CORE
Metadata, citation and similar papers at core.ac.uk
Provided by International Institute for Science, Technology and Education (IISTE): E-Journals

Journal of Information Engineering and Applications
ISSN 2224-5782 (print) ISSN 2225-0506 (online)
Vol.6, No.4, 2016
www.iiste.org

The Efficacy of the "Big Data" Syndrome and Organizational Information Governance

Antonia Enefaa Bestman
Office and Information Management, Rivers State University of Science and Technology
Port Harcourt, Rivers State

Henry Wogboroma
Office and Information Management, Rivers State University of Science and Technology
Port Harcourt, Rivers State

Abstract
This paper addresses the challenge of big data for the design of organisations in governance. Big data refers to the availability to organisations of massive amounts of heterogeneous and continuously updated information. Practitioners agree that the availability of such information creates challenges and opportunities for organiations that have never been seen before. The article presented here takes up this challenge and discusses avenues for future research and practice on organsiation design in the era of big data. The importance of digital technologies for social and economic developments and a growing focus on data collection and privacy concerns have made the internet a salient and visible issue in global politics. Surprisingly, little research has explored questions about the relations between business, governance and the internet. Government organisations are feverishly exploring ways of taking advantage of the big data phenomenon. This paper seeks to expand our knowledge of the intersections between business management, global governance and the digital domain.

1. Introduction

Big data burst upon the scene in the first decade of the 21st century, and the first organisations to embrace it were online and startup firms. Arguably, firms like Google, eBay, LinkedIn, and Facebook were built around big data from the beginning. They did not have to reconcile or integrate big data with more traditional sources of data and the analytics performed upon them, because they didn't have those traditional forms. They did not have to merge big data technologies with their traditional IT because those infrastructures did not exist. Big data could stand alone, big data analytics could be the only focus of analytics, and big data technology architectures could be the only architecture. Consider, however, the position of large government organisations. Big data in those environments should not be separate, but must be integrated with everything else that is going on in the organization.

Analytics on big data have to coexist with analytics on other types of data. Data scientists must somehow get along and work jointly with mere quantitative analysts. Big data is a broad term for data sets so large or complex that traditional data processing applications are inadequate. Challenges include analysis, capture, curation, search, sharing, storage, transfer, visualisation, and information privacy. The term often refers simply to the use of predictive analytics or other certain advanced methods to extract value from data, and seldom to a particular size of data set. Analysis of data sets can find new correlations, to "spot business trends, prevent diseases, combat crime and so on." Scientists, practitioners of media and advertising and governments alike regularly meet difficulties with large data sets in areas including interest search, finance and business informatics. Data sets grow in size in part because they are increasingly being gathered by cheap and numerous information-sensing mobile devices, aerial (remote sensing), software logs, cameras, microphones, radio-frequency identification (RFID) readers, and wireless sensor networks. The world's technological per-capita capacity to store information has roughly doubled every 40 months since the 1980s; as of 2012, every day 2.5 exabytes ($2.5x10^{18}$) of data were created; The challenge for large enterprises is determining who should own big data initiatives that straddle the entire organisation. Relational database management systems and desktop statistics and visualization packages often have difficulty handling big data. The work instead requires 'massively parallel software running on tens, hundreds, or even thousands of servers". What is considered "big data" varies depending on the capabilities of the users and their tools, and expanding capabilities make Big Data a moving target. Thus, what is considered to be "Big" in one year will become ordinary in later years. "For some organizations, facing hundreds of gigabytes of data for the first time may trigger a need to reconsider data management options. For others, it may take tens or hundreds of terabytes before data *size* becomes a significant consideration."

The use and adoption of Big Data within governmental processes is beneficial and allows efficiencies in terms of cost, productivity, and innovation. That said, this process does not come without its flaws. Data analysis often requires multiple parts of government (central and local) to work in collaboration and create new and innovative processes to deliver the desired outcome. We are witnessing a growing gulf between those companies

that see the value of business analytics and are transforming themselves to take advantage of these new found opportunities, and those that are yet to embrace them. Successful leaders are infusing analytics throughout their enterprise to drive smarter decisions, enable faster actions and optimize outcome.

In the right hands, big data - the massive amounts of information organizations collect today, can become a valuable new asset. With the emergence and increasing adoption of big data, enterprises are discovering new ways to compete and win — transforming themselves to take advantage of the vast array of available information to improve decision-making and performance. In today's competitive marketplace, executive leaders are racing to convert enterprise insights into meaningful results. With the help of big data, health insurance companies can reduce fraudulent claims, pharmaceutical companies can improve drug efficacy and safety, manufacturers can create predictive supply chains, financial service firms can manage risk effectively, and retailers can master real-time inventory and pricing. The possibilities are endless. Winners are those with the ability to sift through the mounds of new and emerging data to uncover the few precious nuggets with significant business value - the ability to make big data seem small.

2. Data Governance

Every emerging global technology trend comes with increased opportunities, challenges and also risks. People and technology are the major drivers of these trends. It is amazing how technology has changed information in itself. Data is increasing every day with the increased use of connected devices, the internet and social media. Information is now generated and consumed at greater speed than ever. This trend will get even bigger with the future and with Internet of Things.

Some interesting statistics show that over 2 billion people are now connected to the internet, and that every minute 27,778 blog posts are made, 47,000 downloads from the Apple platform, over 100,000 tweets sent, 200,000 search queries on Google, it actually will take an individual about 5 years to read the content on the internet, personnel usually take 2 hours searching for information they kept somewhere by themselves. It is simply amazing. How these information overload have turned into unstructured data. All of that happening with great speed shows that unstructured data is now becoming a major challenge for big corporations and even the government. These unstructured data constitute the Big Data. A report by Commvault shows that 80% of the data in an organization are unstructured and that 93% of executives believe they are losing revenue opportunities by not leveraging existing data.

Experts at the G1TEX conference in Dubai have opined that despite the challenges around big data, it also offers great opportunities for businesses as these trends are at the centre of transformation of businesses. People are beginning to look at the potential risks with the increased information at their disposal; others are seeing the value when this information are analyzed and structured and then come the cost of storing these data. The well-known futurist, Gerd Leonhard has said that data is the "new oil". Massively parallel and distributed database technologies available today can process a lot of this data, in rest and in motion, and provide insights in real time or near real time. Enterprises can use this stream of timely, relevant insights to become not only more operationally efficient but also more responsive to the business changes. With such fantastic gains to be had, the future certainly looks data driven and as with every major change, it pays to adapt well in advance to "get it right".

Till recently, Data served primarily as application's infrastructure and remained fragmented across business processes and IT systems. This cannot and will not be the case anymore. Implementing an information governance program, through the establishment of an enterprise-wide strategic locus on information infrastructure whose importance will be more or less equal to that of enterprise application management, is the key for any industry's success in this digital age. New analytics tools and methods are expanding the possibilities for how enterprises can derive value from existing data within their organizations and from freely available external information sources, such as software as a service (SaaS), social media and commercial data sources. While traditional business intelligence has generally targeted "structured data" that can be easily parsed and analyzed, advances in analytics methods now allow examination of more varied data types.

Information security, audit and governance professionals should take a holistic approach and understand the business case of big data analytics and the potential technical risk when evaluating the use and deployment of big data analytics in their organizations. For information security, audit and governance professionals, lack of clarity about the business case may stifle organizational success and lead to role and responsibility confusion. By looking at how these analytics techniques are transforming enterprises in real-world scenarios, the value becomes apparent as enterprises start to realize dramatic gains in the efficiency, efficacy and performance of mission-critical business processes. Understanding this business case can help security, audit and governance practitioners in two ways: it helps them to understand the motivation and rationale driving their business partners who want to apply big data analytics techniques within their enterprises, and it helps balance the risk equation so that technical risk and business risk are addressed. Specifically, while some new areas of technical risk may arise as a result of more voluminous and concentrated data, the business consequences of not adopting big data

analytics may outweigh the technology risk.

The opportunity of big data is huge, and the biggest analytical opportunity existing within is the use of predictive analytics. The data shows companies favor taking advantage of the opportunities in front of them rather than minimizing risk. Technology is playing a role here and making predictive capabilities even easier to use, embedding them in business processes, automating model creation. The idea of data creating business value is not new; however, the effective use of data is becoming the basis of competition.

Business has always wanted to derive insights from information in order to make better, smarter, real time, fact-based decisions: it is this demand for depth of knowledge that has fueled the growth of big data tools and platforms.

Those leading the change are now including big data from both within and outside the enterprise, including structured and unstructured data, machine data, and online and mobile data to supplement their organizational data and provide the basis for historical and forward-looking (statistical and predictive) views. Big data will fundamentally change the way businesses compete and operate. Companies that invest in and successfully derive value from their data will have a distinct advantage over their competitors — a performance gap that will continue to grow as more relevant data is generated, emerging technologies and digital channels offer better acquisition and delivery mechanisms, and the technologies that enable faster, easier data analysis continue to develop.

While the ability to capture and store vast amounts of data has grown at an unprecedented rate, the technical capacity to aggregate and analyze these disparate volumes of information is only just now catching up.

3. The Scope of Big Data

Big data refers to the dynamic, large and disparate volumes of data being created by people, tools and machines; it requires new, innovative and scalable technology to collect, host and analytically process the vast amount of data gathered in order to derive real-time business insights that relate to consumers, risk, profit, performance, productivity management and enhanced shareholder value. Big data includes information garnered from social media, data from internet-enabled devices (including smartphones and tablets), machine data, video and voice recordings, and the continued preservation and logging of structured and unstructured data. It is typically characterized by the four "V's":

Volume: The amount of data being created is vast compared to traditional data sources.
Variety: Data comes from different sources and is being created by machines as well as people.
Velocity: Data is being generated extremely fast a process that never stops, even while we sleep.
Veracity: Big data is sourced from many different places; as a result you need to test the veracity/quality of the data.

Evolving technology has brought data analysis out of IT backrooms, and extended the potential of using data-driven results into every facet of an organization. However, while advances in software and hardware have enabled the age of big data, technology is not the only consideration. Organizations need to take a holistic view that recognizes that success is built upon the integration of' people, process, technology and data; this means being able to incorporate data into their business routines, their strategy and their daily operations.

Organizations must understand what insights they need in order to make good strategic and operational decisions. The first part of the challenge is sorting through all of the available data to identify trends and correlations that will drive beneficial changes in business behavior. The next step is enriching this organizational information with that from sources outside the enterprise, this will include familiar big data sources, such as those created and stored online. In a business environment that constantly and rapidly changes, future prediction becomes more important than the simple visualization of historical or current perspectives. For effective future prediction, data analysis using statistical and predictive modeling techniques may be applied to enhance and support the organization's business strategy. The collection and aggregation of big data, and other information from outside the enterprise, enables the business to develop their own analytic capacity and capability, which for many years has only been available to a few larger organisations.

4. Trends in Technology

"Companies who implement Big Data strategies are seen to perform 20% better" Regina Casonato, the Managing Vice President at Gartner.

Big data is generating an intense amount of attention among governments, businesses, media and even consumers, along with analytics, cloud-based technologies, digital channels and data visualization. These are all part of the current diverse ecosystem created by the technology megatrends. Some even herald the potential transformative power of the current trends as rivaling that of' the internet. Yet, as in the early days of the internet, there is uncertainty about just what big data is, its potential benefits and the associated risks. Big data is generating an intense amount of attention among businesses, media and even consumers, along with analytics, cloud-based technologies, digital channels and data visualization. These are all part of the current diverse

ecosystem created by the technology megatrends. Some even herald the potential transformative power of' the current trends as rivaling that of the internet. Yet, as in the early days of the internet, there is uncertainty about just what big data is, its potential benefits and the associated risks.

The term Big Data has become a major theme of the technology media, but it has also increasingly made its way into many compliance, internal audit and fraud risk management- related discussions. Forensic data analytics (FDA) technologies are available to help organizations keep pace with increasing data volumes, as well as business and regulatory complexities; examples can include real-time analytical processing engines that make rapid business decisions, such as stopping a potentially improper payment or business transaction, or leveraging anti-fraud/anti-corruption monitoring controls that integrate data visualization, statistical analysis and text mining. Yet despite their availability, many companies have not scaled up their data usage to take advantage of these effective tools, and may be missing important fraud prevention and detection opportunities by not mining larger data sets to more robustly monitor business activities.

Decisions can be made with a structured approach through data-driven insight, including: Customer and product profitability, Customer acquisition and retention strategies, Customer satisfaction strategies, Marketing segmentation, Operations and performance management and Supply chain and delivery channel strategy. This new trend has made it necessary for businesses to look at what they need to do with the deluge of data they have at their disposal. Businesses will have to keep up with speed at which information is generated, the volume of these information and then the various forms in which they occur, these are obvious concern today. And that is what big data is all about.

5. New Attributes of Big Data

Firms and other organizations have been using large databases and analytics for the last couple of decades. Transactions are stored in data warehouses and analyzed with data-mining algorithms to extract insights. Nowadays, there are more — and different kinds — of data. In the past, it was stored, structured data. This data was largely from transactions and was stored as rows and columns. Today, we store unstructured data from a variety of sources. The data could be photos from a mobile phone, maps from a GPS device, video from a surveillance camera, audio from a call center, e-mails, tweets, and text messages. All of this data can be digitized, analyzed, and stored. Secondly, this new data is accessible in real time. Before, data in the data warehouse was historical and described outcomes that had already occurred. Now, we can receive data about events as they are happening and perhaps influence their outcomes. Historically, credit card companies stored all of their transactions in a database and analyzed them with fraud-detecting algorithms. Fraudulent transactions were then turned over to the police to investigate. The companies could distinguish chronically late payers from people who had lost their jobs. Customer service could then take the appropriate actions with each group.

Today, a fraudulent transaction can be detected while the fraudster is still at the checkout counter. An algorithm operating in real time can determine that the transaction is a charge on a stolen credit card. The clerk at the checkout counter can be advised to delay the suspect. Store security can be informed to apprehend the person and confiscate the credit card. Thus, real-time data allow us to influence the outcome and prevent bad outcomes before they happen. This capability is new. However, this new capability is only possible if we have an organization that is designed to Operate in real time. We need to design a decision process that uses real-time data, analyzes it to produce instant insights, and processes those insights to arrive at real-time decisions, Using real-time decisions, organizations can take quick action.
We need much faster-acting companies in order to profit from big data.

6. Power Shift

Before an organization can make real-time decisions, it must get data scientists and analytics experts embedded into decision processes. This will require a shift in power from experienced and judgmental decision makers to digital decision makers. Every organization has an establishment, a power structure with a vested interest in the status quo. The establishment is currently making investment decisions, setting customer priorities, and deciding on new product features. These are the same decisions that new insights from big data can improve. In order to be successful, the organization needs to execute a shift in power to the digital experts who generate new insights from big data. A shift in power is necessary to accomplish the changes that are needed to fully embed the big data analytics capability.

For big data, a new set of skills and competences are required. New set of strategies are developed, you now see information being measured in terms of volume, velocity and variety; new professionals like Chief Data Officer, Data science manager emerge; ideas like having a Data science Lab are becoming common, information governance is now more relevant than before with the nexus of information coming to play.

7. Competence- Enhancing

One factor that will determine the magnitude of the power shift is the amount of resistance that the big data

proponents will encounter from the establishment. The amount of resistance will depend on whether this new capability is competence-enhancing or competence-destroying. (Tushman & Anderson, 1996). For example, when c-commerce came along, Dell adopted it immediately. Taking orders over the Internet was competence-enhancing for Dell; it enhanced the company's direct sales to end-users business model. Hewlett-Packard's strength was its relationship with resellers and retailers. E-commerce was a competence-destroying innovation, it would disinter mediate HP's resellers. HP was slower to adopt c-commerce, keeping the firm's resellers in the distribution chain at a higher cost. So, with respect to any particular big data initiative, companies need to determine where they are on the competence enhancing-destroying continuum. Procter & Gamble is an example of a company for whom big data is competence-enhancing. P&G is a very analytical company and has had an analytics group since 1992. Plus, it will try anything that might increase its understanding of consumer behavior. As a result, P&G is adopting big data practices ahead of most other companies. The big data initiative is led by the CIO and supported by the CEO. P&G has adapted its hiring practices to bring in more data scientists. For the past five years, P&G and Google have exchanged teams of people annually. Google wants to learn about advertising, and P&G wants to learn from Google's digital acumen. At P&G, all managers are upgrading their digital skills. Moreover, every manager's digital and analytical performance gets assessed in the performance management process. The ClO and the business leaders have identified 88 business processes that are being redesigned and accelerated to operate in real time. So, P&G is a good example of a company that has embraced big data.

In contrast, a good example of a competence-destroying situation is the arrival of big data in spoils. In the baseball movie Moneyball, Billy Beane, the general manager of the Oakland Athletics, brings in "sabre-metrics" expert Peter Brand. Billy wants Peter to advise him on putting together the best team possible, but on Oakland's very low budget. Peter is 25 years old and an economics graduate from Yale who has never played the game. Of course, they run into the chief scout, Grady, and his grey-haired scouting team. Grady first tries to keep Peter out of the meeting. "Does Pete really need to be here?" Then he shifts to, "You can't put a team together with a computer." The meeting is a clash between Peter's data and Grady's opinions. Billy makes his decisions based on Peter's data, but then they run into the manager, Art Howe, who will not play a data-chosen player. He is insubordinate when Billy commands him to play the player. Billy then makes a trade so that Art has to play Billy's choice. Thus, the arrival of data and analytics at the Oakland Athletics destroyed or diminished the experience-based competence of Grady and his fellow talent scouts. Today, almost all American baseball teams and European soccer teams use big data and analytics to some degree.

Such scenes will play out in many companies where an analog establishment is making product and marketing decisions based on years of experience and historical data. The digital newcomers will clash with these old POS and lose if the leadership, like Billy Beane, does not support them.

8. Objectives for Big Data

Like many new information technologies, big data can bring about dramatic cost reductions, substantial improvements in the time required to perform a computing task, or new product and service offerings. Like traditional analytics, it can also support internal business decisions. The technologies and concepts behind big data allow organizations to achieve a variety of objectives, but most organizations focus on just one or two. The chosen objectives have implications for not only the outcome and financial benefits from big data, but also the process—who leads the initiative, where it fits within the organization, and how to manage the project.

9. Cost Reduction from Big Data Technologies

Some organizations pursuing big data believe strongly that MIPS and terabyte storage for structured data are now most cheaply delivered through big data technologies like Hadoop clusters. One company's cost comparison, for example, estimated that the cost of storing one terabyte for a year was $37,000 for a traditional relational database, $5,000 for a database appliance, and only $2,000 for a Hadoop cluster. Of course, these figures are not directly comparable, in that the more traditional technologies may be somewhat more reliable and easily managed. Data security approaches, for example, are not yet fully developed in the Hadoop cluster environment.

Organizations that were focused on cost reduction made the decision to adopt big data tools primarily within the IT organization on largely technical and economic criteria.

10. Time Reduction from Big Data

The second common objective of big data technologies and solutions is time reduction. Macy's merchandise pricing optimization application provides a classic example of reducing the cycle time for complex and large-scale analytical calculations from hours or even days to minutes or seconds. The department store chain has been able to reduce the time to optimize pricing of its 73 million items for sale from over 27 hours to just over 1 hour. Described by some as "big data analytics," this capability set obviously makes it possible for Macy's to re-price

items much more frequently to adapt to changing conditions in the retail marketplace. This big data analytics application takes data out of a Hadoop cluster and puts it into other parallel computing and in-memory software architectures. Macy's also says it achieved 70% hardware cost reductions. KeremTomak, VP of Analytics at Macys.com, is using similar approaches to time reduction for marketing offers to Macy's customers (see the, "Big Data at Macys.com," case study). He notes that the company can run a lot more models with this time savings:

Generating hundreds of thousands of models on granular data versus only 10, 20 or the 100 that we used to be able to run on aggregate data is really the key difference between what we can do now and what we will be able to do with high performance computing. Tomak also makes extensive use of visual analytics tools for his big data results, which is common with big data. Another key objective involving time reduction is to be able to interact with the customer in real time, using analytics and data derived from the customer experience, if the customer has "left the building," targeted offers and services are likely to be much less effective. This means rapid data capture, aggregation, processing, and analytics (see the "Big Data at Caesars Entertainment" case study)

11. Big data and analytics

Big data poses both opportunities and challenges for organizations. In order to extract value from big data, it must be processed and analyzed in a timely manner, and the results need to be available in such a way as to be able to effect positive change or influence business decisions. The effectiveness also relies on an organization having the right combination of people, process and technology.

By pure definition, analytics is the discovery and communication of meaningful patterns in data but for business, analytics should be viewed as the extensive use of data, statistical and quantitative analysis, using explanatory and predictive models to drive fact-based business management decisions and actions. Analytics helps to optimize key processes, functions and roles. It can be leveraged to aggregate both internal and external data. It enables organizations to meet stakeholder reporting demands, manage massive data volumes, create market advantages, manage risk, improve controls and, ultimately, enhance organizational performance by turning information into intelligence.

Analytics can identify innovative opportunities in key processes, functions and roles. It creates a catalyst for innovation and change — and by challenging the status quo; it can help to create new possibilities for the business and its customers. Sophisticated techniques can allow companies to discover root causes, analyze micro segments of their markets, transform processes and make accurate predictions about future events or customers' propensity to buy, churn or engage. it is no longer enough for companies to simply understand current process or operations with a view on improving what already exists, when there is now the capacity to question if a process is relevant to the business, or whether there is a new way of solving a particular issue. The key driver for innovation within organizations is to constantly challenge existing practices rather than consistently accept the same.

Most organizations have complex and fragmented architecture landscapes that make the cohesive collation and dissemination of data difficult. New analytic solutions are playing an important role in enabling an effective Intelligent Enterprise (IE). An IE helps to create a single view across an organization by utilizing a combination of standard reporting and data visualization.

Data from multiple source systems is cleansed, normalized and collated. External feeds can be gathered from the latest research, best practice guidelines, benchmarks and other online repositories. Use of enhanced visualization techniques, benchmarking indexes and dashboards can inform management and consumers via smartphones, laptops, tablets, etc., in-house or remotely. All companies need to start thinking about collecting and using relevant big data. Data- driven decisions can reduce inefficiency between the business, legal and IT, optimize existing information assets and address disconnects between different functions of an organization. However, it is worth noting that the best data and the most advanced analytical tools and techniques mean nothing if they are not being leveraged by people who are asking the right questions. Big data, emerging storage technology platforms and the latest analytical algorithms are enablers to business success — not a guarantee of it.

Big data can be a powerful way to identify opportunities, but when combined with traditional organizational information, the volumes of data collected can be vast and traditional storage methods can be prohibitively expensive and do not necessarily scale effectively.

Organizational information is typically historical, incomplete and inaccurate. For a forward-looking perspective (using statistical and predictive modeling) it needs to be enriched with external information.

12. The benefits and risks of big data

A key success factor for companies is the availability of relevant information at the right time. While there is no doubt that the big data revolution has created substantial benefits to businesses and consumers alike, there are commensurate risks that go along with using big data. The need to secure sensitive data, to protect private

information and to manage data quality, exists whether data sets are big or small. However, the specific properties of big data (volume, variety, velocity, veracity) create new types of risks that necessitate a comprehensive strategy to enable a company to utilize big data while avoiding the pitfalls. This should be done in a prioritized fashion so that companies can start to realize the benefits of big data in step with managing the risks.

13. Governance

Good governance is vital to the success of Big data initiatives in any business; it encompasses consistent guidance, procedures and clear management decision-making. Organizations need to ensure standard and exhaustive data capture they need not protect all the data, but they need to start sharing data with in-built protections with the right levels and functions of' the organization.

14. Benefits

The MIT Center for Digital Business states "When it comes to big data, the 'right' governance model depends on the maturity level of the organization regarding data driven decisions." It obviously also highly depends on if big data is used to create new business or to drive more sales. To unleash the power of big data, first of all data must be available and made fit for sharing. When it comes to, for example, medical data, respect for privacy and trust is inevitable. Standardization in governance structures, with an integrated combination of technical, organizational and legal measures and safeguards, will help to increase trust. This is especially important when integrating governmental, institutional "open" and company data.

An example of this in action is seen with a European association that aims to build up a big data services platform for the health sector in their local region. It's a unique collaboration between health institutions, government, education and knowledge institutions and major IT service providers, addressing both the clinical and research sides of the health sector. The core solution is comprised of a "vendor neutral hub" — a platform that works independently of vendors and data "owners" — where data can be captured, safely and durably stored, processed and distributed and finally, shared, if permitted by the data owner. This may be the case when patients are being treated by multidisciplinary teams or having treatment in various locations, or for research purposes using large sets of anonymous data. The framework offers solutions for the immediate need to access data, to substantially lower the cost of data storage and, more importantly, to do more with the rapidly growing amount of unstructured data.

15. Risks and Considerations

16. Traditional risks

There is continued regulatory pressure on companies to meet a variety of policies and laws (e.g., Basel II, MiFID, SOX). Compliance governance is an expensive and complex problem to deal with, but failing to meet regulations can mean safety risks, hefty penalties, loss of reputation or even bankruptcy. In a global and continuously and rapidly changing legal and IT landscape it is not always clear exactly what legal and regulatory compliance entails (Who is responsible? Who is liable?), or how best to translate abstract rules from laws into organizational and technical measures within a company.

Companies need to balance contradictive rules and regulations e.g., obligations based upon the US Patriot Act and the EU Data Protection Directive (and its many local implementation).

17. New risks

Managers will need to learn to embrace the evidence-based decision-making process. Organizations have to redefine their understanding of "judgments" of the outcome of big data analytics. Data can be of great value, but companies have to consider ownership and privacy issues before using big data results. In the case of medical data, it is sometimes not clear who is the owner of the data, but using the data without the right legal foundation or consent of the patient may cause big problems. Big data may bring about intellectual property issues, e.g. copyright and database rights infringements. It will be a challenge to make sure that employees are not sharing inappropriate information, or too much data outside of the organization.

18. Management

Integrating and moving data across the organization is traditionally constrained by data storage platforms such as relational databases or batch files with limited ability to process very large volumes of data, data with complex structure or without structure at all, or data generated or received at very high speeds. Organizations need to start managing data through different sources, and integrating its usefulness via a range of technologies in the market.

19. Benefits

Big data overcomes traditional restraints in a cost-effective manner and opens opportunities to ingest, store and

process data from new sources such as external social media data, market data, communications, interaction with customers via digital channels, etc. By some estimates, more than 80% of the data within organizations is unstructured and unfit for traditional processing. Using big data will enable the processing of this unstructured data and increased system intelligence which can be used to improve performance in sales, increase understanding of customer needs, reinforce the internal risk management function, support marketing initiatives and enhance fraud monitoring.

Big data capability allows organizations to integrate multiple data sources with relatively low effort in a short timeframe. Combined with a lower cost of storage per gigabyte, this enables organizations to build, for example, a federated view of customers by shifting customer data from various separate business departments into a single infrastructure, and then to run consolidated analytics and reporting on it. Big data technologies release organizations from the traditional accuracy vs. cost challenge by enabling them to store data at the lowest level of detail, keeping all data history under reasonable costs and with less effort.

20. Risks and Considerations

21. Traditional risks

There is a long implementation cycle for data warehousing and reporting solutions. Challenges over unifying data definitions are made even more complex across multiple business lines. Modeling, storage and processing challenges arise from the growing volumes of data with dynamic structures.

22. New risks

Simplified access to diverse sources of data and easy-to-ingest large amounts of information may result in increasing amount of “noise” in data and decrease in the overall level of data quality. Many new technology market players don’t have mature enterprise-ready capabilities around implementation, support, training, etc. New big data methods, architecture and volume variety impose additional risks of lack of control and governance over data, and this requires additional organizational focus. Under the context of the complex data landscape, it is especially important to establish and maintain data lineage.

Organizations may struggle with finding the right skills and building internal capabilities for handling big data as most of the technologies and methods are relatively new, and market resources are in short supply.

23. Managing Big Data

Most technical errors originate from human error rather than technical glitch. Even more reason then, for organizations to put governance programmes in place to ensure that with big data comes great reward, rather than incorrect data and unexpected costs. This is certainly easier said than done - but not impossible. Big Data is both a marketing and a technical term that refers to a valuable organization asset — information. It also represents a trend in technology that is leading the way to a new approach in understanding the world and making business decisions.

These decisions are made based on very large amounts of structured, unstructured and complex data - such as tweets, videos, commercial transactions — which have become difficult to process using basic database and warehouse management tools. The primary objective of analyzing big data is to support organizations in making better business decisions to have a positive effect on:

- Product development
- Market development
- Operational efficiency
- Customer experience and loyalty
- Market demand predictions.

When new initiatives, such as adoption of big data, are properly aligned to the business, existing governance structures can be easily adjusted to address security, assurance and a general approach to embracing new technologies. Here are seven steps that will help address risk and improve the organization’s ability to use big data so that it can meet its business objectives:

1. The concept of big data risk management is still at the infancy stage for many organizations, and data security policies and procedures are still under construction. Provide insight by monitoring all data that runs in the company, and analyze and then take action based on the results.
2. For data to be used productively, the organization needs to consider a corporate data lifecycle process. Data quality in any system is a constant battle, and big data systems are no exception. After all, big data insights are only as good as the quality of the data themselves. Certain types of’ data are business critical, while others are not. Ensure that critical processes get precedence. Figuring out the end goal is vital.
3. Do not be afraid to seek the advice and guidance of external data experts when needed. Talk to companies and cloud service integrators, and consider companies that run platforms for big data

analytics.

4. The faster and easier it is to access big data, the greater the risk to all of that valuable information. Organizations must get a proper insight into the performance of their data handling processes to minimize the risks. Do not forget to check service level agreements with clients and adapt them where necessary.
5. Make sure your organization's employees, data, networks, partners and customers are protected end-to-end. To minimize the potential for damage resulting from inaccurate or fraudulent data, organizations need to consider all the data sources they are pulling into their analyses and assess each source for vulnerabilities.
6. Ensure future-proof systems. This means that not only the right systems, but also the right tools and processes are implemented for big data today and can cope with the inevitable data growth of in the future. Companies should invest in tools that help ensure that their data is accurate, up-to-date and clean at all times. Figure out how you can incorporate what you already have.
7. Logical and physical access security controls are needed to prevent unauthorized access to sensitive and/or valuable data. Stay informed of the legislative proposals, such as those of the European Commission, and use the opportunity to employ data lifecycle best practices.

The cloud offers a new option in the storage and use of data. The proper controls must be in place to enable businesses to trust cloud service providers with their sensitive and/or valuable data. Ideally, companies start using a private cloud solution and gradually move towards a secure hybrid version. Big data should be a revenue-enhancing asset, not a revenue-haemorrhaging liability. Organizations must take a good look at their big data opportunity and take steps to manage and manipulate the growing data within the organization.

24. Implementation of Big Data

In implementing Big Data strategy, organizations are first faced with the challenge of the Technology view around big data and the business angle to it. The growing concern with big data is centered on managing increased data growth, big data in terms of business versus the concept of technology, the growing cost associated with it, needed skills and competence, and securing these data. Around these challenges are the value and opportunities with big data.

Regina Casonato, the Managing Vice President at Gartner says that organizations should have a holistic approach to big data management. They should build a strong team made up of personnel from different departments of the organization, this will help the organization build their data science lab and implement a robust big data strategy. The business value of big data comes from the fact that we can access information we never could, this information can be analyzed in new ways, it can be done faster and cheaper and then the organization can use the analyzed data to transform their businesses better by 20%.

Several sectors and businesses all over the world are using big data to transform their businesses and remain globally competitive. Governments around the world are using big data to become very effective and efficient. The open data project in places like the UK and the US has increased interaction between the government and the people, thus, their governments have become very interactive and responsive. Big data enables the government to provide relevant services to its citizens in a more organized, convenient and open platform. Big data have also impacted on the health care sector in no small measure ensuring effective health care delivery, the manufacturing, the transportation, Retail market and many others are also witnessing transformation. The major strategies on getting the best from implementing big data is first asking the write questions, understanding what analytic capacity your organization presently has, getting the right people with the right skill set and then education around it. Getting the right skills for big data analysis is seen as a major challenge in the industry as statistics show that of 4.4 million IT related jobs only 1 out of 3 of these jobs are filled. Other challenges include Software maturity, security, risk, reusability and adoption.

Mr. Massimo Cannizzo of Accenture Middle East said that the major challenges he sees around big data is extracting the right information from the noise, analyzing these information/data before they are stored. Ronald Raffensberger, CTO, IT solution Sales at Huawei Technologies added that implementing big data strategy makes businesses agile, makes for evidence based decisions, reduced risks involved with data loss, and then it improves efficiency and increased productivity in an organization.

These experts have advised organizations to take a multi-layered approach to security issues around big data and also the need for organizations to have Data science Lab with a team of skilled experts led by the Chief Data Office (CD). In addition to that organizations should consider business alignment with technology, engaging trusted partners in the implementation of big data strategies, and the need to always access their current capabilities and architecture at all time.

Companies, both large and small, learning to cope with big data are using a myriad of strategies from employing private and public cloud enabled technology to completely overhauling their view on support of data types, device types, and overall data strategy. One critical approach to managing the big data challenge is the

creation of a data governance policy. While data governance has been used by IT for years to establish control over organizations' numerous types of data, big data has unique characteristics that affect how it is governed. For example:

- The sheer volume of data. Big data tends to grow exponentially. Without defined governance policies in place, it quickly can become impossible for organizations to search, classify, and manage huge amounts of information.
- The wide variety of data. Traditionally, data governance has focused on information stored in relational databases. Big data, however, involves many different forms of information such as non-relational databases and other types of unstructured data, like information generated by social media.
- Combining knowledge of big data challenges with a well-thought out data governance policy, however, enables organizations to increase the value of their information and transform it into a highly available view of a company's legacy knowledge and intellectual property. When users enjoy access to data that is consistent, accurate and available when they need it, that translates into better business decision-making and faster responses to market trends and customer needs.

25. Neglecting Data Governance - A Risky Proposition

It is not uncommon for organizations to view data creation as a separate activity from data governance. That approach can be risky, however, if data governance falls by the wayside and is forgotten or neglected. Dangers associated with ungoverned big data include:

- *Data that is difficult to search and analyze.* When organizations do not have a coordinated approach to data management and governance, business users often have a hard time finding the information they need to make decisions and they may not trust the validity of the data they can access. This problem is compounded as data volumes increase. On average a company's data doubles every 18 months. This means that the time for getting control over this information is now.
- *Lack of compliance with regulations or internal controls.* Regulations such as the Federal Rules of Civil Procedure (FRCP), the Federal Rules of Evidence (FRE), the Health Insurance Portability and Accountability Act (HIPAA), Sarbanes-Oxley (SOX), and others are complex because organizations must proactively demonstrate compliance with standards related to electronically stored information, if steps to ensure compliance arc not accumulated in a data governance policy and then followed, compliance issues can arise.
- *Potential or financial and reputational damages.* If organizations do not have a clear idea of the lifecycle for different types of information and the systems where information resides, the risk of data breaches and theft increases. That can result in fines, as well as reputation damage. Unclear retention policies for recent types of information. Storing big data indefinitely can quickly become a costly matter. A robust data governance policy should define how long different types of data are retained.

26. Harnessing Big Data's Potential through Data Governance

Big data, when used wisely, can deliver tremendous value to organizations. The importance of data governance in this equation is gaining visibility. A recent report from the Institute for Health Technology Transformation, for example, indicated that a standardized format for data governance is essential for healthcare organizations to leverage the power of big data. The authors indicate that the first and most critical priority is to develop a carefully structured framework for enterprise data governance. Whether you are developing a policy from scratch or enhancing an existing one, here are four ways to strengthen your data governance model:

1. **Develop a Data Governance Strategy:** This should be consistent with the overarching business strategy and should include guiding principles for how big data will be governed. This means deciding who owns different types of information, who can access it, and how data is used. Key issues to consider include data quality, regulatory requirements, security and privacy, and information lifecycle management.
2. **Use a Cross-Functional Approach:** This is particularly important for compliance purposes. Data and information systems often touch many different departments and no one individual has a complete view. A cross-functional team is best positioned to develop a holistic view of the organization's big data, including controls, documentation and auditable proof of compliance.
3. **Make Decisions About Data-Related End-Of-Life Issues:** All aspects of the data lifecycle arc relevant when addressing data governance, but end-of-life issues should not be overlooked. One standard retention schedule won't fit all needs. Different types of data will have different requirements for retention periods. Organizations may elect to archive data in order to enhance application performance.
4. **Consider how Technology can Support Data Governance Efforts:** With big data, organizations must estimate how quickly data volumes will grow, as well as how costly it will be to store information. Data governance policies should define when information is moved to archiving systems which offer less expensive forms of storage, while maintaining easy access for end users and taking performance loads off of other

applications.

Big data has great potential for helping organizations do business better, but simply having the data is not enough. To derive the greatest value and minimize risks, a data governance infrastructure is essential of IT governance and is a critical component of your IT strategy. You need a level of confidence in how any data is handled to make sure your organization is not at risk of a nasty, often public data exposure. That extends to privacy for all your data, including big data sets that are increasingly becoming part of the mainstream IT environment.

Privacy is also related to issues like data monetization. If the data that you have is not secure, high-quality or fit for purpose, can you trust the monetary value placed on that data? And, as the amount of data grows, do you have a strategy for larger privacy efforts, or big data privacy, in your organization?

The recommended approach is to blend your business rules and IT rules. If you can accomplish this collaborative effort through the use of governance solutions to establish a big data privacy framework within your IT environment, it would be to the advantage of that organization.

27. Conclusion

Even though it hasn't been long since the advent of big data, these attributes add up to a new cra it is clear from our research that large organizations across industries are joining the data economy. They are not keeping traditional analytics and big data separate, but are combining them to form a new synthesis. It means change in skills, leadership, organizational structures, technologies, and architectures. it is perhaps the most sweeping change in what we do to get value from data since the 1980s.

It is important to remember that the primary value from big data comes not from the data in its raw form, but from the processing and analysis of it and the insights, products, and services that emerge from analysis. The sweeping changes in big data technologies and management approaches need to be accompanied by similarly dramatic shifts in how data supports decisions and product/service innovation. There is little doubt that analytics can transform organizations, and the firms that utilize big data will seize the most value. Big data is often portrayed as a potential opportunity for business firms and other types of organizations. Firms skilled in the use of big data, however, are already using their analytics capability to create strong competitive advantages. Those firms are now making many important decisions in real time and are able to keep pace with the rapidly changing environments of the digital age. Their experiences can be instructive for other interested organizations because the organizational, managerial, and cultural changes required by a big data analytics capability are considerable.

References

According to a repon. published by Cisco Systems, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update," February 6, 2013.

Andrew Leonard, "How Netflix Is Turning Viewers into Puppets," February 1, 2013, http://www.salon.com/2O 13/02/01 /how_netflix_is_turning_viewers_into puppets/SAS 2013 Big Data Survey, page 1: http://www.sas.com/resources/whitepaper/wp_5 8466 .pdf

Banjo S. 2013 (June 19). Nike takes a page from Apple's playbook. The Wall Street Journal, Technology Section.

Big Data: The Next Frontier for innovation, Competition, and Creativity," McKinsey Global Institute, 2011.

Chandler AD. 1962. Strategy and Structure. MIT Press, Cambridge, MA.

Critical Design Challenges. Jossey-Bass, San Francisco, CA.

Gaibraith JR. 20 14.Designing Organisations.3rd edition.Jossey-Bass, San Francisco, CA.

http://www.cisco.co,/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white paper c 1 11http://wikibon.org/wiki/v/Financial Comparison of Big Data MPP Solution and D ata Warehouse Appliance SAS 2013 Big Data Survey, page 4: http://www.sas.com/resources/whitepaper/wp 58466.pdf

Joel Schectman, "Ad Firm Finds Way to Cut Big Data Costs," Wall Street Journal ClO Journal website, February 8, 2013, http://blogs.wsj.com/cio/201 3/02/08/ad-firm-find s-way-to-cut-big-data-costs/

Kates A, Galbraith JR. 2010. *Designing Your Organization*: Using the Star Model to Solve 5

Kerem Tomak, in "Two Expert Perspectives on High-Performance Analytics," Intelligence Quarterly (a SAS publication), 2d quarter 2012, p. 6.

Morozov E. 2011. The Net Delusion: *The Dark Side of Internet Freedom* Penguin, London, U.K.

Morrissey B. 2010 (July 15). How Old Spice ruled the real time web. *AdWeek.*

Peter Evans and Marco Annunziata, "industrial Internet: Pushing the Boundaries of Minds and Machines," GE report, Nov. 26, 2012. www.ge. com/docs/chapters/industrial internet. pdf

Piskorski M, Johnson R. 2012. Social strategy at Nike. Harvard Business School Case 9-712-484.

Schmidt E, Cohen J. 2013. The New Digital Age: *Reshaping the Future of People, Nations and Business.* Knopf New York, NY.

Shetty B, Wind J. 2013 (February 15). Advertisers should act more like newsrooms. Harvard Business Review Blog Network.*The Economist*. 2013. Less guff, more puff May 18; 69-70.

Tom Vanderbilt, "Let the Robot Drive: The Autonomous Car of the Future is Here," Wired, January 20, 2012, http://www.wired.com/magazine/20/ff_autonomouscarsf

Tushman M, Anderson P. 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly* 3 1: 439-465.

Winsor J, Wind J. 2013 (May 9). The end of traditional ad agencies. Harvard Business Review Blog Network. New Vantage Partners, "Big Data Executive Survey: Themes and Trends," 2012