BIG DATA TECHNOLOGY AND LEVERAGING OPEN DATA FOR ELECTORAL PROCESSES IN NIGERIA

By

Dr. Omoniyi Victor Ajulor and Ejalonibu, Ganiyu Layi

Department of Public Administration,

Faculty of Management Science,

Lagos State University, Ojo

Email – ajulorvictor@yahoo.com & ejaloniganiyu@yahoo.com

Telephone Nos. - +2348023600421, +2348032177394

ABSTRACT

What comes to our minds in Nigerians during electioneering processes are the humongous rallies, colourful campaigning, myriad party symbols and distribution of items and monies now called " Stomach Infrastructure". The election results are marred with electoral frauds such as rigging, snatching of ballot boxes and consequent political violence. The political communication with citizens is changing and experiencing serious revolution globally through the use of Big Data and electronic technologies. Nigeria is gradually introducing electronic technologies in the conduct of election but much has not been achieved in this direction. This paper therefore examines Big Data Technologies and how Open Data can be leveraged for electoral processes in Nigeria. The study makes use of secondary data in the form of content analysis of books, journals and internet materials. The paper discovers that the combination of Big Data and computational politics allow for massive, latent data collection and sophisticated modeling. There is an increase in the capacity of those with resources and access to use these tools to carry out highly effective and unaccountable campaign of persuasion and social engineering in political, civic and commercial spheres. It is therefore recommended that, the initiative of Big Data Technology and the leveraging Open Data for electoral processes in Nigeria should be encouraged in order to reduce election rigging and malpractices. However, not only Independent National Electoral Commission (INEC) should be involved but also by other various bodies, which have well-defined roles for a greater coherence in achieving quality electoral objectives.

Keywords: Big, Data, Electioneering, Media, Nigeria, Social, Technology

Introduction

The world of electronic technology (especially Internet service) works because of the existence of basic standards of data exchange. In many areas of commerce and government there exist Electronic Transaction Standards (ETS) that facilitate Electronic Data Interchange (EDI). An EDI provides a defined format for the exchange of data for every specific transaction in question (Alvarez and Hall, 2005). Big data is increasingly becoming a factor in governance, production, market competitiveness and, therefore facilitating growth. That is, the cutting-edge analysis technologies such as Big Data analytics are making inroads into all areas of life and changing our day-to-day existence including political life. Sensor technology, biometric identification and the general trend towards a convergence of information and communication technologies have, for a while now, been driving the big data movement in business and recently showing up in politics.

There are a lot of challenges facing nations associated with manual/paper based ballot. The traditional paper-based voting system consists of a voter manually marking the paper ballot and the ballot being counted by hand by election officials. The method of voting used in five out of six past elections in Nigeria in 1979,1983, 1999, 2003, 2007 and 2011 was the Open Ballot System (OBS) in which the prospective voter goes through a process of accreditation, receives a ballot paper from the appropriate poll official and thereafter makes the confidential thumb impression in favour of the political party or candidate of choice in a secret voting compartment before dropping the ballot in the box positioned in the open, in the full glare of officials, security and party agents. The Modified Open Ballot System was adopted in the 1993 elections, in which voters filed behind the party symbol or photograph of the candidate of choice. Voters were physically counted at the close of polls and the results declared to officials, security and party agents (Mohammed and Bashir, 2009).

In this age, governments and public authorities across the world are launching Open Data initiatives. Research indicates that by October 2011, twenty eight nations around the world had established Open Data portals (www.slideshare.net, 2011). One of the benefits of Big Data in business fields (which can also be applied in political fields) is that it enables organizations to analyze business problems in the context of a more complete view of processes and their interactions, analyzing a greater number of scenarios quicker and more cheaply. Thus, Big Data holds great potential to change the whole electoral system value chain, from voters' registration, to electioneering, to accreditation of voters, to actual voting processes for improved electoral outcomes and acceptability of result, to a legitimate and good government.

In 2002, Professor Robert Done published an article entitled “Internet Voting: Bringing Elections to the Desktop”. In that article, he addressed some of these challenges, emphasized the need to discard paper-based election and explained while this should also be on the reform agenda. He believed that in this age of advanced communication technology, human elements in election management should be reduced to a barest minimum (Done, 2002).

Nigeria’s experience with paper-based balloting has produced challenges such as the snatching of ballot boxes and alteration of election results and these challenges remained. Conducting a credible election in Nigeria is almost a mirage. It is therefore necessary to examine big data technology and how it can be used for leveraging open data for electoral processes in Nigeria.

Conceptual and Theoretical Framework

Big Data as a Concept

The term “Big data” has a variety of definitions and has been used in different ways. According to Einav and Levin (2013) most definitions reflect the growing technological ability to capture, aggregate, and process an ever-greater volume, velocity, and variety of data. In the same vein, IDC quoted in Piai and Claps (2013) sees Big Data technologies as a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data produced every day. This enables high velocity capturing, discovery and analysis. In other words, data is now available faster, has greater coverage and scope, and includes new types of observations and measurements that previously were not available.

Big data refers to the vast amount of data that is now generated and captured in a variety of formats and from a number of disparate sources. A widely accepted definition of big data is the one provided by Gartner (2013) who sees it as “…high-volume, high velocity and/or high variety information assets that demand cost-effective innovative forms of information processing for enhanced insight, decision making and process optimization”. Big data exists in both structured and unstructured forms, including data generated by machines such as sensors, machine logs, mobile devices, GPS signals, transactional records and automated streams of information exchanged under initiative such as Standard Business Reporting (www.sbr.gov.au, 2014).

According to McKinsey (2012) Big Data refers to datasets whose size are beyond the ability of typical database software tools to capture, store, manage and analyze. There is no explicit definition of how big a dataset should be in order to be considered Big Data. New technology has

to be in place to manage this Big Data phenomenon. According to O'Reilly, "Big data is data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or does not fit the structures of existing database architectures. To gain value from these data, there must be an alternative way to process it."

Following from the above definitions, it has been discovered that the common themes for big data are Volume–Variety–Velocity–Complexity–Variability. It is important to include all of these characteristics because otherwise the term "big data" may continue to be applied to a variety of circumstances diluting the real meaning. The use of the term "big data" should carry all of these characteristics. Nevertheless, definition still remains an issue and a big enough issue that establishing a clear definition is one of the priorities of the Tech America Big Data Commission (NASCIO, 2012).

As Manovich (2011) observes, it has been used in the sciences to refer to data sets large enough to require supercomputers, but what once required such machines can now be analyzed on desktop computers with standard software. There is little doubt that the quantities of data now available are often quite large, but that is not the defining characteristic of this new data ecosystem. In fact, some of the data encompassed by Big Data (e.g., all Twitter messages about a particular topic) are not nearly as large as earlier data sets that were not considered Big Data (e.g., census data). Big Data is less about data that is big than it is about a capacity to search, aggregate, and cross-reference large data sets.

Electoral Process

Election is described as the process of choosing people for particular jobs by voting. There are two major types of election (direct election and indirect election). The former emphasizes direct participation of voters in election. Each voter goes to the poll and records a vote in favour of one candidate or another. The candidate with the maximum number of votes is declared elected. This method is said to be the most popular and is used in all democratic countries (Ojo, 2007).

Electoral process on the other hand can be described as "the wider set of activities that creates and maintains the broad institutional framework in which voting and electoral competition take place". For example, in Nigeria, elections take place every four years to vote into power a President and Members of National Parliament or governor and members of State House of Assembly in case of the State. The electoral procedure involves many processes. These processes

include voters' registration, voter register exhibition, voting, vote counting, collation and publication of results (Mozaffar and Schedler, 2002:7).

Seemingly, electoral process relates to the entire cycle ranging from the provision of voter education to the dissolution of the National Assembly. INEC quoted in Elekwa (2008), identifies the different phases of the electoral process such as: delimitation of electoral boundaries; registration of voters; notice of elections; nomination of candidates; election campaigns, elections, announcement of results and completion of tribunal sittings; participation of other organizations; resolution of electoral conflicts resulting from the participation of other organizations, people, groups, etc.

Whatever phases of electoral process we may want to follow, the point here is that a good electoral system must guarantee the right of the electorate to freely exercise their franchise while simultaneously creating a level playing field for all contenders in the electoral competition. Electoral process operates at three basic levels: rule making (defining the basic rules governing electoral competition), rule application (implementing electoral rules), and rule adjudication (lodging and disposing election petitions). In practical terms, electoral process does not operate in a vacuum, but in a political context characterized by competition among various political interests and constituencies for a head start. Partisanship defines electoral system across democracies and, as shown by Agersinger (2004) in his study of American election laws, matured democracies are not immune to this syndrome thus confirming Sartori's (1976) thesis that electoral system is the most manipulated instrument of politics.

The framework for electoral process in Nigeria is anchored on two instruments: the 1999 Constitution and the Electoral Act both of which were recently amended by Nigeria's federal legislature. These two documents contain rules and regulations that drive electoral governance in Nigeria. These regulations are aimed at protecting the integrity of the electoral process. Although these regulations are ambitious, they have failed to ensure credible elections. The failure of these documents is attributed to two major reasons: the inherent weaknesses/ inadequacies of some of the provisions of these documents; and weak enforcement institutions. These two factors were brazenly exploited by the politicians during the first post transition elections in 2003 and were taken to an absurd level during the 2007 polls (Animashaun, 2010).

The current model of parliamentary representation in Nigeria is the First-Past-the-Post (FPTP) system. In this system, according to Animashaun (2010: 19):

> The candidate with the simple plurality of the total votes emerges as the representative in a single-member constituency even if he does not command half of the total votes. This has been the only electoral system in the electoral history of Nigeria and has considerably strained inter-group relations in the country. It is against this background that stakeholders have been canvassing for the introduction of proportional representation (PR) model. Proportional representation is an electoral system, which allocates parliamentary seats to parties according to their share of the national votes.

Apart from being more inclusive than the First-Past-The-Post (FPTP) system, it ensures representation of the minorities whose votes carry no electoral weight under the majority principle of the FPTP regime. There is a direct relationship between votes and parliamentary seats under the proportional representation system which allows for a minimal number of wasted votes.

Although, the proportional representation is complex to administer, it is the best option in order to tackle the plethora of problem inherited in FPTP. Such problem can also be solved with the use of ICT. The issues that are raised in this section about leveraging Big Data for electoral process need to be put under search light through a best suited theoretical framework.

Theoretical Framework

This study focuses on the theoretical foundations of computational social choice. Computational Social choice theory as a scientific discipline with sound mathematical foundations came into existence in 1951 with the publication of the Ph.D. thesis of Kenneth J. Arrow who introduced the axiomatic method into the study of aggregation methods and whose seminal Impossibility Theorem shows that any such method that satisfies a list of seemingly basic fairness requirements must in fact amount to a dictatorial rule. Social choice theory was originally developed as an abstraction of problems that arise in political science and economics. Generally, social choice theory provides a useful theoretical framework for the precise mathematical study of the normative foundations of collective decision making, in a wide range of areas, involving not only human decision-makers but also autonomous software agents (Brandt, Conitzer & Endriss, 2013).

Since inception, much of the work in classical social choice theory has focused on results concerning the formal possibility and impossibility of aggregation methods that combine certain desirable properties, like Pareto-optimality, monotonicity, or non-manipulability, without resulting in an unacceptable concentration of power (Arrow, 1963). Also, some of the landmark results include Sen's characterization of preference domains allowing for consistent majority decisions (Sen,1966) and the Satterthwaite Theorem which establishes the impossibility of devising a reasonable, general voting rule that is immune to strategic manipulation (Satterthwaite, 1975).

It should however be noted that this theory has actual and potential application domains, going beyond political elections and collective decision making both in private and public sectors, where the theory can be put to good use. The first such example comes from the domain of Internet search engines where a Meta search engine that combines the search results of several engines is designed. This may be likened to or has a lot in common with preference aggregation. Aggregating preferences means asking each individual agent for a ranking over the set of alternatives and then amalgamating this information into a single such ranking that adequately represents the preferences of the group. For the Meta search engine, each individual search engine is asked for a ranking of its own, say, 3 top results and then have to aggregate this information to produce the Meta ranking (Brandt, Conitzer & Endriss, 2013).

In the 1990 and in the first few years of the 21st century, as the relevance of social choice to artificial intelligence, multi-agent systems, computational politics and electronic commerce became apparent, the frequency of contributions on problems related to social choice with a computational flavour suddenly intensified. Although, the field was still lacking a name, by 2005 contributions in what is now called "computational social choice" had become a regular feature at several of the major conferences in artificial intelligence. The first workshop specifically dedicated to computational social choice, and the first event to explicitly use this name, took place in 2006 (Endriss & Lang, 2006). Around the same time, Chevaleyre et al. (2007) attempted the first classification of research in the area by distinguishing the nature of the social choice problem addressed, and the type of formal or computational technique used.

The Use of Big Data and Leveraging Open Data for Electoral System in Nigeria

Recently, digital technologies have given rise to a new combination of big data and computational practices which allow for massive, latent data collection and sophisticated

computational modeling, increasing the capacity of those with resources and access to use these tools to carry out highly effective, opaque and unaccountable campaigns of persuasion and social engineering in political, civic and commercial spheres. For example, the political communication with citizens is changing and experiencing serious revolution through the use of big data and computational techniques (Tufekci, 2014).

There are major areas through which Big Data technologies can be leveraged on for efficient effective Electoral Process in Nigeria. The new found technology can be of good use in (1) voters registration (e-registration) (2) electioneering, political campaigns and presentation of party programmes and manifestoes (3) technology-based election or electronic voting and (4) franchise for Nigerians in Diaspora (5) Election Reporting.

Election Reporting

Diverse groups argue about the potential benefits and costs of analyzing genetic sequences, social media interactions, health records, phone logs, government records, and other digital traces left by people. According to Tufekci (2014) Big data has the potential to turn political communication into an increasingly personalized, private transaction and thus fundamentally capable to reshape the public sphere, first and foremost by making it less and less public as these approaches can be used to both profile and interact individually with voters outside the public sphere such as Facebook aiming at a particular voter, seen only by her.

In contrast to broadcast technologies, the Internet offers expansive possibilities for horizontal communication among citizens, while drastically lowering the costs of organizing and access to information (Shirky, 2008). Indeed, the Internet has been a critical tool for many social movements (Tufekci and Freelon, 2013). However, Internet's propensity for citizen empowerment is neither unidirectional, nor straightforward. The same digital technologies have also given rise to a data-analytic environment that favors the powerful, data-rich incumbents, and the technologically adept, especially in the context of political campaigns. These counter-trends arises specifically from an increased exploitation on big data, that is, very large datasets of information gleaned from online footprints and other sources, along with analytic and computational tools. Big data is often hailed for its ability to add to our knowledge in novel ways and to enrich our understanding (Lazer et al., 2009; Lohr, 2012).

One other important aspect of electoral process where Big Data (Open Data) can be leveraged is election reporting. For example, India recently concluded their general elections. The Indian General Elections have one perspective which often does not figure in the most buoyant thoughts held of it. It has been well observed that election in India is a classic Big Data problem and the 2014 general elections were the biggest of them all. To be sure, 300 parties, 8000 candidates, 800 Million voters, 1 Million booths served by 20 Million officials. The heady mix is further embellished with variety of structured & unstructured information candidate histories, crime records, declared assets and audacious election manifestos. Mixed with the above is the frenetic activity on the day of results. Live streaming of results of about1000 votes to be counted per second, from all corners of the country spanning an area of 1 million square miles (www.gramener.com).

Before thinking about the imminent elections was the task of assimilating historical data. Data from 1950s was to be gathered from multiple PDF files published on the Election Commission of India (ECI) website and followed with the task of cleaning, correction & collation. Synthesizing this data with various other credible data sources completed the painstaking exercise of building an integrated, structured elections data source. This master data source served as the backdrop for 2014 election with over 60 years of election data. Gramener's proprietary analytics and visualization technology was used to complete this task and this history was hosted through a web page (https://gramener.com/election/parliament).

For the first time, the electorate in India was exposed to data-based reporting with hard facts and numbers, which were easy to consume as data stories. Then when the polling phase began the same analytics and visualization techniques were used to report the 5 weeks of polling, by providing live facts on voter turnouts, exit poll surveys etc. Accuracy, speed and ease of consumption were put to test in front of the entire nation which had reached a crescendo in anticipation. Gramener's analytics and visualization engine, Gramex® was programmed to arrive at an election results dashboard which was to meet the heavy expectations (www.gramener.com).

Still in the category of data uses, scholars have observed that big data (i.e. remotely collected data about large groups such as populations or users of a particular technology) can also be seen as a tool for configuring communities and actions (sometimes serendipitously) through data science and aggregation. An example is the data science that accompanied US President

Obama's 2012 campaign, which categorised and then mobilised voters based on an unprecedentedly detailed level of data on the individual level (technologyreview.com….voters/).

Voters Registration

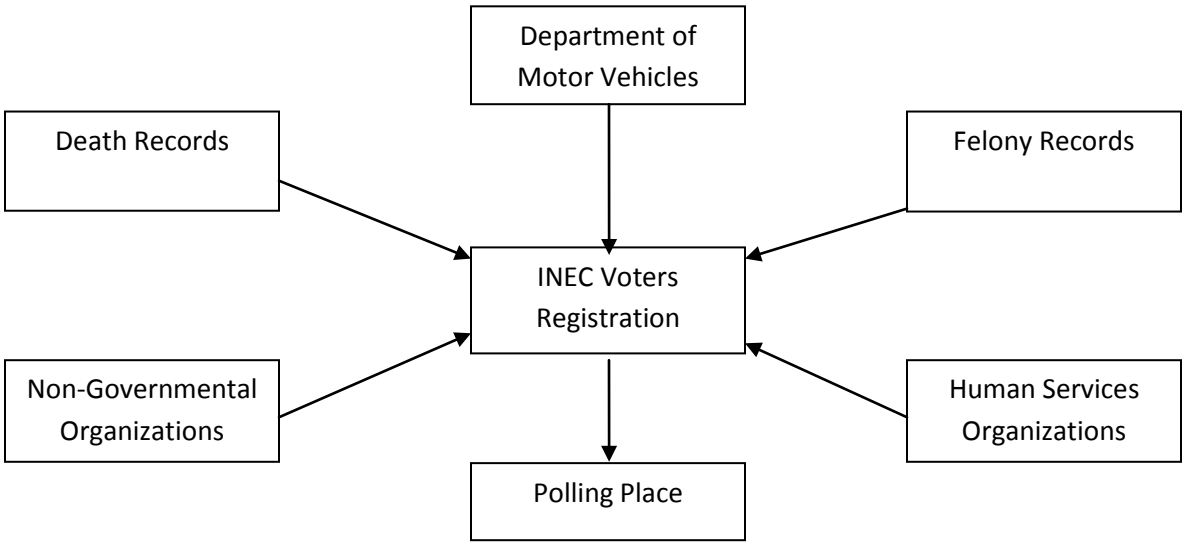
Registration of voters is the process of enabling an eligible voter to have his/her name entered into a document (Voter Register) with the aim of offering the person the opportunity to exercise his/her franchise on the appointed day of voting. Nigerian laws peg an eligible voter as one who is 18 years of age or above, a national and resident in the country. The Voter Register is considered provisional until such time when the Permanent Voter card is available. In February 2011, INEC carried out a computerized nationwide voter registration exercise that is centralized at the federal level. However, for a proper monitoring and future update keeping of such data, INEC will need to link their voter registration system with the states and local government databases, including those governing an individual's felony status (if applicable) and death records. Although not explicitly required, the database also needs to be able to coordinate with the state's department of motor vehicles and the federal Social Security Administration's database; both of these linkages are needed so that information from new registrants can be compared to either of these external databases for verification.

To be sure, there are many other entities (governmental and non-governmental) with which INEC voter registration system needs to be able to interface in order to keep the voter registration system up-to-date (see figure 1 below). With a common protocol, the transmission of data can occur in a couple of ways. First, it can run through a data center, where individuals convert the data from one electronic format to another, which often requires reformatting the data or re-entering parts of the data. Second, the data may have to be completely hand-entered by the election officials in charge of voter registration. This process of reformatting or re-entry introduces opportunities for data entry errors, errors that can result in voters not being listed correctly on the voter rolls at their polling place. When this occurs, a voter often has to cast a provisional ballot, which slows polling place operations on Election Day and results in the voter's ballot not being counted (Alvarez and Hall, 2005).

The growth of social media use in society is generating large quantities of new digital information about individuals, organizations and institutions that is now commonly labeled Big Social Data. Social media analytics is a term we use here to refer to the collection, storage,

analysis, and reporting of these new data (Vatrapu, 2013). These social data sets carry valuable information and if analysed utilizing proper methods, techniques, and tools of computational social science in particular and data science in general. They can provide meaningful facts and actionable insights that go beyond traditional social science research methods. For example, recent studies have shown that social data on Face-book can be analysed for investigating political discourse on online public spheres for the United States Election (Robertson, Vatrapu, and Medina, 2010), and social data from twitter has been used for predicting Hollywood movies' box-office revenues (Asur and B. Huberman, 2010). Conte et al (2012) also point that Computational Social Science is a model based science that analyses electronic trace data, builds predictive models and intends to provide instruments for enabling social science to inform decision makers for societal and organizational challenges.

Figure 1: The Voter Registration Network



Source: Adapted from Alvarez and Hall (2005).

Additionally, there are many legal and social factors that affect the need for data uniformity with voter registration systems across the nation and various sectors in Nigeria. For instance, mobility of voters has a greater impact on election administration, and uniform protocols for voter registration would improve the elections process. Very often in Nigeria, there are rural-urban migrations and most of these moves are intra-state moves, where those involved may be unable to vote. All of these individuals potentially created a two-part voter registration issue: (1) the

need to register to vote in their new state, and (2) the need to un-register to vote in their previous state of residence. This mobility rate means that every general election year, millions of people could be voting in a new state. Whereas it is our believe that the concept behind the requirement for nationwide voter registration systems is, in part, intended to address the voter re-registration problems associated with short moves.

Individual data sets held by INEC and viewed in isolation are only of limited use when it comes to analysing big data. Only once several data points, some from different sources, are merged does it become possible to extract certain patterns. Very few organizations offer such a diverse range of products and services that they can accumulate a sufficiently broad variety of customer information under one roof (Dapp, 2014). If this information contains enough personal data, it can be used to create comprehensive and relatively detailed profiles of people. The majority of the various data being stored, archived and analyzed originates from digital advertising, information, transaction and other web channels. It is relatively easy to collect information about the respective internet provider, IP or e-mail address and the search engine used.

Dapp (2014) further observes that the growing data collections held by the stakeholders makes it relatively easy to link the real and the virtual world, making even physical addresses and telephone numbers of individuals fairly simple to identify. On top of that it is possible to make further deductions. What would happen if personal data was also linked to other information from social networks, data on election matters, location data (GPS tracking), and data on age, status and personality? The resulting data correlations would in turn allow new, relevant conclusions to be made about personal preferences and characteristics.

Electioneering

Big data can play a substantial role in guiding the election strategy. For example, it was used in Obama 2012 campaign in the US. The campaign had a large data analytics team, who used data from social media (including Face-book and Twitter), alongside data from their own party database, which included information on approximately 180 million voters. By looking for correlations in past voter characteristics and behaviour, they were able to build up profiles of the kinds of people who might vote for them, and to target resources more efficiently. The TV adverts were broadcast when they were known to have the most impact with the targeted swing voters, rather than in premium generic prime-time slots. Analytics was also used to determine which households to target door-to-door. These approaches have not yet been taken up to the

same extent in the UK, although they are likely to become more prominent in the 2015 general election. However, differences in data regulation and campaign spending may affect how widely social media data analysis is used in UK politics (Enos and Fowler, 2014).

Howard (2005) observes that social media platforms that are increasingly integral to the practice of computational politics have fully blossomed only recently. These new practices build upon the growing ability of campaigns to use technology to "manage" the electorate, a dynamic which has so far been examined in case studies of Barack Obama's campaigns, as well as an ethnographic account of a congressional effort (Kreiss, 2012; Bimber, 2014; Nielsen, 2012). The Obama campaign is the most recent example, best-studied and most relevant one. Further, 2012 also saw computational methods besides polling spread to outside of campaigns, such as that of Nate Silver's simulation models; that, however, is beyond the conceptual scope of this paper which focuses on campaigns and political actors. Rather, this is a conceptual and theoretical-building paper that grapples with the consequences newly emergent computational politics.

Electoral Prediction

Although unstated, it is assumed that any method to predict electoral results from Twitter data is an algorithm; otherwise, it would be impractical and pointless. Therefore, such methods process some collection of tweets to make predictions; they are parameterized to adapt to different scenarios; and, finally, predictions can be more or less detailed (for instance, just providing the winner or vote rates for the different candidates) and they should be eventually evaluated against the actual results. Thus, there are a number of characteristics and sub-characteristics defining any method to predict electoral results from Twitter as mentioned by Gayo-Avello (2012), namely:

1. Period and method of collection: i.e., the dates when tweets were collected, and the parameterization used to collect them.

2. Data cleansing measures:

- Purity: i.e. to guarantee that only tweets from prospective voters are used to make the prediction.
- Debiasing: i.e. to guarantee that any demographic bias in the Twitter user base is removed.

- Denoising: i.e. to remove tweets not dealing with voters' opinions (e.g. spam or disinformation) or even users not corresponding to actual prospective voters (e.g. spammers, robots, or propagandists).

3. Prediction method and its nature:

- The method to infer voting intentions from tweets.
- The nature of the inference: i.e., whether the method predicts individual votes or aggregated vote rates.
- The nature of the prediction: i.e., whether the method predicts just a winner or vote rates for each candidate.
- Granularity: i.e., the level at which the prediction is made (e.g. district, state, or national).

4. Performance evaluation: i.e., the way in which the prediction is compared with the actual outcome of the election.

Technology-Based Election

Election is the heartbeat of democratic setting. This is because it gives the citizens opportunity to choose their representative through casting of vote. In Nigeria, voting in elections are usually done manually in form of paper-based elections. In the paper-based elections an eligible voter goes to the polling station where his name is registered with his voter identification card and casts vote through a ballot paper issued by an electoral officer. This process most of the times in Nigeria has resulted to electoral malpractices because of the human factors.

Technology-based election most especially the use of Electronic Voting and counting technologies are being increasingly used around the world. India, the world's largest democracy, now uses electronic voting machines exclusively for national and provincial elections. Brazil, Belgium and the Philippines also use electronic voting or counting technologies for all of their national elections. Countries such as Estonia, Indonesia, Kazakhstan, Nepal, Norway, Pakistan, Russia and the United States are at various stages of piloting or partially using electronic voting and counting technologies, including the use of Internet voting. In elections where the use of electronic voting is involved one or both of these processes are automated using an electronic device. In electronic voting an electronic device records the voting preference of the voter. This

voting device may be located at the polling station or a remote location; for example, a personal computer is used to cast a ballot over the Internet or a mobile phone.

The adoption of E-Voting system in Nigeria is considered long overdue because of the enormity of abundant human and material resources endowed the country (Olateru_Olagbegi, 2007). The shortcomings notwithstanding, the system has the potentials of solving a lot of the electoral problems. E-voting, which must be adopted with caution, could be disastrous if rushed on the electorates without regards to the associated risks.

It is also believed that *Social Media* can play a great role in electoral process in Nigeria such as conducting opinion polls among the general populace be it in rural or urban centers. For instance, sentiment analysis of Twitter has been used to reveal insights in real-time of users' political opinions. During the 2010 United Kingdom General Election, this technique was used to create a visual display of Twitter users' reactions to televised political debates (Anstead and O'Loughlin, 2012). Also, during the 2012 Presidential Election in the United States of America, the political news site "*Politico*" used sentiment analysis to examine large volumes of public and private data on Facebook as a complement to traditional methods of polling. However, there have been doubts expressed as to how useful analysis of the often spontaneous, emotional content of social media is for predicting the potentially more calculated decisions of voters during elections.

A good example of computational politics can also be seen in the recently concluded Indian general elections. Elections in India are classic Big Data problem and the 2014 general elections were the biggest of them all. The Indian General Elections of 2014 involved 300 parties, 8000 candidates, 800 Million voters, 1 Million booths served/secured by 20 Million officials. The heady mix is further embellished with variety of structured & unstructured information such as candidate histories, crime records, declared assets and audacious election manifestos. Mixed with the above is the frenetic activity on the day of results. Live streaming of results: 21000 votes to be counted per second, from all corner of the country spanning an area of 1 million square miles (Ramachandran, 2014). While technology may be able to process this humongous data, how all these information can be consumed and understood by millions of people in India is a relevant question at this junction? Gramener an engineering company where Ramachandran is the CEO took the task of analyzing such large volumes of data into meaningful information for the general public to consume easily.

Franchise for Nigerians in Diaspora: Nigerian nationals abroad have never been accommodated in the transition process particularly as voters. Section 77(2) of the 1999 constitution allows only eligible voters resident in Nigeria to be registered as voters. This has denied Nigerian citizens in the Diaspora the opportunity to participate in leadership recruitment in their country. Even with the verdict of an Abuja High Court delivered on January 27 2009 in a suit filed by some Nigerians resident abroad (Aluko, 2009), this is yet to be in operation based on logistics. However, one of the transformative potentials of big data is that open data initiatives make data available to the public via integrated web portals and automated interfaces (El-Darwiche, Koch, Meer, Shehadi, & Tohme, 2014).

Internet voting has become a hot topic in recent years and most governments in Europe and elsewhere are planning to experiment with it, and to implement it. Many technology development projects have been undertaken in recent years, and the technological standards are being established. At the same time, a lot of legal and philosophical issues are at stake, as the system, form and technologies for voting do have normative implications. This makes the politico-technical arena in which the development and implementation of e-democracy systems in general and e-voting systems in particular so difficult and complex.

The Challenges of Technology Based Election Processes

Technology-based elections are not, however, without their own challenges. The potential for Big Data is still generally untapped and in practice only 3% of potentially useful data is tagged and even less is analysed. This is not "just" a matter of semantics and data interoperability, it is more holistically a matter of understanding what set of methodologies, skills, regulatory, and organizational changes are necessary to leverage the benefits of Big Data (Piai and Claps, 2013). Some of the challenges facing the use of Big Data to leverage elections include high costs of procuring the needed technology as well as the limitations imposed by the high level of illiteracy. Other challenges are discussed below.

Claims to Objectivity and Accuracy are Misleading: Sociology has been obsessed by the goal of becoming a quantitative science. As Latour (2010) puts it ‘numbers, numbers, numbers’. Whereas Sociology in Latour’s view, has never reached this goal, because of where it draws the line between what is and is not quantifiable knowledge in the social domain. In reality, working with Big Data is still subjective, and what it quantifies does not necessarily have a closer claim

on objective truth – particularly when considering messages from social media sites. But there remains a mistaken belief that qualitative researchers are in the business of interpreting stories and quantitative researchers are in the business of producing facts. In this way, Big Data risks rein-scribing established divisions in the long running debates about scientific method and the legitimacy of social science and humanistic inquiry.

Bigger Data are Not Always Better Data: Social scientists have long argued that what make their work rigorous is rooted in their systematic approach to data collection and analysis (McClosky, 1985). While many scholars are conscientious about discussing the limitations of Twitter data in their publications, the public discourse around such research tends to focus on the raw number of tweets available. Twitter does not represent 'all people', and it is an error to assume 'people' and 'Twitter users' are synonymous: they are a very particular sub-set. Furthermore, the notion of an 'active' account is problematic. While some users post content frequently through Twitter, others participate as 'viewers' (Crawford, 2009: 532).

Just Because it is Accessible Doesn't Make it Ethical: In 2006, a Harvard-based research group started gathering the profiles of 1,700 college based Facebook users to study how their interests and friendships changed over time (Lewis et al. 2008). This supposedly anonymous data was released to the world, allowing other researchers to explore and analyze it. What other researchers quickly discovered was that it was possible to identify parts of the dataset thereby compromising the privacy of students, none of whom were aware their data was being collected (Zimmer, 2008).

Limited Access to Big Data Creates New Digital Divides: In an essay on Big Data, Scott Golder (2010) quotes sociologist George Homans (1974): 'The methods of social science are dear in time and money and getting dearer every day.' Historically speaking, collecting data has been hard, time consuming, and resource intensive. Much of the enthusiasm surrounding Big Data stems from the perception that it offers easy access to massive amounts of data.

Concluding Remarks

There are ever growing pools of data everywhere around us. While surfing internet, doing day to day transactions with IT tools, talking on mobile phones, travelling on plane, walking and interacting with digital gadgets, sending a tweet or a posting a comment on Face-book, all these actions are accumulating data. These very large sets of data are commonly termed as Big Data. Big data sets demand exploration and evaluation of new analytical methods, techniques, and tools as existing data analysis techniques are increasingly becoming inadequate (Mukkamala, Hussain & Vatrapu, 2013) and require multi-disciplinary skills. Leveraging the Big Data opportunity will also therefore require an end-to-end strategy where IT is the technical enabler but where new process and organization aspects are led by key executives that will also set the overall business objectives.

Big Data is good but it is not a straightforward task. It starts with government willingness to open up data. Once there is a broad buy-in to opening up public data, governance structures should be established to oversee all Open Data initiatives. Also, Governments will need to decide on the type of data that is prioritized for release and how best to offer this data to developers.

The final step is to enable governments to allow Open Data to flourish by to driving its uptake both by citizens and by developers. It is believed that Open Data will be the essential characteristic of future public policy (Maude, 2012). Governments need to develop institutions with an explicit mandate to frame and encourage the development of Open Data. Thus, the Open Data initiative in Nigeria should be advocated and encouraged not only by INEC but also by other various bodies, which have well-defined roles for a greater coherence in achieving quality electoral objectives.

It is important to note that it is not the technology that is used that matters most, but the way in which the technology is employed ultimately determines the success of the election technology project (Hall, 2010). Thus, before implementation, there should be proper consideration of all factors influencing the decision whether to adopt Big Data technologies and also all stakeholders should be given the opportunity to understand and express their opinions during the process. This guide provides a solid basis for the decision making process involved in whether or not to use these technologies (Goldsmith, 2011).

REFERENCE

Agersinger, P. (2004). Electoral Reform and Partisan Jugglery. *Political Science Review. 119*(3): 5

Altman, A. & Tennenholtz, M. (2008). Axiomatic Foundations for Ranking Systems. *Journal of Artificial Intelligence Research, 3*1:473–495.

Aluko, M. (2009). Star Ruling: Federal High Court in Nigeria Orders INEC to Allow Nigerians abroad to vote. Accessed from: *www.nigeriavilagesquare.com/tar-r...* on 26/12/2014.

Alvarez, R.M. and Hall, T.E. (2005). *The Next Big Election Challenge: Developing Electronic Data Transaction Standards for Election Administration.* California: IBM Centre for the Business of Government.

Animashaun, K. (2010). Regime Character, Electoral Crisis and Prospects of Electoral Reform in Nigeria. *Journal of Nigeria Studies*, 1(1): 1-33

Anstead, N. and O'Loughlin, B. (2012). *Semantic Polling*. London: LSE

Arrow, K. J. (1963). *Social Choice and Individual Values*. New Haven: Cowles Foundation.

Asur, S. and Huberman, B. (2010). Predicting the Future with Social Media, in *Web Intelligence and Intelligent Agent Technology (WI-IAT) 2010 IEEE/WIC/ACM International Conference*, 1(1, 2): 492–499.

Bartholdi, J. III & Orlin, J. B. (1991). Single Transferable Vote Resists Strategic Voting. *Social Choice and Welfare*, 8(4):341–354.

Bartholdi, J. III, Tovey, C.A. & Trick. M.A. (1989). The Computational Difficulty of Manipulating an Election. *Social Choice and Welfare*, 6(3):227–241.

Bimber, B. (2014). Digital Media in the Obama Campaigns of 2008 and 2012: Adaptation to the Personalized Political Communication Environment. *Journal of Information Technology & Politics*, 11(2), 130–150.

Brandt, F., Conitzer V. & Endriss, U. (2013). Computational Social Choice. Available @ www.BrandtEtAlMAS2013.com. Accessed 20/1/2015

Chevaleyre, Y. Endriss, U. Lang, J. & Maudet, N. (2007). A Short Introduction to Computational Social Choice. *In Proceedings of the 33rd Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM)*, Volume 4362 of Lecture Notes in Computer Science (LNCS): 51–69.

Conte, R.; Gilbert, N.; Bonelli, G.; Cioffi-Revilla, C.; Deffuant, G.; Kertesz, J.; Loreto, V.; Moat, S.; Nadal, J.P.; Sanchez, A.; Nowak, A. Flache, A. San Miguel, M. and Helbing, D. (2012). Manifesto of Computational Social Science. *The European Physical Journal Special Topics*, 214(1).

Crawford, K. (2009). Following you: Disciplines of Listening in Social Media. *Continuum: Journal of Media & Cultural Studies*, 23(4): 532-33.

Dapp, T. F. (2014). *Big data – the Untamed Force*. Frankfurt: Deutsche Bank Research

Einav, L. and Levin, J. (2013). The Data Revolution and Economic Analysis. Working Paper, No. 19035, National Bureau of Economic Research,*http://www.nber.org/papers/w19035*

El-Darwiche, B. Koch, V. Meer, D. Shehadi, R. T. & Tohme, W. (2014). Big Data Maturity: An Action Plan for Policymakers and Executives, in Bilbao-Osorio, B., Dutta, S. & Lanvin, B. *The Global Information Technology Report 2014: Rewards and Risks of Big Data.* Geneva: World Economic Forum.

Endriss U. & Lang, J. (2006). *Proceedings of the 1st International Workshop on Computational Social Choice (COMSOC-2006).* ILLC, University of Amsterdam.

Enos, R. D. & Fowler, A. (2014). The Effects of Large-Scale Campaigns on Voter Turnout: Evidence from 400 Million Voter Contacts. *Accessed from www.enosfowlera-ggregate..*

Gartner (2013).The Importance of 'Big Data: A Definition. *http://www.gartner.com/id=2057415* Accessed 7 November 2014

Gayo-Avello, D. (2012). *A Meta- Analysis of State-of-the- Art Electoral Prediction from Twitter Data.* Oviedo (Spain): University of Oviedo.

Goldsmith, B. (2011). *Electronic Voting & Counting Technologies: A Guide to Conducting Feasibility Studies.* Washington, D.C.: International Foundation for Electoral Systems (IFES)

Hall, T. (2010). Voting Technology and Election Experience: The 2009 Gubernatorial Races in New Jersey and Virginia. *E-VOTE2010 Conference*. Bregenz: Austria.

Homans, G.C. (1974). *Social Behavior: Its Elementary Forms*. Cambridge, MA: Harvard University Press.

Howard, P. N. (2005). *New Media Campaigns and the Managed Citizen*. Cambridge: University Press.

http://www.technologyreview.com/featuredstory/509026/how-obamas-team-used-big-data-to-rally-voters/

Kreiss, D. (2012). *Taking Our Country Back: The Crafting of Networked Politics from Howard Dean to Barack Obama*. Oxford: Oxford University Press.

Latour, B. (2009). Tarde's Idea of Quantification, in Candea, M. (ed.) *The Social After Gabriel Tarde: Debates and Assessments.* London: Routledge: 145-162. Available at: *http://www.bruno-latour.fr/articles/article/116-TARDE-CANDEA* on 26/12/ 2014.

Lazer, D.; Pentland, A; Adamic, L; Aral, S; Barabási, A.L; Brewer, D; Christakis, N; Contractor, N; Fowler, J; Gutmann, M; Jebara, T; King, G; Macy, M; Roy D & Van Alstyne, M. (2009). Computational Social Science. *Science,* 323(5915): 721–723.

Lemos, R. (2003). E-voting Machine Fails Inspection. Available @ *http://news.com.com/2102*

Lewis, K., Kaufman, J., Gonzalez, M.,Wimmer, A., & Christakis, N. (2008) 'Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks,* 30: 330-342.

Lohr, (2012). Steve Lohr, 2012. The Age of Big Data. The New York Times, Available at *http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html*.

McKinsey J. et al. (2012). Big data: The next frontier for innovation, competition, and productivity. Available from: *http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation*Accessed 10/11/214.

Maude, F. (2012). *Open Data Innovation Community*. London: UK Cabinet Office.

McCloskey, D. N. (1985). From Methodology to Rhetoric, In McCloskey, D. N. *The Rhetoric of Economics au*. Madison: University of Wisconsin Press: 20-35.

Mohammed, S. and Bashir, A. (2010). Assuring the Integrity of the Electoral Process Through Automation. *Being a paper, presented at fourth African Regional conference on sustainable development, Delta state University, Abraka, Nigeria between June 9 and 10*

Mozaffar, S. and Schedler, A. (2002). The Comparative Study of Electoral Governance-Introduction." International Political Science Review. 23(1): 3.

Mukkamala, R.R.; Hussain, A. & Vatrapu, R. (2013). *Towards a Formal Model of Social Data*. Copenhagen: IT University Technical Report Series.

NASCIO (2012). *The Big Data Revolution–Impacts for State Government–Timing is Everything. L*exington: NASCIO

Nielsen, R. K. (2012). *Ground wars: Personalized Communication in Political Campaigns*. Princeton, New Jessy: Princeton University Press.

Ojo, E. (2007). Media and Elections. *Discussion Paper presented at the Post Election Conference on 2007 Elections and the Way Forward organized by IFES Nigeria*. Abuja. August 22-23.

Olateru_Olagbegi 0. E. (2007). *E-Voting: The Way Out*. Punch on the web. Feb. 6th.

Page, L. Brin, S. Motwani, R. & Winograd, T. (1999). The Page Rank Citation Ranking: Bringing Order to the Web. *Technical Report 1999–66*. Stanford: Stanford University.

Pennock, D. M. Horvitz, E. & Giles, C. L. (2000). Social Choice Theory and Recommender Systems: Analysis of the Axiomatic Foundations of Collaborative Filtering. *In Proceedings of the 17th National Conference on Artificial Intelligence* (AAAI Press): 729–734.

Ramachandran, J. (2014). Indian Elections 2014 – Big Data for Billion People! Available @ *www.IndiasporaForum*. Retrieved on 21/12/2014

Robertson, S. Vatrapu, R. K. and Medina, R. (2010). Online Video Friends Social Networking: Overlapping Online Public Spheres in the 2008 US Presidential Election. *Journal of Information Technology & Politics*, 7(2-3): 182–201

Sartori, G. (1976). *Parties and Party Systems: A Framework for Analysis*. Cambridge: Cambridge University Press.

Satterthwaite, M. A. (1975). Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions. *Journal of Economic Theory*, 10:187–217.

Sen. A. K. (1966). A Possibility Theorem on Majority Decisions. *Econometrica*, 34(2):491–499.

Shirky, C. (2008). *Here Comes Everybody: The Power of Organizing Without Organizations*. Penguin Books.

Tinholt, D. (2012). *The Open Data Economy: Unlocking Economic Value by Opening Government and Public Data*. UK: Capgemini Consulting

Tufekci, Z. & Freelon, D. (2013). Introduction to the Special Issue on New Media and Social Unrest. American Behavioral Scientist. *doi: 10.1177/0002764213479376*

Tufekci, Z. (2014). Engineering the Public: Big Data, Surveillance and Computational Politics. *Forthcoming*, 19(7): 1-39

Vatrapu, R. (2013). Understanding Social Business in Emerging Dimensions of Technology Management. *Springer*: 147–158.

Zimmer, M. (2008). More on the 'Anonymity' of the Facebook Dataset – It's Harvard College. Accessed from: *http://www.michaelzimmer.org/2008/01/03/more-on-the-anonymity-of-the-facebookdataset-its-harvard-college/* on 20/12/ 2014.