessments of creditworthiness and risk (Kemp, 2014; Kshetri, 2016), for example. These applications enable firms to address varied customer needs and behaviors by being able to create innovative products and services more responsively (Kim et al., 2014; Weill & Woerner, 2015). Other applications of big data are also found in medical services for establishing and predicting patterns of disease occurrences (Lazer et al., 2014) and more accurate diagnostics and treatment prescriptions, in insurance firms for better risk assessments and premiums valuations, and in public security for more reliable threat assessments (Demirkan & Delen, 2013).

2.3 Big data and the finance and banking sector

Within business disciplines, finance and budget management has been the top performer in terms of big data use (LaValle et al., 2011). The finance and banking sector is one of the bigger and more obvious targets for big data applications, getting significantly ahead of other industries in terms of its digitization efforts (Madche, 2015). Big data has permeated into the industry's various processes, including banking applications, financial markets, and network infrastructures (Fuschi & Tvaronavičienė, 2011). This has been made more prominent by the introduction of electronic transfers, effectively replacing cash-based transactions, and hereby reducing the limitations of spatial and temporal constraints brought about by the requirement of facilitating transactions by cash. Retail banking has focused on aggregating all available information about a single customer while corporate banking is more focused on coming economic data with credit and transactions information and consumer behaviors on bank patronage (Connors et al., 2013; Forest et al., 2014). There are vast and complex amounts of data and information (Bedeley & Iyer, 2014; Gozman et al., 2015) and flows (Munar et al., 2014) existing on a daily basis on stocks and bonds exchanges, personal and corporate banking transactions, and online and onsite purchases, to name a few. On top of these daily transactions, such data and information can be acquired through customer service records, real-time social media feeds and posts, and even correspondences (Barton & Court, 2012; He et al., 2015), providing for a very rich, yet unstructured, pool (Munar et al., 2014). Banks and other financial institutions can take advantage of this volume, velocity, variety, variability, complexity and value for better creation and management of inventories, customer behaviors, and market patterns (Bedeley & Iyer, 2014). This is further multiplied by the significant volumes and varieties of bank accounts, ranging from small individuals to multi-million dollar corporations maintaining several accounts for different purposes.

With no physical products to manufacture (Turner et al., 2013), there has been some initiatives to examine how big data technologies and capabilities can be adopted in the finance and banking sector. Some applications include analyzing transaction networks involving transfers, payments, monitoring interbank and intermarket activities, developing accounts and portfolio databases (Hu et al., 2012), creating better risk models for wider access to financial products and services (Kemp, 2014; Kshetri, 2016), creating anti-fraud monitoring and detection systems (Akter & Wamba, 2016; Coppolino et al., 2015), and employing more accurate sentiment analysis tools for better product cross-selling, lead and referral management, sales forecasting (Srivastava & Gopalkrishnan, 2015), and personalized financial advising (Madche, 2015). Because of these, the resulting benefits derived from such applications include easier assessments of creditworthiness and risk profiles, resulting to increases in the number of potential borrowers at reduced transactions costs (Kemp, 2014; Kshetri, 2016). Furthermore, banks can also have better management of their regulatory compliances, reputational risks, and financial crime prevention (Srivastava & Gopalkrishnan, 2015). This is typically done by a technological integration of customer touch points such as ATMs, customer care centers, bank branches, credit card facilities, and mortgage units; and financial and business forecast channels, such as industry data, securities trading data, analyst reports, and even social media notifications (Connors et al., 2013; EY, 2014; Oracle, 2015). This integration results in an unprecedented scale and scope of data consumption, stored and analyzed with significantly less effort, yielding greater and more efficient processing outputs. At a national level, central banks are also seen to potentially benefit from big data use (Bholat, 2015; IFC, 2015). Numerous applications such as economic forecasting at various levels of analysis, business cycle and financial stability analysis, identification of market and credit risks, monitoring and understanding capital flows and market behavior, and conducting research on pricing dynamics are just a few of the wide array of things central banks can do with big data technologies.

However, these big data applications involve some concerns. For instance, in creating anti-fraud systems, an integration of customers' bank transactions with their other payments to other products and services is a necessity for a more accurate predictive model (Coppolino et al., 2015). Combining transaction data with customers' purchase history, web logs, social feed, and geospatial location data from smartphone applications is another method (Akter & Wamba, 2016). Also, because of the significantly increased ability to acquire data on the same person from a multitude of sources and hereby making authentication and verification of customer information easier (Che et al., 2013), background checks and personalization efforts on potential and existing clients have been made much more cost-efficient as well. A third example is on the efforts to identify and predict mortgage repayments to avoid shortfalls and reduce arrears, which requires integrating data from different payment systems tracking mortgage customers' spending habits with loan-by-loan mortgage data (Bholat, 2015). Even central banks using big data technologies are not spared from the issues on restrictions on data-sharing practices and confidentiality concerns for legal and privacy-related reasons (IFC, 2015). However, all of these practices pose a significant risk, given how finance and banking customers can be very diverse and demanding (Turner et al., 2013).

As illustrated in Figure 1 below, big data technologies employed to, for example assess a customer's creditworthiness would typically require looking into simply the bank-related transactions of payments of utilities and credit card, both of which can be made through the bank. However, to get a more accurate assessment, the bank goes beyond and looks into specific purchases made via credit card, being able to acquire more data and information on the customer's spending habits. And these spending habits can be further verified by looking into the same customer's social media posts, which are more often than not made public. Therefore, not only can the bank look into purely transactional data, but it can also look at the lifestyle data, which are unstructured, to confirm the structured transactional data. This is but a simple example of what big data technologies can do.

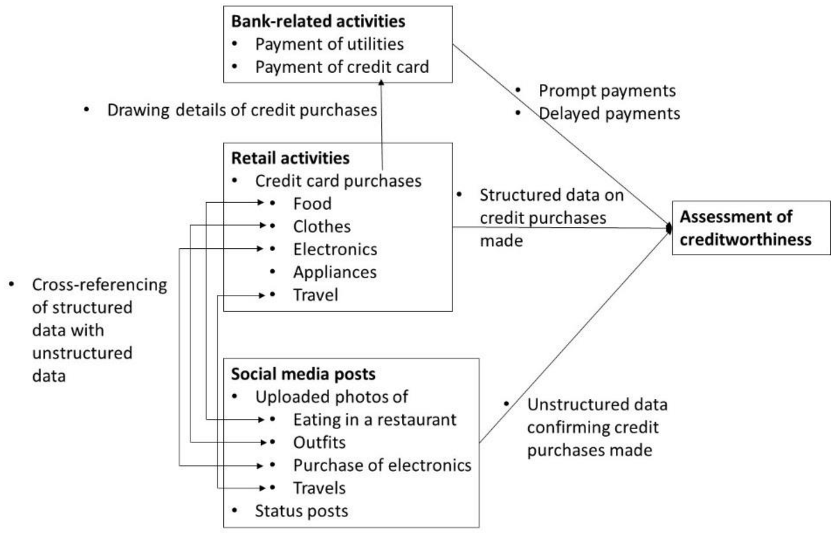

Figure 1: Illustration of big data applications in assessments of creditworthiness

This highlights one of the biggest concerns regarding big data. Big data has widened the scope and scale in which a person can be directly identified, employing several methods and channels to do so (Crawford & Schultz, 2014). Customers' personal and personalized information from an organization are being integrated with public data (Kiron et al., 2013), their social media information (Che et al., 2013), their browsing patterns, login counts, past purchases (Akter & Wamba, 2016) and other external large data sets (Katal et al., 2013). The acquisition and combination of such data from different channels to create deeper and more valuable data mining and integrative analytic and predictive models has been a significant question, bordering on trust and privacy concerns. Especially for proprietary information, issues such as ownership protection and informed consent are still critical

(Boyd & Crawford, 2012). Additionally, these particular big data capabilities enables organizations to de-anonymize and re-identify people in data economy (Akter & Wamba, 2016; Tene & Polonetsky, 2012), creating real and authentic identities that would otherwise be kept anonymous online. This goes well beyond the current levels of data sharing agreements linked into data protection and privacy mechanisms, which includes provisions for anonymization for open data, access control, rights management, and data usage control (George et al., 2014). But privacy concerns in this context also goes into the analytics and prediction part of big data, where the developed models and the generated results creating highly accurate customer pseudo-profiles can be potentially encroaching as well (Crawford & Schultz, 2014).

The easiest way to ensure that there is an acceptable level of standardized technology use within the finance and banking industry is to leverage existing industry practices (Madche, 2015), and even best practices (IFC, 2015), which is already subject to certain regulations that would ensure its integrity. But the problem with simply relying on this measure is that industry best practices are constantly changing and that these regulations that govern industry practices are not up-to-date with what big data can potentially do (Crawford & Schultz, 2014; Tene & Polonetsky, 2012). Therefore, more alternatives are being advocated to fill in the gaps. One of the proposed ways to create a better fit between big data applications and finance and banking activities is the examination of the financial industry business ontology (FIBO) (Bennett, 2013), which essentially governs data management practices within the financial services industry. It is important to understand that big data applications in the financial services sector have three key issues to consider (Kshetri, 2016). The first is the barriers and challenges related to accessing financial services. The second is the high transactions costs and inefficient processes. The third is the presence of informational opacity, moral hazard, and adverse selection processes. Almost immediately, a challenge lies on how this FIBO framework can help overcome these three key issues hounding the application of big data technologies, especially in light of the fact that the this industry is one of the most stringent in terms of data generation and exchanges, subject to a wide and complex variety of regulatory constraints (Munar et al., 2014).

FIBO has originally developed within the context of formal information systems management theory, providing some direction in the creation of common semantics to reconcile and compare data and of a framework to use semantic technology applications to analyze data in new ways (Bennett, 2013). What is interesting about FIBO is that its potential to address these regulatory concerns is consistent with what some sectors propose as a procedural or operational approach as an alternative. This proposed procedural data due process (Crawford & Schultz, 2014) and a code of conduct for big data practices (Forest et al., 2014) both would regulate the fairness of how organizations employ big data in any process, including using data and metadata to assign attributes to, create profiles for, and segment and categorize individuals. In competitive environments where business processes are the one of the last key opportunities for differentiation (Davenport, 2006), it does make sense to focus on the interactions between regulations governing big data processes and how organizations employ such processes to their advantage. Currently, consumer banking and financial services maintain their own data warehouses and business intelligence and analytics tools (Oracle, 2015).

In this particular case, FIBO-compliant approaches require two main considerations (Bennett, 2013): The model needed to be understood in terms of some formal logic. And the content of the model needed to be represented in terms that business stakeholders could understand without technical notation. This simply means that combining and translating big data terms into the appropriate business language that the users will easily understand. To develop the appropriate big data models that are easy to adopt and implement, meanings must be derived from real world business constructs.

2.4 Big data, government intervention, and the legal framework

Governments wanting to get involved with big data must recognize that data comes not only from multiple channels, but also from different sources, resulting to a wide array of data exchanges at different levels of government (Kim et al., 2014). The current array of research points to policy creation, government-business relationship building, and infrastructure development as the means in which government can intervene (Chang et al., 2014; Kshetri, 2016). However, data and information used in big data modeling often falls outside the scope of current governance and regulation (Oracle, 2015), as has been previously pointed out as well. Despite rapid developments on the technological side, there is little progress at the strategic level, especially in terms of realizing value and managing risks, resulting in the presence of significant legal and regulatory hazards that include data protection, competition laws, and intellectual property rights, among others (EY, 2014). The data issue has fallen into a broader regulatory situation beyond defining what data and information is and should do. This issue now covers the details of systems and processes that banks and other financial institutions and organizations employ to manage their data and information (Hall et al., 2014).

Big data is highly touted that it can do current practices better and differently and completely new things all together, can co-create value with customers, and can better monetize data (Frizzo-Barker et al., 2016). Furthermore, the interactions of various skilled people, technology, organizations, and shared information within certain contexts of language, laws, measures, and models are the main sources of this value and monetization generation (Demirkan & Delen, 2013). But how exactly to do these within the boundaries of what is socially acceptable or legal is still an ongoing issue (Che et al., 2013; Kemp, 2014; Tene & Polonetsky, 2012), since there are multiple ways to go about addressing this. What makes it harder to craft such context is the difficulties to completely know in advance what, when, and how big data technologies can make predictions that are encroaching on legitimate privacy concerns (Crawford & Schultz, 2014).

One very basic argument that poses an daunting challenge is that just because the data is accessible in a public sphere does not mean that its outright acquisition is automatically ethical (Boyd & Crawford, 2012). As an example of this dilemma, organizations employing some big data technologies and practices can choose to not collect data from any first or third party, and therefore be under no obligation in the context of current privacy regulations to give notice to or gather consent from its customers in the same way that direct collection methods and practices require (Crawford & Schultz, 2014). This in itself already poses a direct challenge to existing understandings on what informed consent is under current views on technology use and adoption.

A fundamental starting point for the development of the appropriate legal framework for big data is to determine the nature of data and information (EY, 2014; Kemp, 2014). By having a common understanding of what data and information should be and should do, it will definitely be less challenging to address the current disparities regarding perceptions and dispositions on how big data applications can be properly harnessed for greater benefits. The pressure to come up with standardized practices and approaches to big data applications is always present, and in fact it is actually growing (Madche, 2015). It is also important to keep in mind that big data does not always and necessarily mean more and better data (Boyd & Crawford, 2012; Lazer et al., 2014). The quantity of data does not immediately and necessarily equate to an equivalent quality of data (Forest et al., 2014). Not all data are useful, with much of it are just noise, and for some instances, the noise is growing faster than the actual signal. Interesting, one proposal in that direction points to a possible viewing the identifiability of data as a continuum as opposed to the current dichotomy of it being explicitly personally identifiable vs non-personally identifiable (Tene & Polonetsky, 2012). Furthermore, attempting to reconcile whatever common understanding with the finance and banking industry’s best practices involving data management provides additional significant insight on developing the concrete steps needed for government to legally intervene in big data use.

The growing knowledge of information and communications technologies, coupled with the continuous evolution of the Internet, has made for the convergence of architectures, infrastructures, business processes, Web 2.0 and 3.0, cloud computing, mobile Internet, and other technologies necessary to produce new services and service platforms, all exploiting the power of big data (Demirkan & Delen, 2013). This is because big data technologies has grown to include not just the data storage and analytics, but also infrastructures and data processing and management practices (Forest et al., 2014; Oracle, 2015). However, this convergence is still subject to numerous issues. Therefore, a second fundamental starting point is to identify where big data technologies and practices currently exploit the weaknesses in existing regulatory paradigms. Big data can elude current pertinent and relevant laws and regulations primarily because of the very dynamic and unpredictable nature of its analytics and prediction capabilities (Crawford & Schultz, 2014). For instance, one cannot assess the privacy risks from the collection of a single data point since current regulations do not touch on this particular practice. The same goes for the data and information transfer and processing. One cannot predict when a certain practice of data and information processing will produce privacy harms because the data that ends up being personally identifying does not yet exist during significant parts of the transfer and processing flow. Therefore, explicitly identifying the scope of information subject to privacy law in light of big data technologies must be done (Tene & Polonetsky, 2012) to start plugging the weaknesses in the existing regulatory frameworks.

As far as interventions of central banks are concerned, it is a question of how to best integrate their supervisory and regulatory methods and practices with existing analytical practices that big data technologies can afford, which can then spill over to an entire nation's financial system (Bholat, 2015). While central bank officials have indeed expressed interest in big data at the senior policy level, actual involvement is still very much limited, influenced by the restrictions of present legislations and bilateral and multilateral agreements (IFC, 2015).

2.5 Big data, informed consent theory, and privacy concerns

Information systems research is an interaction of technological and social systems existing in an intersection of knowledge and discussion about physical machines and human behavior (Gregor, 2006). Socio-political questions as to how various disciplines can aid in understanding the use of certain technologies in the human behavior context, moderated by the relevant social, ethical, and political issues, are important to appreciate this interaction. In addition, the importance of having the appropriate organizational and cultural support, usually the burden of management, is a significant contributing factor (Davenport, 2006). In this particular case, all of the theoretical efforts must be focused on answering the fundamental question on what are the appropriate regulatory requirements governing the collection, manipulation, storage, and distribution of data (Bholat, 2015), especially now in light of technologies that challenge traditional and existing regulations.

With big data, privacy concerns are extended beyond the usual quantity of data collected from people (Forest et al., 2014; Tene & Polonetsky, 2012). Even before the advent of big data, the issue of collecting data for a specific purpose, and yet using them for another, and oftentimes undisclosed, reason has been present. There are a number of instances, and therefore an equitable number of previous studies, where the information is gathered for one purpose but then used for other purposes, especially with its ever-increasing availability in electronic form (Bellman, Johnson, Kobrin, & Lohsed, 2004; Cate, 1997; Culnan & Armstrong, 1999; Culnan & Williams, 2009; Foxman & Kilcoyne, 1993; Goodwin, 1991; Hann, Hui, Lee, & Png, 2007; Laudon, Laudon, & Brabston, 2005; Nowak & Phelps, 1995; Phelps, Nowak, & Ferrell, 2000; Premazzi et al., 2010; Solove, 2004; Tavani, 2007; Taylor, Davis, & Jillapalli, 2009). What makes this more critical is the fact that the dynamic nature of information collection and manipulation reduces the customer's ability to keep track of their information that they provided to the organization (Smith, Milberg, & Burke, 1996). And big data technologies are no exception to this dynamic nature. This simply reinforces the urgency that elements of the privacy framework should adjust to reflect existing technological and organizational realities (Tene & Polonetsky, 2012).

In addition, the nature of the data and information collected, processed, and analyzed vis-à-vis prevalent privacy concerns is a second issue. Organizations ask for their customers' personal information in exchange for tailor-fit or personalized products and services (Culnan & Armstrong, 1999; Hann et al., 2007; Hoffman, Novak, & Peralta, 1999; Malhotra, Kim, & Agarwal, 2004; Zeithaml, Parasuraman, & Malhotra, 2002), but some personal information may be considered as sensitive by the consumers (Malhotra et al., 2004). The evaluation of whether a piece of personal information is sensitive or not is based on its type (Anderson & Agarwal, 2011), which is usually categorized into either personally identifying information or non-personally identifying information. But big data technologies have been blurring the lines between these two categories, mainly due to its power to piece together several pieces of personally identifying and non-personally identifying information to create an accurate pseudo profile of people. Big data challenges these fundamental principles (Tene & Polonetsky, 2012).

Despite widespread hype and use in the business community, there is proportionately little published management scholarship that tackles the challenges of using big data to explore opportunities for new theories and practices (Gozman et al., 2015). By understanding all of this in light of big data scenarios, this will allow both academics and practitioners to refine the model for data sharing and data rights, which could be universally beneficial and define big data collaborations (George et al., 2014). A theoretical starting point is to reexamine the meaning of informed consent, which posits that an organization must let its customers know about its data and information management policies, and that there must be voluntary agreement from its customers regarding these policies. (Kracher & Corritore, 2004; Tavani, 2007; Warkentin, Johnston, & Shropshire, 2011). However, this has already been made contentious by the ongoing debate on the ownership of personal digital footprints left by consumers (EY, 2014; Hann et al., 2007; Kracher & Corritore, 2004; Malhotra et al., 2004; Tavani, 2007) in cyberspace, whether it is a private or public domain under whose jurisdiction. Big data technologies are exacerbating this problem. The practice of collecting different pieces of data for different purposes and then integrated into a predictive model for yet another purpose puts a significant strain on technical and legal concerns (Bholat, 2015). The extraordinary societal benefits of big data must be reconciled with increased risks to individuals' privacy (Tene & Polonetsky, 2012), as demonstrated in this research.

3. Proposed conceptual research framework development

3.1 Developing the conceptual research framework

Interestingly, previous researches on big data depart from the traditional academic practices of grounding the analysis and discussion on some fundamental theory, concept, or framework. Much of the discussion have been practice-driven (George et al., 2014). Because of the disruptiveness of big data (Frizzo-Barker et al., 2016), fundamental academic discourses based on theoretical underpinnings have been scarce, especially surprisingly in the management field (Gozman et al., 2015). Instead, it has relied more on a practical understanding of their technical capabilities vis-à-vis feasibility issues, mostly to bolster efforts of marketing big data to organizations (Gandomi & Haider, 2015). As a result, big data has become an imprecise (Crawford & Schultz, 2014) and poorly-explored concept, obstructing its theoretical and practical development (Akter & Wamba, 2016) and making for the current incoherent understanding of its strategic implications (Hall et al., 2014). This represents the lack of research and understanding due to the fact that the acceptance and success any piece of technology is largely, if not completely, subject to human intervention and interaction (Gregor, 2006).

For a better fit involving technology adoption and use, the discussion must provide a set of organizational (Ebner et al., 2014; Fuschi & Tvaronavičienė, 2011) and cultural (Akter & Wamba, 2016) structured and unstructured factors, processes (Madche, 2015; Munar et al., 2014), and decision making environment (Boyd & Crawford, 2012; Işık et al., 2013) that influence the implementation of the different big data strategies that are available for consideration. There are numerous examples that highlight the importance of technology fit, and all of them point to the integral role of management (Davenport, 2006) to influence these different factors.

And since big data technologies have to integrate themselves with existing platforms, systems, and practices, this process of integration must also be subject to some form of regulations and regulatory environments depending on the context (Boyd & Crawford, 2012; Fuschi & Tvaronavičienė, 2011; Işık et al., 2013). Examining the proper context of this integration is yet another key consideration for this research. One fundamental way of doing so is to examine its different components (Gandomi & Haider, 2015) and characteristics of velocity, veracity, variety, volume, complexity, and value (Akter & Wamba, 2016) in terms of how each can fit into and subsequently affect and change existing processes and systems. This therefore goes back to the potential of employing a combination of FIBO-based approaches (Bennett, 2013) and the proposed procedural data due process (Crawford & Schultz, 2014), since this examination and determination of fit is directed at precisely the big data-based processes and practices.

This also requires a significantly holistic and integrative view, bringing together scientists, practitioners, and even politicians and prominent members of the civil society (Boyd & Crawford, 2012; Crawford & Schultz, 2014; Fuschi & Tvaronavičienė, 2011; Lazer et al., 2014), and even central bank officials (Bholat, 2015; IFC, 2015). The interaction of many different fields such as the arts, design, computer science, law and mathematics and of many different backgrounds such as psychology, management, sociology, and philosophy is an integral part of developing the proper theoretical foundation needed to study the behaviors of this phenomenon (Gregor, 2006). On top of the widely-accepted dynamism and complexity of big data itself, this is due to the fact that the application of due process to data requires significant discourse, imagination, and debate to design the appropriate processes and procedures (Crawford & Schultz, 2014), in light of several legal, compliance, ethical, and commercial risks involved in data practices (Forest et al., 2014; Tene & Polonetsky, 2012). It does go without saying how difficult, and time-consuming this design can be.

Regardless of the outcome of this integration and collaboration, both private organizations and regulators must define a big data strategy with a business-centric blueprint built on big data's capabilities that coincide with acceptable business practices (Turner et al., 2013) and operational decisions (EY, 2014). Therefore, it is an interaction of technology, practices, regulation, and compliance (Gozman et al., 2015), built on both relevant theoretical and practical underpinnings.

What is presented in Figure 2 is the proposed conceptual research framework, painting the broad strokes that set the agenda and the direction for the formulation of the regulatory framework. The first step is clearly defining what big data capabilities can do when it comes to privacy concerns and to categorizing personal data. This research has made some connections regarding what some of these are. What is more difficult and time-consuming is on addressing these issues on privacy concerns and on the categorizing personal data. This is the part in which the issues on collecting data for one purpose and yet its use is for another purpose and on the definition of what private vs. public and what is personally identifying vs non-personally identifying must be addressed. It is only after this where the efforts can be then directed towards building the regulatory framework in terms of what organizations can do, especially in light of what data should be and can do. However, transition must be made with the perspective of informed consent. Informed consent is still a relevant theoretical foundation, which is being challenged with these new technologies.

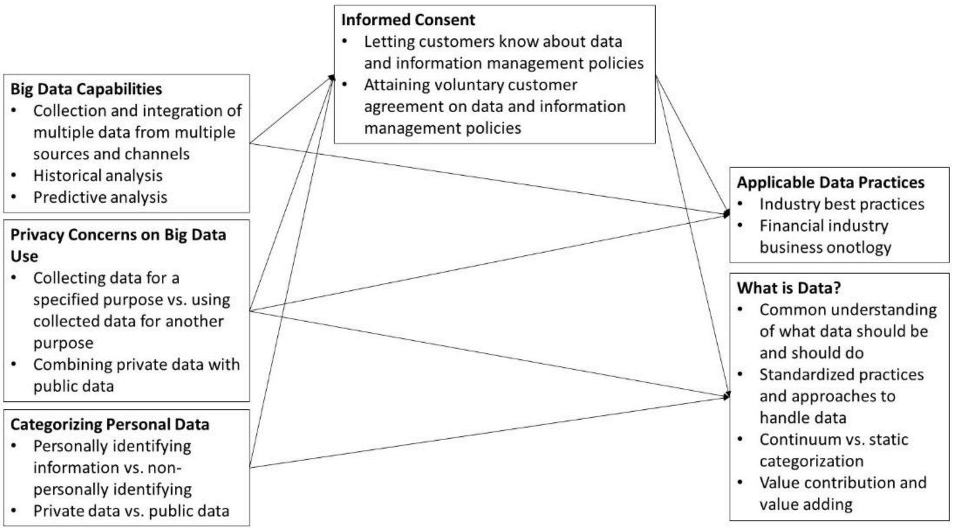

Figure 2: The proposed conceptual research framework

4. Concluding Remarks

In its present form, this research presents the current trends of big data, but at the same time casts a veil of caution over being carried away by these trends. Indeed, as big and as significant as the potential is, important support structures are unfortunately lagging behind, threatening the overall feasibility and sustainability of its full use. This research presents one of those important support structures in the form of the regulatory environment to govern the use of big data technologies, specifically for the finance and banking industry. To reiterate, this research has three broad objectives, each of which are detailed, based on the literature review and the subsequent arguments and discussions, as follows.

To explore the possible components and capabilities of big data that are appropriate to meet specific finance and banking needs and requirements.

It is clear that big data has the capabilities to enable finance and banking organizations provide better products and services to its clientele, while at the same time protecting itself from the inherent risks of providing these products and services. In addition, organizations can do so with significantly less resources and effort to do so, especially in the long run. Big data not only provides a very efficient means to collect all sorts of structured and unstructured data and information both from the customer side and the market and industry side from all sorts of different channels, but it also provides powerful computing power to develop and run both more accurate historical analyses and more reliable predictive analyses. If employed properly with the right set of human skills, the opportunities for more cost-efficient and effective research and development of innovative products and services can grow exponentially.

To reiterate, for the case of the finance and banking industry, it is all about making better sense of customer behaviors to make more accurate assessments of their creditworthiness, risk appetites, and qualifications for higher valued products and services other than the usual deposit accounts and debit and credit cards. Especially nowadays wherein banks are bundling fixed-income securities and equities into investment portfolios, and even packaging them further with different insurance policies, there is a more obvious and urgent need to make more accurate assessments beyond the typical risk profiling that are being doing.

To identify possible gaps between the available big data technologies and the finance and banking business processes.

On top of the significant manpower shortage, the most significant gap between big data and business practices is the lack of a concrete understanding of what big data is, especially in light of what the finance and banking industry requires. There is an obvious dilemma regarding the relatively urgent need of banks to protect itself from credit frauds and mortgage arrears, for example, but employing big data technologies to address this raises the ethical questions stemming from informed consent. As discussed, current practices and regulations put the banks under no obligation to give notice to or attain consent from its customers to analyze their personal data acquired from digital footprints existing in the public cyberspace. Furthermore, the practice of combining several pieces of personal information as required by big data technologies to work effectively and efficiently raises the question on the same informed consent paradigm. Legally and ethically, banks are bound to limit the use of the data and information acquired from customers as stipulated in whatever agreements the customer has either explicitly or implicitly conformed to when they filled out the customer information sheet or an application form for a certain product or service. However, as illustrated, big data technologies blur those limits.

This presents another gap pointing to the level of appreciation and understanding bank employees, managers, and executives have not only on the use of big data technologies, but also on the benefits and risks of employing them. With the question of what can banks do and cannot do looming over their heads, coming to a unified and standardized understanding of this issue is of utmost concern as well.

To determine if the available big data technologies vis-à-vis finance and banking practices adhere to the appropriate governance and regulatory environments.
As discussed, there is a serious lack of a regulatory framework to govern what big data can and cannot do. Unfortunately, current regulations on data and information use, especially those that are acquired from customers, are insufficient. The disruptiveness of big data technologies have obviously blurred the boundaries that categorize what is private data and what is public data; what is personally identifiable data and what is non-personally identifiable data. The traditional theoretical and practical underpinnings of what data and information privacy is has been significantly challenged not only by what big data can do, but also of when and how it can execute its numerous capabilities.

Looking forward, this research proposes to further answer the following questions to better understand what needs to be done to address these regulatory issues on big data technology adoption and use.

How will government proactively get involved? What is required for government to effectively get involved?
There is no question that the government, through the appropriate departments, ministries, bureaus, and agencies, must proactively get involved. As the entity in charge of protecting, preserving, and defending the rights of its citizens as stipulated in the nation's Constitution, which includes the right to data and information privacy, the government must ensure that big data technologies, no matter how beneficial they are, do not encroach on people that would result in some negative ethical or legal ramifications. But on the same thread, as the enforcer of the same Constitution, the government must also exert every appropriate effort to make sure that big data technologies are used to contribute to the continuing progress of the nation. This research has mentioned some of the applications that help both the national and local levels of government do its job, but it must do so without violating existing citizens' rights.

What is further interesting is that this requires a collaboration of different national government departments to effectively address. It is highly impossible for one department to do this alone. The collaboration will depend on what the national government deems as necessary and urgent applications of big data. For example, in the case of this context of finance and banking, the central bank and the finance department are prime targets for collaboration, on top of the science and technology department, and the newly formed information and communications department. Especially for the last one, it is a big task ahead of them.

Will the private sector take the lead? How will the private sector take the lead?
It is obvious that the private sector has significantly more resources to invest in big data technologies, including the skills and training that are required to properly operate these technologies. Furthermore, it is the private sector that has a more direct contact with customers. As leaders in the industry, setting the trend for what are the acceptable practices and therefore lobbying and contributing for the proper regulatory measures is an effective mechanism to start things off from the their end, while the public sector prepares for their own contributions to craft this regulatory environment as necessary.

Addressing the gaps caused by big data technologies is something that cannot be done overnight, but would take significant amount of time and effort to do so. The regulatory environment is just one of the many gaps. The lack of empirical research alone on the adoption, acceptance, and use of big data technologies is something seriously worth considering, especially in light of the significant imbalance between the empirical and the conceptual research. This is not something that either a purely academic or a purely managerial approach can accomplish on its own.

References

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: A systematic review and agenda for future research. *Electronic Markets, 26*(2), 173-194.

Anderson, C. L., & Agarwal, R. (2011). The digitization of healthcare: Boundary risks, emotion, and consumer willingness to disclose personal health information. *Information Systems Research, 22*(3), 469-490.

Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard Business Review, 90*(10), 78-83.

Bedeley, R. T., & Iyer, L. S. (2014). *Big Data opportunities and challenges: The case of banking industry.* Paper presented at the Southern Association for Information Systems Conference, Macon, GA.

Bellman, S., Johnson, E. J., Kobrin, S. J., & Lohsed, G. L. (2004). International differences in information privacy concerns: A global survey of consumers. *The Information Society, 20*(5), 313-324.

Bennett, M. (2013). The financial industry business ontology: Best practice for big data. *Journal of Banking Regulation, 14*(3-4), 255-268.

Bholat, D. (2015). Big data and central banks *Quarterly Bulletin* (Vol. 1, pp. 86-93): Bank of England.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662-679.

Cate, F. H. (1997). *Privacy in the information age*. Washington, D.C.: The Brookings Institution.

Chang, R. M., Kauffman, R. J., & Kwon, Y. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems, 63*(1), 67-80.

Che, D., Safran, M., & Peng, Z. (2013). From big data to big data mining: challenges, issues, and opportunities. In B. Hong, X. Meng, L. Chen, W. Winiwarter & W. Song (Eds.), *Database Systems for Advanced Applications* (pp. 1-15). Berlin, Germany: Springer Berlin Heidelberg.

Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences, 275*(1), 314-347.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From Big Data to big impact. *MIS Quarterly, 36*(4), 1165-1188.

Connors, S., Courbe, J., & Waishampayan, V. (2013). Where have you been all my life? How the financial services industry can unlock the value in Big Data: PricewaterhouseCoppers LLP.

Coppolino, L., D'Antonio, S., Romano, L., Campanile, F., & de Carvalho, A. V. (2015). Effective visualization of a Big Data banking application. In E. Damiani, R. J. Howlett, L. C. Jain, L. Gallo & G. De Pietro (Eds.), *Intelligent Interactive Multimedia Systems and Services* (pp. 57-68). Switzerland: Springer International Publishing.

Crawford, K., & Schultz, J. (2014). Big data and due process: Toward a framework to redress predictive privacy harms. *Boston College Law Review, 55*(1), 93-128.
Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science, 10*(1), 104-115.
Culnan, M. J., & Williams, C. C. (2009). How ethics can enhance organizational privacy: Lessons from the ChoicePoint and TJX data breaches. *MIS Quarterly, 33*(4), 673-687.
Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review, 84*(1), 98-107.
Demirkan, H., & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems, 55*(1), 412-421.
Ebner, K., Bühnen, T., & Urbach, N. (2014). *Think big with Big Data: Identifying suitable Big Data strategies in corporate environments.* Paper presented at the 47th Hawaii International Conference on System Sciences, Waikoloa, HI.
EY. (2014). Ready for takeoff? Overcoming the practical and legal difficulties in identifying and realizing the value of data: Ernst & Young Global Limited.
Forest, H., Foo, E., Rose, D., & Berenzon, D. (2014). Big Data: How it can become a differentiator: Deutsche Bank AG Global Transaction Banking.
Foxman, E. R., & Kilcoyne, P. (1993). Information technology, marketing practice, and consumer privacy: Ethical issues. *Journal of Public Policy and Marketing, 12*(1), 106-119.
Frizzo-Barker, J., Chow-White, P. A., Mozafari, M., & Ha, D. (2016). An empirical study of the rise of big data in business scholarship. *International Journal of Information Management, 36*(3), 403-413.
Fuschi, D. L., & Tvaronavičienė, M. (2011). Sustainable development, Big Data and supervisory control: Service quality in banking sector. *Journal of Security and Sustainability Issues, 3*(3), 5-14.
Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137-144.
George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. *Academy of Management Journal, 57*(2), 321-326.
Goodwin, C. (1991). Privacy: Recognition of a consumer right. *Journal of Public Policy and Marketing, 10*(1), 149-166.
Gozman, D., Currie, W., & Seddon, J. (2015). The role of big data in governance: A regulatory and legal perspective of analytics in global financial services: SWIFT Institute.
Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly, 30*(3), 611-642.
Hall, J., Luscombe, P., Urry, N., Richardson, G., Hart, B. J., Stork, P., . . . Kelly, H. (2014). Managing the data challenge in banking: Why is it so hard? : KPMG International Cooperative.
Hann, I. H., Hui, K. L., Lee, S. Y. T., & Png, I. P. L. (2007). Overcoming online information privacy concerns: An information-processing theory approach. *Journal of Management Information Systems, 24*(2), 13-42.
Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems, 47*(1), 98-115.
He, W., Wu, H., Yan, G., Akula, V., & Shen, J. (2015). A novel social media competitive analytics framework with sentiment benchmarks. *Information & Management, 52*(7), 801-812.
Hoffman, D. L., Novak, T. P., & Peralta, M. A. (1999). Building consumer trust online. *Communications of the ACM, 42*(4), 80-85.
Hu, D., Zhao, J. L., Hua, Z., & Wong, M. C. (2012). Network-based modeling and analysis of systemic risk in banking systems. *MIS Quarterly, 36*(4), 1269-1291.
IFC. (2015). Central banks' use of and interest in "big data": Irving Fisher Committee on Central Bank Statistics, Bank for International Settlements.
Işık, Ö., Jones, M. C., & Sidorova, A. (2013). Business intelligence success: The roles of BI capabilities and decision environments. *Information & Management, 50*(1), 13-23.
Katal, A., Wazid, M., & Goudar, R. H. (2013). *Big data: Issues, challenges, tools and good practices*. Paper presented at the 2013 Sixth International Conference Contemporary Computing (IC3) Noida, India.

Kemp, R. (2014). Legal aspects of managing Big Data. *Computer Law & Security Review, 30*(5), 482-491.
Kim, G.-H., Trimi, S., & Chung, J.-H. (2014). Big-data applications in the government sector. *Communications of the ACM, 57*(3), 78-85.
Kiron, D., Ferguson, R. B., & Prentice, P. K. (2013). From value to vision: Reimagining the possible with data analytics (pp. 1-19): MIT Sloan/SAS.
Kracher, B., & Corritore, C. L. (2004). Is there a special e-commerce ethics? *Business Ethics Quarterly, 14*(1), 71-94.
Kshetri, N. (2016). Big data's role in expanding access to financial services in China. *International Journal of Information Management, 36*(3), 297-308.
Laudon, K. C., Laudon, J. P., & Brabston, M. E. (2005). *Management information systems: Managing the digital firm*. Toronto, Ontario: Pearson Education Canada, Inc.
LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review, 52*(2), 21-31.
Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in Big Data analysis. *Science, 343*(6176), 1203-1205.
Madche, A. (2015). Interview with Wolfgang Gaertner on ‘‘Digitalization in retail banking: Differentiation and standardization through IT’’. *Business & Information Systems Engineering, 57*(1), 83-85.
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.
Mansour, R. F. (2016). Understanding how big data leads to social networking vulnerability. *Computers in Human Behavior, 57*(1), 348-351.
Munar, A., Chiner, E., & Sales, I. (2014). *A big data financial information management architecture for global banking.* Paper presented at the 2014 International Confernece on Future Internet of Things and Cloud (FiCloud), Barcelona, Spain.
Nowak, G. J., & Phelps, J. (1995). Direct marketing and the use of individual-level consumer information: Determining how and when "privacy" matters. *Journal of Direct Marketing, 9*(3), 46-60.
Oracle. (2015). Big Data in financial services and banking: Architect's guide and reference architecture introduction. Redwood Shores, CA: Oracle Corporation.
Phelps, J., Nowak, G., & Ferrell, E. (2000). Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy and Marketing, 19*(1), 27-41.
Premazzi, K., Castaldo, S., Grosso, M., Raman, P., Brudwig, S., & Hofacker, C. F. (2010). Customer information sharing with e-vendors: The roles of incentives and trust. *International Journal of Electronic Commerce, 14*(3), 63-91.
Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly, 20*(2), 167-196.
Solove, D. J. (2004). *The digital person: Technology and privacy in the information age*. New York, NY: New York University Press.
Srivastava, U., & Gopalkrishnan, S. (2015). Impact of Big Data Analytics on banking sector: Learning for Indian banks. *Procedia Computer Science, 50*(1), 643-652.
Tambe, P. (2014). Big Data investment, skills, and firm value. *Management Science, 60*(6), 1452-1469.
Tavani, H. T. (2007). *Ethics & technology: Ethical issues in an age of information and communication technology*. Hoboken, NJ.: John Wiley & Sons, Inc.
Taylor, D. G., Davis, D. F., & Jillapalli, R. (2009). Privacy concern and online personalization: The moderating effects of information control and compensation. *Electronic Commerce Research, 9*(3), 203-223.
Tene, O., & Polonetsky, J. (2012). Big data for all: Privacy and user control in the age of analytics. *Northwestern Journal of Technology and Intellectual Property, 11*(5), 240-273.
Turner, D., Schroeck, M., & Shockley, R. (2013). Analytics: The real-world use of big data in financial services: How innovative banking and financial markets organizations extract value

from uncertain data: IBM Institute for Business Value and Saïd Business School at the University of Oxford.

Warkentin, M., Johnston, A. C., & Shropshire, J. (2011). The influence of the informal social learning environment on information privacy policy compliance efficacy and intention. *European Journal of Information Systems, 20*, 267-284.

Weill, P., & Woerner, S. L. (2015). Thriving in an increasingly digital ecosystem. *MIT Sloan Management Review*, *56*(4), 27-34.

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2002). Service quality delivery through web sites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science, 30*(4), 362-375.