

SURVEY OF BIG DATA ANALYTICS: PROVIDING UBIQUITOUS BENEFITS

¹I. A. Braimoh and ²A. B. Olaniyi

¹Department of Computer Sciences, Federal College of Education, Okene

²Department of Computer Science, Federal Polytechnic, Idah

Corresponding Author: abdulbramoih@yahoo.com

Abstract: Big Data and analytics are becoming a new technology focus both in science and in industry, thereby motivate technology shift to data centric architecture and operational models. While the quantum of large datasets is significantly rising, big data growth has brought about many challenging issues demanding remediation. This paper discusses the trends of big data and its applications in various disciplines, methods of analysis, benefits and its security challenge. For each component, detailed overviews of relevance and its applications were made as well as possible risks and benefits on general scale and generally applied to Nigeria landscape. Challenges of big data and analytics are similarly outlined while proffering big data technology adoption to the Nigeria business arena with a view to solving the gamut of her problems. Finally, this survey is concluded with some recommendations and future directions.

Keywords: Big Data, Technology, Analytic, Real-Time, Business, infrastructure.

INTRODUCTION

The world today is experiencing an exponential growth of the volume of data generated, processed and stored in a manner that require an alternative technology to managing the data scenarios. There is the possibility of harnessing big data potentials in enhancing the society, infrastructure, governance and stimulating the nations economies. In similar manner, according to André *et al.* (2015) advances in technology have resulted in the proliferation of data with different formats from most various domains (e.g. health care, banking, government or logistics) and sources (e.g. sensors, social networks or mobile devices) to various platforms of sizes, volumes and orientations. Government agencies, businesses, cooperate and private organisations are regularly generating enormous quantum of data too big to be

processed and analysed by Data Warehouses (DWs) only.

Big data refers to humongous volumes of data that cannot be processed effectively with the traditional applications that exist (Avantika, 2016). Big data typically is a massive and complex datasets, which are made of a variety of data structures (structured, semi-structured and unstructured data) from a multitude of sources (André *et al.*, 2015) that need more real-time analysis. Big Data Analytics refers to a set of procedures and statistical models to extract information from a large variety of data sets (Raghavendra *et al.*, no date). In addition, big data also brings about new opportunities for discovering new values, helps us to gain an in-depth understanding of the hidden values, and also incurs new challenges, for instance, how to effectively organize and manage such

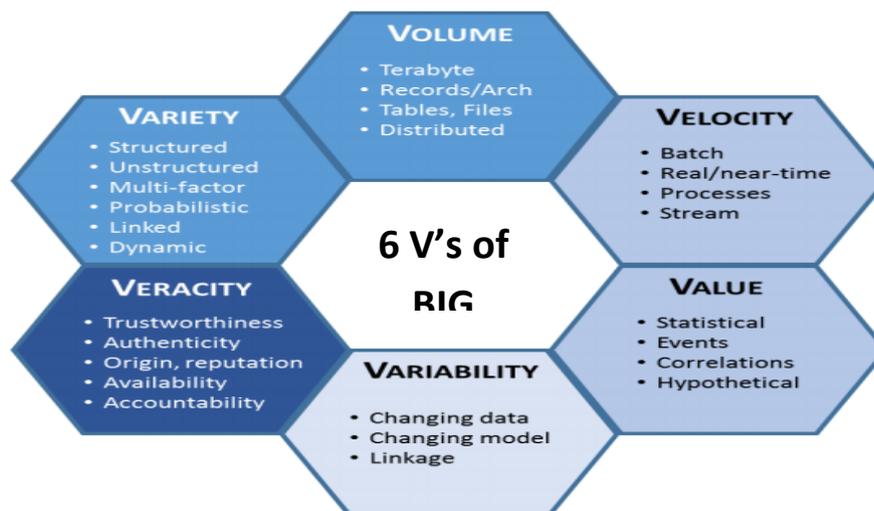


Fig 1: A pictorial representation of 6 V's of Big Data
Source: Demchenko (2013)

datasets (Chen *et al.*, 2014). Big Data has been found to be characterised (see figure 1) by three Vs: *volume* (amount of data), *velocity* (speed of data in and out) and *variety* [kinds of data types and sources – availability]

(Beyer and Laney, 2012). Other added characteristics of Vs include: *variability* (data changeability), *veracity* (data reliability), *value* (data worth) and *visualisation* [readable and accessible] (MCnulty-Holmes, 2014).

According to Kuketz (2017), the common characteristics of big data insights include addressing speed and scalability, mobility and security, flexibility and stability while integrating both structured and unstructured data. Big data growth has brought about many challenging issues demanding remediation. Advances in Information Technology (IT) for instance provided more data generation within and between organisations. The niche created by the advent of computer printers to facilitate the voyage of *paperless society* in the 90's brought the spontaneous leap in the mass production of data storage media (DISK, TAPES, CDROM, FLASH etc) with backup capabilities revolution. During this era systems were used to extract, transform and load data into giant data warehouses that had business intelligence (BI) solutions built over them for reporting. Periodically, all the systems would backup and combine the data into a database where reports could be run and everyone could get insight into what was going on. According to www.martech.zone (2017), the problem was that the database technology simply could not handle multiple continuous streams of data. And reporting tools were similarly lacking as a consequence it could not handle anything other than a relational query on the back-end. These defects created a scenario that seems to jeopardise growth in data interpretation, real-time access and provision of secured data warehouse and networks. More so, Oloruntoba and Sodiya (2015) noted other challenges as including data privacy, security of data, data analysis, data retrieval and lack of IT infrastructure, skills and expertise. Advances in Information and Communication Technology (ICT) came with an unprecedented explosion in computer networks and mobile networks interface facilities. Large volumes of data Backup domicile in offices, room, warehouse, banks were transferred or accessed remotely. Social media also were highly contributing to the worsening of this situation. Facebook, for instance, has an average of 4.75 billion pieces of content shared among friends every day (Andre' *et al.* 2015). According to Chen *et al.* (2014), the rapid growth of cloud computing and the Internet of Things (IoT) further promote the sharp growth of data. Data representation aims to make data more meaningful for computer analysis and user interpretation. Nevertheless, an improper data representation will reduce the value of the original data and may even obstruct effective data analysis (Chen *et al.*, 2014)

Adopting Big Data based technologies therefore, not only mitigates the problems presented above, but also opens new perspectives that allow extracting value from Big Data. Big Data-based technologies are being applied with success in multiple scenarios (Agrawal *et al.*, 2011). The history of previous trends in IT investment and innovation and its impact on competitiveness and productivity strongly suggest that big data can have a similar power, namely the ability to transform our lives. The same preconditions that allowed previous waves of IT-enabled innovation to power productivity, i.e., technology innovations followed by the adoption of complementary management innovations, are in place

for Big Data, and we expect suppliers of Big Data technology and advanced analytic capabilities to have at least as much ongoing impact on productivity as suppliers of other kinds of technology (McGuire *et al.*, 2012).

The study aims at surveying the trends of big data and its applications in various disciplines, its benefits and security challenges. More specifically the objectives are to: present a systematic introduction of big data initiative, identify emerging cases deployed in different critical sectors, identify the security and risk challenges of the use of big data services and outline recommendations for strengthening the security of big data infrastructures and services in Nigeria.

LITERATURE REVIEW

In the work of Singh and Reddy (2014), an attempt was made to survey different hardware platforms available for big data analytics and assesses the advantages and drawbacks of each of these platforms based on various metrics such as scalability, data I/O rate, fault tolerance, real-time processing, data size supported and iterative task support. This study failed to provide insight on how security challenges are mitigated in hardware platforms and software frameworks.

Also, in the work of Oloruntoba and Sodiya (2015), where differential privacy; a noise perturbation methodology that revolve around hiding the presence of an individual in datasets was proposed. In their proposal, any query on big data analytics datasets is returned with perturbed response with generated noise from Laplace distribution which is hoped to enhance privacy of big data in the cloud. Though, the work is a proposal, however, possible risk tendencies that could be embedded in the system implementation and deployment were not highlighted.

Chen *et al.* (2014), made a survey on big data and reviewed related technologies, such as cloud computing, Internet of Things, data centres and Hadoop. They laid emphasis on the four phases of the value chain of big data, i.e., data generation, data acquisition, data storage, and data analysis. Their discussion was aimed at providing a comprehensive overview and big-picture to readers of this exciting area.

METHODS OF BIG DATA ANALYSIS

On a general outlook Big Data has four layered abstraction model as shown in figure 2, the layers from the bottom up are *physical layer*, *data layer*, *computing layer* and *data analytics layer* respectively. Physical layer takes care of the data organization over several distributed data storage, high speed networks, and partitioned cluster of nodes. Data layer addresses the global namespace for the data access and logical expansion of the data, without knowing the underlying physical layer structure. Compute layer offers several computing methodologies and analytic offers several technologies for analysis of the data for decision making. Methods adopted to address big data problem is stemmed addressing each layer in the abstraction model.

A critical area that demands attention is the compute and programming layers.

Various programming models exist in Big Data technologies which include: distributed infrastructure such as cloud and multi-core computing, distributed storage such as Google file system (GFS), Hadoop Distributed file system (HDFS), Data Model/Indexing such as NoSQL Dbase, Neo4, BigTable etc.

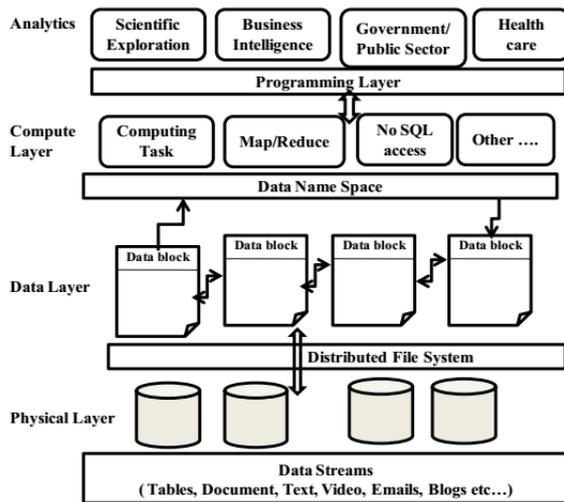


Figure 2. Big Data Abstraction

models such as MapReduce, Spark, NAIAD According to Press (2016) some of the hottest big data analysers based on Forrester’s analysis include Predictive Analytics, NoSQL Databases, Search and Knowledge Discoveries, Stream Analytics, In-Memory Data Fabric, Distributed Data Store, etc.

Figure 3 is MapReduce programming model which is the basic data processing scheme used in Hadoop that includes breaking the entire task into two parts, known as mappers and reducers. At a high-level, mappers read the data from HDFS, process it and generate intermediate results to the reducers. Reducers are then used to combine the intermediate results to generate the final output which is again written to HDFS. It is a two stage fixed

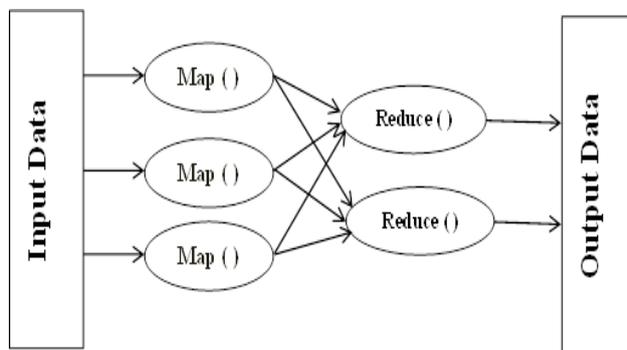


Figure 3. MapReduce Programming Model

data flow. Spark programming model is a next generation paradigm for big data processing (Singh and Reddy, 2014) using more flexible data flow designs

using the Directed Acyclic Graph (DAG) and overcome the disk/IO and iteration limitations of Mapreduce plus increase performance in speed and in-memory computations.

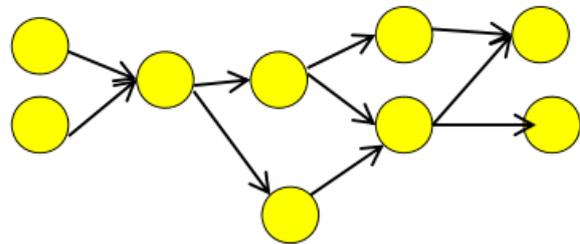


Fig 4. Spark Model using DAG data flow

APPLICATION IN VARIOUS UBIQUITOUS DISCIPLINES

André *et al.* (2015) had stated that big data-based technologies are being applied with success in multiple scenarios which was conceptualised into: (1) e-commerce and marketing, (2) government and public health, (3) transportation, industry and surveillance, with improve real-time estimates and smart use of resources. In a specific detailing of these application based scenarios, the following are some considerations among others. According to Analytics Vidhya Content Team (2015) and other authors, here are the outlines:

Internet Search: Many existing search engines like Google, Yahoo, Bing, Ask, AOL, Duckduckgo make use of data analytic algorithms to deliver the best result for many searched query in fraction of seconds. Because, Google processes more than 20 petabytes of data every day.

Airline Route Planning

Airline Industry across the world is known to bear heavy losses owing to many factors. The situation is critical with high rise in maintenance cost, overhead cost and the need to offer significant discounts to customers. With the deployment of big data and data analytic to identify the strategic areas of improvements, the airline companies according to Analytics Vidhya Content Team (2015) have been found to:

- (1) Predict flight delay
- (2) Decide which class of airplanes to buy
- (3) Effectively drive customer loyalty programs

Fraud and Risk Detection

Over the years, banking business learned to divide and conquer data via customer profiling, past expenditures and other essential variables to analyse the probabilities of risk and default. More so, real-time monitoring that provides alerts on regular intervals tends to come in as a rescue alternative. This problem tends to increase with big data given the volume and velocity of data stream (Cloud Security Alliance, 2012). Finance (Credit Risk, Fraud), Human Resources and many other tasks are easily accomplished using data analytic in these disciplines (Avantika, 2016).

Delivery logistics

Logistic companies like DHL, FedEx, EMS Speed Post and UPS have been said to use big data and analytic technology to improve their operational efficiency. Using big data have enabled these companies discovered the best routes to distribute, the best suited time to deliver and the best mode of transport to opt. These approaches are found to show the way to quick arrival, cost and time efficiency and many other benefits. Furthermore, the data that these companies generate using the GPS installed devices provides them a lot of possibilities to explore using data analytics everyday (Analytics Vidhya Content Team, 2015).

Price Comparison Websites

At a basic level, these websites are being driven by huge amount of data obtained by using APIs and RSS Feeds. Using these websites bring the convenience of comparing the price of a product from multiple vendors at one location. In present time, price comparison website can be found in almost every field such as technology, hospitality, automobiles, cosmetics, appliances etc.

Gaming

EA Sports, Zynga, Sony, Nintendo, Activision-Blizzard have led gaming experience to the next level using data analytic. Games are now designed using machine learning algorithms which improve/upgrade themselves as the player moves up to a higher level. In motion gaming also, your opponent (computer) analyses your previous moves and accordingly shapes up its game.

Speech Recognition: Some of the best example of speech recognition products is Google Voice, Siri, Cortana etc. Using speech recognition feature, even if you aren't in a position to type a message, your life would not stop. Simply speak out the message and it will be converted to text.

Image Recognition

Uploading images with friends on Facebook and with the suggestions to tag your friends is an area of enterprise. This automatic tag suggestion feature uses face recognition algorithm. Similarly, while using whatsapp web, you scan a barcode in your web browser using your mobile phone. In addition, Google provides you the option to search for images by uploading them. It uses image recognition and provides related search results.

Digital Advertisements (Targeted Advertising and re-targeting): Starting from the display banners on various websites to the digital bill boards at the various city locations – almost all of them are decided by using data analytic algorithms (Analytics Vidhya Content Team, 2015). This is the reason why digital ads have been able to get a lot higher CTR than traditional advertisements. They can be targeted based on users past behaviour. This is the reason why we see ads of analytics trainings while others see ads of apparels in the same place at the same time.

Recommender Systems

Recommender systems are a subclass of information filtering systems that are meant to predict the preferences or ratings that a user would give to a product (Bhargav, 2016). They not only help you find relevant products from billions of products available with them, but also add a lot to the user experience. A lot of companies have passionately used this engine/system to promote their products/suggestions in accordance with user's interest and relevance of information. Internet giants like Amazon, Twitter, Google Play, Netflix, Linkedin, imdb and many more uses this system to improve user experience. The recommendations are made based on previous search results for a user.

Miscellaneous

Using data analytic, the marketing departments of companies decide which products are best for up selling and cross selling, based on the behavioural data from customers. In addition, predicting the wallet share of a customer, which customer is likely to churn, which customer should be pitched for high value product and many other questions can be easily answered by big data technology.

Communications: Gaining new subscribers, retaining customers, and expanding within current subscriber bases are top priorities for telecommunication service providers. The solutions to these challenges lie in the ability to combine and analyse the masses of customer generated data and machine generated data that is being created every day (Avantika, 2016). ***Big Data for Retail:*** Brick and Mortar or an online e-retailer, the answer to staying the game and being competitive understands the customer better to serve them. This requires the ability to analyze all the disparate data sources that companies deal with every day, including the weblogs, customer transaction data, social media, and store branded credit card data, and loyalty program data (Avantika, 2016).

Healthcare: The main challenge for hospitals with cost pressures tightens is to treat as many patients as they can efficiently, keeping in mind the improvement of quality of care. Instrument and machine data is being used increasingly to track as well as optimise patient flow, treatment, and equipment use in the hospitals. It is estimated that there will be a 1% efficiency gain that could yield more than \$63 billion in the global health care savings (Avantika, 2016).

Travel: Data analytics can optimise the buying experience through the mobile/web log and the social media data analysis. Travel sights can gain insights into the customer's desires and preferences. Products can be up-sold by correlating the current sales to the subsequent browsing increase browse-to-buy conversions via customised packages and offers. Personalized travel recommendations can also be delivered by data analytics based on social media data (Avantika, 2016).

Gaming: Data Analytics helps in collecting data to optimise and spend within as well as across games. Game companies gain insight into the dislikes, the relationships, and the likes of the users (Avantika, 2016).

Energy Management: Most firms are using data analytic for energy management, including smart-grid management, energy optimisation, energy distribution, and building automation in utility companies. The application here is centred on the controlling and monitoring of network devices, dispatch crews and manage service outages. Utilities are given the ability to integrate millions of data points in the network performance and lets the engineers use the analytics to monitor the network (Avantika, 2016).

BENEFITS OF BIG DATA

According to Douglas (2015), big data solutions offer cloud hosting, highly indexed and optimised data structures, automatic archival and extraction capabilities, and reporting interfaces have been designed to provide more accurate analyses that enable businesses to make better decisions. Better decisions that reduce costs and increase marketing and sales effectiveness can be achieved by companies which in effect help to reduce the risk. This also implies that companies will have far more intelligence at their disposal to make accurate decisions and predictions on business operations. Nigeria Companies will in effect be combining marketing, sales, customer data, transactional data, social conversations and even external data like stock prices, weather and news to identify correlation and causation statistically valid models to help them make more accurate decisions. In a brief summary, Douglas (2015) puts the benefits in this order:

- (1) *Big Data is Timely* – 60% of each workday, knowledge workers spend attempting to find and manage data.
- (2) *Big Data is Accessible* – Half of senior executives' report that accessing the right data is difficult.
- (3) *Big Data is Holistic* – Information is currently kept in silos within the organization. Marketing data for example, might be found in web analytics, mobile analytics, social analytics, CRMs, A/B Testing tools, email marketing systems and more, each with focus on its silo.
- (4) *Big Data is Trustworthy* – 29% of companies measure the monetary cost of poor data quality. Things as simple as monitoring multiple systems for customer contact information updates can save millions of dollars.
- (5) *Big Data is Relevant* – 43% of companies are dissatisfied with their tools ability to filter out irrelevant data. Something as simple as filtering customers from web analytics can provide a lot of insight into acquisition efforts.
- (6) *Big Data is Secure* –The secure infrastructures being built by big data hosting and technology partners can save the average company 1.6% of annual revenues.
- (7) *Big Data is Authoritative* – 80% of organisations struggle with multiple versions of the truth

depending on the source of their data. By combining multiple, vetted sources, more companies can produce highly accurate intelligence sources.

- (8) **Big Data is Actionable** – Outdated or bad data results in 46% of companies making bad decisions that can cost billions.

Businesses are using the power of insights provided by big data to instantaneously establish who did what, when, where and the effective enterprise decision-making. To obtain the maximum business impact, this process also requires a precise combination of people, process and analytic tools. Some of the potential business benefits from implementing an effective big data insights methodology include (Kuketz, 2017):

- (1) Timely insights from the vast amounts of data. This includes those already stored in company databases, from external third-party sources, the Internet, social media and remote sources.
- (2) Real-time monitoring and forecasting of events that impact either business performance or operation
- (3) Ability to find, acquire, extract, manipulate, analyse, connect and visualize data with the tools of choice (SAP HANA, SAP Sybase, SAP Intelligence Analysis for Public Sector application by Palantir, Kapow and Hadoop).
- (4) Convergence of the BDI solution for variety with the speed of SAP HANA for velocity
- (5) The capability of Hadoop for volumes to manage vast amounts of data, in or out of the Cloud, with validation and verification.
- (6) Identifying significant information that can improve decision quality
- (7) Mitigating risk by optimizing the complex decisions of unplanned events more rapidly

According to research from the McKinsey Global Institute (MGI) and McKinsey & Company's Business Technology Office, the sheer volume of data generated, stored, and mined for insights has become economically relevant to businesses, government, and consumers (McGuire *et al.*, 2012). For example, some retailers embracing big data see the potential to increase their operating margins by 60 per cent (McGuire *et al.*, 2012), with the following observations:

1. *Big Data can unlock significant value by making information transparent.* There is still a significant amount of information that is not yet captured in digital form, e.g., data that are on paper, or not made easily accessible and searchable through networks. This effort represents a significant source of inefficiency.
2. *As organisations create and store more transactional data in digital form; they can collect more accurate and detailed performance information on everything from product inventories to sick days and therefore expose variability and boost performance.* In fact, some leading companies are using their ability to collect and analyse big data to conduct controlled experiments to make better management decisions.

3. Big Data allows ever-narrower segmentation of customers and therefore much more precisely tailored products or services.

4. *Sophisticated analytics can substantially improve decision-making, minimise risks, and unearth valuable insights that would otherwise remain hidden.*

5. *Big Data can be used to develop the next generation of products and services.*

For instance, manufacturers can use data obtained from sensors embedded in products to create innovative after-sales service offerings such as proactive maintenance to avoid failures in new products.

(1) *Fraud can be detected the moment it happens and proper measures can be taken to limit the damage.* The financial world is very attractive for criminals. With a real-time safeguard system, attempts to hack into an organisation are notified instantly. IT security department can take immediate appropriate action.

(2) *Cost savings:* The implementation of a Real-Time Big Data Analytics tools may be expensive, it will eventually save a lot of money. There is no waiting time for business leaders and in-memory databases (useful for real-time analytics) also reduce the burden on a company's overall IT landscape, freeing up resources previously devoted to responding to requests for reports.

(3) *Better sales insights, which could lead to additional revenue.* Real-time analytics tell exactly how sales are doing and in case an internet retailer sees that a product is doing extremely well, it can take action to prevent missing out or losing revenue.

(4) *Keep up with customer trends:* Insight into competitive offerings, promotions or your customer movements provides valuable information regarding coming and going customer trends. Faster decisions can be made with real-time analytics that better suit the (current) customer.

THE RISKS ASSOCIATED WITH BIG DATA TECHNOLOGIES

Guillermo (2014) and [www.computer weekly.com](http://www.computerweekly.com) reiterated that though splendid is the adoption of big data technology, however, there abound some associated risk with its usage. These, he itemise as follows:

New technology: This is a new technology for most organisations. And many of these organisations are keying in to adopt the technology for their business outfits.

Vulnerability: Any technology that is not well understood will introduce new vulnerabilities.

Big Data implementations: Typically include open source code, with the potential for unrecognised back doors and default credentials.

Attack surface of nodes: The attack surface of the nodes in a cluster may not have been reviewed and servers adequately hardened.

User authentication: User authentication and access to data from multiple locations may not be sufficiently controlled. An additional problem is that software

commonly used to store big data, such as Hadoop, does not always come with user authentication by default.

User's Privacy: One of the biggest challenges for big data from a security point of view is the protection of user's privacy. Adequate access control mechanisms will be key in protecting the data. Guillermo (2014) suggested that a better approach is to protect the information using encryption that only allows decryption if the entity trying to access the information is authorised by an access control policy.

Regulatory: Regulatory requirements may not be fulfilled, with access to logs and audit trails problematic. Anyone using third party cloud providers to store or process data will need to ensure that the providers are complying with regulations. There is significant opportunity for malicious data input and inadequate data validation ([www.computer weekly.com](http://www.computerweekly.com)).

Encryption Problem: According to Guillermo (2014) when storing the data, organizations may face the problem of encryption. Data cannot be sent encrypted by the users if the cloud needs to perform operations over the data. A solution for this is to use "Fully Homomorphic Encryption" (FHE), which allows data stored in the cloud to perform operations over the encrypted data so that new encrypted data will be created. When the data is decrypted the results will be the same as if the operations were carried out over plain text data. *Real-time security monitoring:* Real-time security monitoring is also a key security component for a big data project. It is important that organisations monitor access to ensure that there is no unauthorised access. It is also important that threat intelligence is in place to ensure that more sophisticated attacks are detected and that the organisations can react to threats accordingly.

Policy Approach: Organisations should run a risk assessment over the data they are collecting. They should consider whether they are collecting any customer information that should be kept private and establish adequate policies that protect the data and the right to privacy of their clients. If the data is shared with other organisations then it should be considered how this is done. Deliberately released data that turns out to infringe on privacy can have a huge impact on an organisation from a reputational and economic point of view.

ADAPTING BIG DATA TO SOLVING GAMUT OF NIGERIA PROBLEMS

According to www.forbes.com (2013), the following subjects were proffered as means to solving scale of problems in a society with desire for growth, of which Nigeria is one.

The New Consumer Conversation: At this moment, big data is enabling a new consumer conversation that uses natural language processing to read and evaluate consumer responses. Big data is often thought as a

purely analytical effort, but sometimes it can be most useful improving communication and transparency. Nigeria stands to gain if utilised for its anticorruption crusade.

Improving Public Safety: We see police men and women out there on our streets and high ways. Yet they have many limitations to providing an all round safety and security for the teeming populace. By combining chatter gleaned from social media with predictive crime analytics, police departments in Nigeria can become even more effective.

Increase Visibility Across the Enterprise: Another advantage to big data is that it can help the entire enterprise work as one functional unit. There is no longer any need for data silos for different functions such as marketing, finance, logistics, etc. Big data techniques allow us to pull work from the same data set and pull out what we need. This is already being put to use by many companies and agencies in Nigeria.

Build Collaboration: One of the great management challenges is helping one group communicate effectively with the other. This is achieved by giving people at the point of service access to expert level information, which extends the entire organization's benefits. The best way to approach big data is not to try to build a better system, but to build a better enterprise. This is a way of how some businesses and organizations could use big data to solve problems in companies and agencies in Nigeria.

Netflix: The popular on-demand Internet streaming service could use big data to vary its pricing. Brandeis University economist Benjamin Shiller combined demographics and Web browsing to develop accurate predictions on how much specific customers would pay for a subscription. These can be adopted by many small-scale businesses in Nigeria.

Universal Postal Service (UPS): UPS, the biggest package shipping company in the world, has been using big data for years to not only increase profits, but to become more efficient and environmentally friendly. For example, "delivery routes were designed to minimise left turns, which require vehicles to wait at intersections for oncoming traffic to clear before proceeding." By doing small measures like this, along with sensors in vehicles to predict when parts break down, UPS is able to save \$50 million per year in fuel, maintenance, and time. Data also helped UPS save "more than 1.5 million gallons of fuel and reduced carbon dioxide emissions by 14,000 metric tons." Postal service providers in Nigeria can adopt this technology to optimise routes, coverage, time and fuel cost.

City Information Gathering: Information about residents, road traffic, ecology, challenge areas, trouble spots, monitoring, fire out-breaks, noise pollution, natural disaster and rescue missions and where catastrophic fires were likeliest to occur are gathered for data analytics.

This is an excellent big data application that could be adopted in Nigeria and will go a long way to mitigate myriads of social vices plaguing the Nigeria society and governance.

CONCLUSION

Big data and data analytics based technologies are being applied to various human endeavours with success in multiple scenarios. Adopting these technologies therefore will not only mitigate many of the problems confronting various sectors, but will also open new perspectives that allow extracting value from big data and data analytics. Nigeria's economy, government and her people stand to gain from the deployment of this technology as it will provide employment, business intelligence and infrastructure for the profit of all.

We have presented a survey trend of big data and its applicability in various disciplines. By this we have been able to realise an enormous availability of data in various formats requiring proper storage, procession and an alternative technology to managing the data scenarios.

Also, we have identified emerging cases where the technology is deployed in different critical sectors with successes in multiple scenarios like in e-commerce and marketing, government and public health, and transportation, industry and communication, with real-time improved estimates and smart use of resources.

Furthermore, security and risk challenges involved in the use of big data services were identified and organisations must make sure that they isolate sensitive information and use third party cloud providers to ensure that the providers are complying with regulations.

Finally, recommendations for strengthening the security of Big Data infrastructures and services were made with mention of government policy adoption, access control policy and periodic security audit and agreed penalties for breach of terms.

RECOMMENDATIONS

In view of the discussion made and the possible implementation of big data technology in Nigeria, here is an outline of the recommendations as related by www.computerweekly.com (2017):

Adapting governance frameworks to handle big data security issues and risk: If an adequate governance framework is not applied to big data then the data collected could be misleading and cause unexpected costs.

Best practices from a security perception: There is not a list of best practices yet that are widely recognized by the security community, however, a number of general security recommendations that can be applied to big data include:

- (1) *Vet the cloud providers:* providers to have adequate protection mechanisms, periodic security audits and agree penalties where standards are not met.
- (2) *Create an adequate access control policy:* Create policies that allow access to authorised users only.

- (3) *Protect the data*: Encryption should be used accordingly to ensure no sensitive data is leaked.
- (4) *Protect communications*: ensure data confidentiality and integrity.
- (5) *Use real-time security monitoring*: Access to the data should be monitored. Threat intelligence should be adopted.

REFERENCES

- Agrawai, D., Das, S. and Abbadi, A. E.** 2011. Big Data and Cloud Computing: Current State and Future Opportunities. DBT, Uppsala, Sweden
- Analytics Vidhya Content Team** 2015. Applications data science. Accessed from <https://www.analyticsvidhya.com/blog/2015/09/applications-data-science/>
- Andre, R., Afonso, S. and Da-Silver, A. R.** 2015. Data Modeling and Data Analytics: A Survey from a Big Data Perspective. *Journal of Software Engineering and Applications*. <http://www.scirp.org/journal/jsea>. 8: 617-634
- Avantika, M.** 2016. Data science versus big data versus data analytics. Accessed from <https://www.simplilearn.com/data-science-vs-big-data-vs-data-analytics-article>.
- Beyer, M. A. and Laney, D.** 2012. The Importance of “Big Data”: A Definition. Gartner. <https://www.gartner.com/doc/2057415>
- Bhargav, R.** 2016. 20 Most popular data science interview questions. Accessed from <https://www.simplilearn.com/data-science-interview-questions-article>.
- Cloud Security Alliance (CSA)** 2012. Top ten big data security and privacy challenges. Accessed from <http://www.cloudsecurityalliance.org>.
- Chen, M., Mao, S. and Liu, Y.** (2014). Big data: a survey. *Mobile Network Applications*. DOI 10.1007/s11036-013-0489-0. 19:171–209.
- Demchenko, Y.** 2013. Big data standisation in industry and research. Euro Cloud Symposium ICS Track: Standards for Big Data in the cloud, Luxembourg.
- Douglas, O.** 2015. IBM Systems Blog: In the Making accessed from <https://www.ibm.com/blogs/systems/author/douglas-flaherty/>
- Duncan, A. D.** 2014. Focus on the “Three Vs” of Big Data Analytics: Variability, Veracity and Value. Gartner. <https://www.gartner.com/doc/2921417/focus-vs-big-data-analytics>.
- Gaitho, M.** 2015. How Applications of Big Data Drive Industries <https://www.simplilearn.com/big-data-applications-in-industries-article>.
- Guillermo, L.** 2014. Big data security challenges and solutions. <https://www.mwrinfosecurity.com/our-thinking/big-data-security-challenges-and-solutions/>
- Hurwitz, J., Nugent, A., Halper, F. and Kaufman, M.** 2013. Big data for dummies. John Wiley & Sons, Hoboken.
- Kirk, M.** 2016. Big data: A European survey on the opportunities and risks of data analytics. Vodafone Institute for Society and Communications.
- Kuketz, D.** 2017. Biggest business benefits from big data. Accessed from <http://archive.utoipai-nc.com/8-blog/381-7-biggest-business-benefits-from-big-data>
- Maryanne, G.** 2011. <https://www.simplilearn.com/big-data-applications-in-industries-article>
- McGuire, T., Manyika, J. and Chui, M.,** 2012. Why big data is the new competitive advantage. Accessed on 23rd May 2017 from <http://iveybusinessjournal.com/publication/why-big-data-is-the-new-competitive-advantage/>.
- McNulty-Holmes, E.** 2014. Understanding big data: The seven v's. BI and Analytic - Data Science. Accessed from <http://dataconomy.com/2014/05/seven-vs-big-data/>
- Oloruntoba, C. O. And Sodiya, A. S.** 2015. A noise perturbation framework towards enhancing privacy in big data computing. *Proceedings of 1st International Conference on Applied Information Technology*. pp 103 - 106.
- Press, G.** 2016. Top 10 hot big data technologies. Accessed from www.forbes.com/sites/gilpress/2016/03/14/top-10-hot-big-data-technologies
- Singh, D. and Reddy, C. K.** 2014. A survey on platforms for big data Analytics”, *Journal of Big Data*. 1:1, 8. <https://www.forbes.com/sites/gregsatell/2013/12/03/yes-big-data-can-solve-real-world-problems/#300e4dbc8896> <https://evolvedesign.ie/2015/09/10/what-is-the-advantage-of-big-data-analytics/> <https://evolvedesign.ie/2015/09/10/what-is-the-advantage-of-big-data-analytics/> <http://www.computerweekly.com/feature/How-to-tackle-big-data-from-a-security-point-of-view>