Towards a Data-driven Smart Governance in Nigeria

[1]Emeka Ogbuju*, [2]Isa Aminu & [3]Abraham Musa Peter

[1]Department of Computer Science, Federal University Lokoja, Kogi State, Nigeria
[2,3]Department of Political Science, Federal University Lokoja, Kogi State, Nigeria
{emeka.ogbuju, isa.aminu, abraham.peter}@fulokoja.edu.ng

Abstract

Nigeria has over the years survived majorly by the exploration and exporting of crude oil. Wealth generated from crude oil has been responsible for virtually everything in Nigeria, ranging from politics, debate on resource allocation, manipulation of primordial sentiments leading to myriad of challenges that have made the formation of national identity elusive. This study is therefore an attempt to enhance a re-focusing of Nigeria on data as the new oil that can aid development. The vast data generated through the social media must be mined and further extracted, processed and analyzed for policy determination using ICT tools and techniques. The study adopted the resource curse theory as its theoretical framework of analysis to explain the contradiction of poverty in the midst of the abundance of natural resources in Nigeria and the urgent need for refocusing on the use of data as a way of collecting feedback from people. The data for the study were collected from secondary sources such as published materials, and internet sources. The study revealed that there are possibilities of a data-driven decision making in the nation. Proper data analysis will definitely make governance SMART – creation of simple policies, measurable plans, attainable procedures, relevant programmes and timely delivery of projects.

Keywords: *Stream analytics, Big data, Sentiment analysis, Smart governance, Opinion mining*

1.0 Introduction

Nigeria has over the years survived majorly by the exploration and exporting of crude oil. Wealth generated from crude oil has been responsible for virtually everything in Nigeria, ranging from politics, the debate on resource allocation and manipulation of primordial sentiments, leading to a myriad of challenges that have made the formation of national identity elusive. Policy direction also had been motivated by the avalanche of crude oil. The overdependence of Nigeria on the income generated by this natural resource had not solved the numerous problems that have bedeviled the country. Among those challenges are insecurity, terrorism, inequality, tribalism, deprivation and poor policy making strategies among others. The critical question, therefore, is: how has Nigeria and her people fared under the crude oil regime and what can the country and her people do to reverse the trend?

Data is all around us in Nigeria. In fact, it has been reported that over 7.2 million Nigerians generate a large amount of data over Facebook social network daily in form of posts, tweets, videos, photos, etc. (Matuluko, 2016). We practically live in a pool of data in Nigeria today. However, we critically under use this pool for meaningful governance. Just like crude oil, data must be mined and further extracted, processed and analyzed for policy determination using ICT tools and techniques.

Apart from having data processed and getting information as done with conventional database environments, government needs insights for decision making. Insights cannot be possible without proper and precise analytics of both raw datasets and information. This can only be made possible through the application of big data standards across all government channels deploying e-government initiatives. It is common that Government data is locked up and operates in silos.
There is no seamless interface of data sharing between one governmental agency and the other. Every agency collects and manages its own data separately. However, most government datasets are text based and commonly available to the public through their various online platforms. This is in itself an advantage for insights generation as these texts could be mined,

extracted and analyzed for informed decisions. Government can be said to be data driven when vital decisions are made from sentiments expressed by these texts.

2.0 Theoretical Framework

This study adopts the Resource Curse Theory as its theoretical framework of analysis. Resource curse explains the misfortune that befalls countries that are richly endowed with natural resources yet have less economic growth, less democracy, and worse development indicators than countries with fewer natural resources.

The scenario that has prevailed in Nigeria is a justification of the study of Acosta and de Renzio (2008) that rent dependent countries have a high tendency to mismanage the economy. This explains the volume of recklessness and mismanagement of public funds through embezzlement, bribery and extra-budgetary spending in Nigeria. This has made the Ibeanu (2008), which asserted that government should be able to wipe out affliction using the available affluence futile. The affluence refers to the abundant resources of the country whereas the affliction refers to the misery, hunger and poor quality of life of the people

This theory explains the apparent contradiction in Nigeria occasioned by the increasing spate of poverty in the midst of abundance. Its utility, therefore, is anchored on the need for the Nigerian government to return to the 'ordinary' Nigerians to garner appropriate data that could be processed through Information and Communication Technologies (ICTs) as a means of embracing appropriate policy options that can reverse the resource curse analogy.

This theory becomes relevant in view of the 'Dutch' disease syndrome that had afflicted Nigeria as a result of heavy reliance on oil as her main source of revenue. The volatility of oil prices has always subjected the Nigerian economy to a cycle of boom and burst.

Judging from the foregoing and the apparent volatility of crude oil prices, it is obvious that Nigeria needs a *new oil*; an oil that cannot be exploited by usage, easily generated by all and freely available to all. In the modern world powered by ICTs, *data is the new oil!* Nigeria needs to discover data and the insights it can generate for effective policy-making and good governance.

3.0 Review of Previous Works

Actually, the use of ICTs in the administration and management of people and resources is not new in Nigeria. In fact, Nigeria had already launched a smart city initiative under the umbrella of the e-Government unit of the National Information Technology Development Agency (NITDA, 2010). Also, Nigeria had been participating in the United Nations E-government Survey since 2003. However, careful analyses of the results show an epileptic growth in the Nigerian e-government ranking. For example, as at 2016, the survey placed Nigeria 7 times backward almost returning to its ranking in 2003. This shows that governance in Nigeria had not started exploiting the power of ICTs fully as it concerns data analytics. The aim of this study is to draw the attention of Nigeria to its new oil and expose the potentials it holds in discovering meaningful insights for better policy formulation and good governance. There are certain agencies already charged with the mandate of collecting data in Nigeria. However, most of the data collected are structured data, which are mainly suited for relational database management system. Usually, these structured datasets are merely good for querying and reports generation. We still face the problem of non-collection of semi-structured and unstructured datasets which are mainly user-generated and real-time. We need a system that can collect citizens' feedbacks in form of tweets or texts and measure their feelings/opinions on any area of governance.

Sentiment analysis is an unsupervised leaning method where a set of variables or inputs are used to discover patterns, structure, relationship or similarities/dissimilarities in the input. It is sometimes called opinion mining or sentiment extraction. As an application of Natural Language Processing, it has been applied in many text based platforms especially in social networks (Edison and Aloysius, 2016; Manu, 2015; Michał and Andrzej, 2015). There are several approaches in performing sentiment analysis, however, three (3) of the approaches are common in the literatures (Fang and Zhan, 2015; Silge and Robinson,

2017). These approaches involve using any of the publicly available general purpose lexicons like:

1. The Bing lexicon which contains 2006 positive words and 4783 negative words (Lui and Hu, 2005)
2. The Finn lexicon which contains short texts typically used in social networks ranked from negative sentiment of -5 to positive sentiment of 5 (Finn, 2015)
3. The nrc lexicon which involves categorization of words into eight (8) basic emotions of anger, anticipation, disgust, fear, joy, sadness, surprise, and trust as well as positive and negative words (Mohammad and Turney, 2013).

The fundamental objective of sentiment analysis is to categorize a collection of texts or sentences into positive, negative or neural sentiments. This categorization can be done at sentence level or document level but each involves a proper analysis of each word in the texts into positive, negative, neutral or the emotion(s) it conveys.

Delenn, Anna, and Jessica (2016) analyzed sentiment of the US presidential elections using the emojis tweeted during the election period. Their work showed that emojis can be used to predict public sentiments of candidates in election period by collecting emotions like happy, sad, fear, laughter, and angry.

4.0 Research methodology

In this paper, we propose a centralized real-time data collection framework that is capable of aggregating all forms of textual data with a predefined set of keywords into a NoSQL document store. We tested this framework with data from prevailing and trending situations in Nigeria on security (keyword: Nnamdi Kanu) and politics (keyword: Buhari). The test data was a real-time data stream collected over a period of three (3) weeks from Twitter, a leading social networking site with the participation of over 7 million Nigerians daily. The storage tool used was MongoDB, a NoSQL document store that organizes tweets in JSON format (See Figure 1).

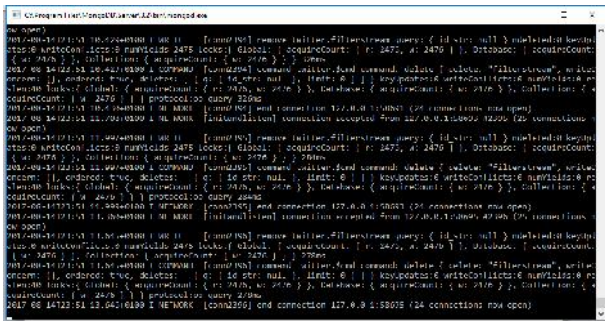

Figure 1. Data stream into MongoDB

The data were collected with the help of an API and processed using R programming. We performed sentiment analysis on the data to determine the opinion of Nigerians on the selected sector and the personalities involved.

5.0 Results and Discussion

A total of 0.621GB of tweets was collected on the keywords with over 24 connections.
The result showed a negative sentiment in certain areas of the sector and positive on the others. It reveals the true feelings of the citizenry on the personalities analyzed and show clearly the policy direction that the government should embark on to steer the sectors aright.

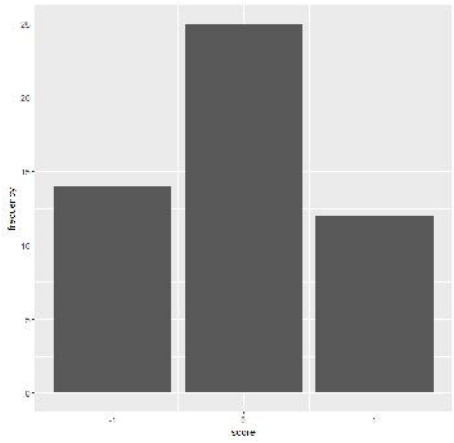

Figure 2. Sentiment Score on Security

The Figure 2 shows the sentiment score on security using Nnamdi Kanu, a secessionist leader in Nigeria campaigning for a state named Biafra. A high neutral sentiment was observed but with higher negative sentiment. A closer look on the result with the word cloud that produced the negative sentiment (See Figure 3) associated him with other secessionist groups like IPOB, Radio Biafra and MASSOB. The word cloud

also describes him as strong and fearless probably in calling for a referendum in Nigeria. The general opinion expressed is negative because the masses see his rearrest as neglect to fundamental human rights.

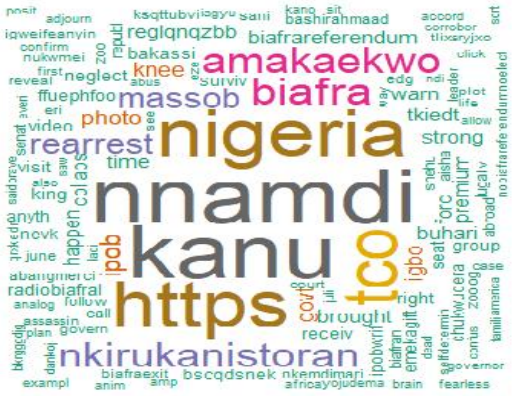

Figure 3. Word Cloud on Nnamdi Kanu

Figure 4 shows the sentiment score on Politics using Buhari, a noted politician and President of the country. A high neutral result is shown. Generally, in sentiment analysis using the Bing lexicon approach, higher neutral results is seen because most people usually tend to be on the objective side of the situation. This explains why the dataset gave more neutral results both on security and politics. Even though there are positive sentiments expressed on Buhari, there are higher negative sentiments.

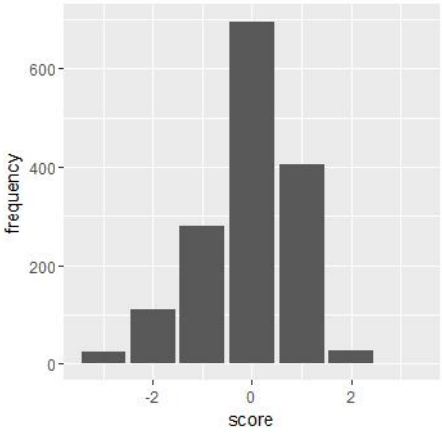

Figure 4. Sentiment Score on Politics

This can be further explained by the word cloud in Figure 5. There are many words that indicate the direction of the negative sentiments the masses hold on him is on his long trip to London on medical grounds away from his civil duties. The dilapidating state of the nation on education (strike), political stance on Biafra/Nnamdi Kanu, poor performances in the economic policies and other issues were the contributing actions that steer the negative public opinion on Buhari and his government.

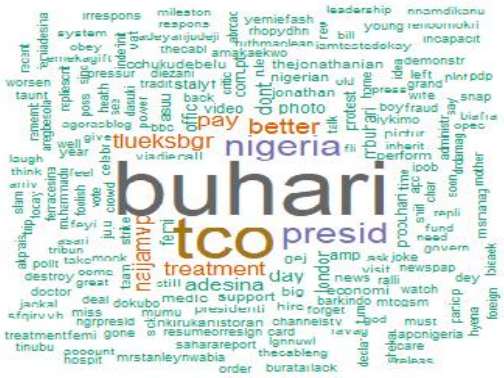

Figure 5. Word Cloud on Buhari

From the forgoing, evidence-based governance is possible. Policies can be formulated and right decisions can be made using the evidence provided by the analysis of these online dataset collected from the masses fully representing their opinions or sentiments on the sectors involved. For instance, a large number of the populace from an ethnic extraction is calling for a referendum with support from other ethnic groups but the government had not given ears to their agitations. Again, there is a high negative opinion about the prolonged absence of the President but the government had not addressed the matter.

It is with this backdrop that we advocate for a smart governance approach where decision making is championed by the collection, collation and analysis of public dataset.

To make governance smart, the application of the centralized real-time data collection framework we used in this work is inevitable. The framework allows a connection to a document store from different social network accounts as depicted in Figure 6.

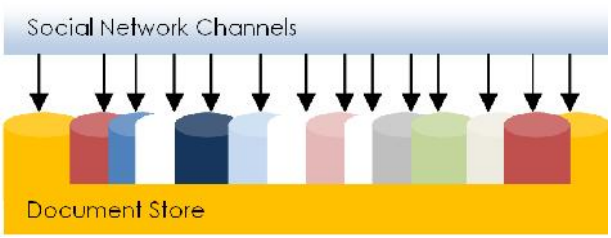

Figure 6. Data Collection Framework

This enables government to make use of the connectedness of social networks to harness public

opinions and make meaningful decisions from them. Other governments of the world already employing the smart governance concept include Australia (Pearce, 2015), USA (Government of the United States, 2014), Brazil (Government of Turkey, 2015), and the proliferation of smart cities in North America, Europe, Asia, South America and Dubai/South Africa (Lee et al., 2014).

6.0 Conclusion

This study reveals the possibilities of a data-driven decision making in the nation. Proper data analysis will definitely make governance SMART – creation of simple policies, measurable plans, attainable procedures, relevant programmes and timely delivery of projects. Also, in line with the smart ideology, NITDA, Nigeria's foremost agency charged with the responsibility of e-governance has as its mandate to ensure simple, moral, accountable, responsive and transparent (SMART) governance to Nigerians using the tools and techniques of Information Technology. Smart governance can only be possible with data, algorithm, and analytics combined to steer policy-making. The applications of the concepts in this paper will drive Nigeria and indeed any developing nation to a route towards true *smartification* of her governance.

7.0 Recommendations

We propose that if the Nigerian government could embrace data – driven smart governance, it will:

a. Help the country to take proper stock of her resources and maximize them appropriately for rapid socio – economic development.
b. Help the country to effectively collect data on people, industries and other organizations, for the purpose of adequate tax payment.
c. Help the government to carryout realistic allocation of resources to prioritize areas so as to forestall misappropriation of resources.
d. Introduce transparency in governance. This will help to gradually eliminate the challenges of mismanagement, misappropriation of funds, corruption and embezzlement.

ACKNOWLEDGEMENT

The authors acknowledge the travel grant provided to the lead author to participate in the Data for Policy Conference by the Sparrho Early Career Researcher Prize, August 2017. The lead author also acknowledges the full fee waiver granted to him by the conference team.

REFERENCES

Acosta, A. M. & de Renzio, P. (2008), Aid, rent and the politics of the budget process, *IDS Working Paper* 311, Sussex: Institute of Development Studies at the University of Sussex Brightan.

Edison, M & Aloysius, A (2016), Concepts and methods of sentiment analysis on big data, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 5, Issue 9, September 2016, Available at: https://www.ijirset.com/upload/2016/september/102_Concepts.pdf

Delenn, C. Anna, Z. & Jessica Z. (2016), Analyzing Twitter Sentiment of the 2016 Presidential Candidates, Available online: https://web.stanford.edu/~jesszhao/files/twitterSentiment.pdf

Fang, X & Zhan, J (2015), Sentiment analysis using product review data, Springer Journal of Big Data, Available at https://journalofbigdata.springeropen.com/articles/10.1186/s40537-015-0015-2

Finn A. N. (2015), https://finnaarupnielsen.wordpress.com/2011/03/16/afinn-a-new-word-list-for-sentiment-analysis

Government of the United States, Federal Emergency Management Agency (2015).'Welcome to FEMA's Flood Map Service Center', [online], available: https://msc.fema.gov/portal. [accessed: 1 November 2015].

Government of Turkey, Ministry of Family and Social Policy (2015).'Turkey's Integrated Social Assistance Services System: Recent Effort, Impressive Virtual Integration, Workshop on Integrated Data and Information Management for Social Protection: Bridging the Gaps between Theory and Practice Jakarta, 11-12 March, [online], available:

http://www.opml.co.uk/sites/default/files/Poster%20Turkey_FINAL.pdf.[accessed: November 2015].

Government of Turkey, Ministry of Family and Social Policy (2015).'Turkey's Integrated Social Assistance Services System: Recent Effort, Impressive Virtual Integration, Workshop on Integrated Data and Information Management for Social Protection: Bridging the Gaps between Theory and Practice Jakarta, 11-12 March, [online], available: http://www.opml.co.uk/sites/default/files/Poster%20Turkey_FINAL.pdf.[accessed: November 2015].

Ibeanu, O. (2008, February 20), Affluence and affliction: The Niger Delta as a critique of political science in Nigeria. An Inaugural Lecture of the University of Nigeria, Nsukka.

Lui, B & Hu, M (2005), https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#lexicon

Lee, Jung Hoon. et al. (2014). 'Towards an effective framework for building smart cities:
Lessons from Seoul and San Francisco', Technological Forecasting and Social Change, vol. 89 (November).

Matuluko, M. (2016). More than 16 million Nigerians use Facebook every month. Retrieved from https://techpoint.ng/2016/02/05/facebook-nigeria-q1-2016/

Mohammad, S & Turney, P (2013), http://saifmohammad.com/WebPages/meetsaif.htm

Michał, S & Andrzej, R. (2015), Sentiment analysis of twitter data within big data distributed environment for stock prediction, Proceedings of the Federated Conference on Computer Science and Information Systems pp. 1349–1354, Available at: https://annals-csis.org/proceedings/2015/pliks/230.pdf

NITDA (2010), The e-Government Framework for Nigeria.

Pearce, Rohan (2015). 'DTO works on Whole-of-Government Service Analytics', Computer World, 2 June 2015, [online], available: http://www.computerworld.com.au/article/576394/dto-works-whole-of-government-service-analytics/. [accessed: 1 November 2015].