**DATA-MINING BASED APPROACH TOWARDS SUSTAINABLE POLICIES AND OPERATIONS OF INTERNATIONAL OIL COMPANIES IN SUB-SAHARAN AFRICA**

A PhD research proposal

Submitted to the University of Bedfordshire

By: Zidiegha Seiyaboh

Supporting Academic: Dr. Paul Sant

Table of Contents

[1.0 Introduction 3](#_Toc519944455)

[2.0 Background 4](#_Toc519944456)

[3.0 Aims and Objectives of the research 5](#_Toc519944457)

[3.1 Research Motivation 5](#_Toc519944458)

[4.0 Research Questions 6](#_Toc519944459)

[5.0 Research approach and Methodology 6](#_Toc519944460)

[6.0 Research Plan 7](#_Toc519944461)

[References 8](#_Toc519944462)

# 1.0 Introduction

The environmental and social impacts of exploration and production operations of industries on communities in oil producing countries in sub-Saharan Africa has been a source of worry and in the media reports since the nineties. International Oil and gas Companies (IOCs) have caused some high-profile negative economic impacts which arose from certain oil and gas activities. Some examples are the Exxon Valdez oil spill that happened in Alaska and the campaign against Shell in Nigeria which led to the death of Ken Saro Wiwa. He was an environmental activist who campaigned against the Nigerian State and Shell Oil Company for devastating the Ogoni land. His cruel execution in 1995 was a shock to the entire world which, aggravated a global condemnation of the then military head of state, General Sani Abacha (Roy, 2016). These incidents and many others led to a widespread criticism of the oil and gas companies (Tomlinson, 2017). It has been observed that the projects carried out by these oil and gas companies generate huge wealth, yet the host communities are not beneficial, rather their living conditions become worse than the advent of the oil and gas companies (Christian Aid, 2004).

Data mining deals with the extraction of new data knowledge from large and complex data sets. (Daniel & Chantel, 2014). Public data and open government reports will be used as the corpus of this research. This study will use unsupervised data mining techniques, investigate and analyze the potential social impact of the policies guiding the exploration and production operations of IOCs on the host communities in sub-Saharan Africa (Katharina, et al., 2012). The process of data analysis and visualization will be used to discover previous unknown data, valid patterns and relationships. It will bring to bear the National legislations, the IOCs’ policies impact and other related socio-economic consequences on the dwellers of the areas where these operations take place (Stibu, 2016).

# 2.0 Background

Proven natural gas reserves in Africa are mainly concentrated in four countries – Algeria, Egypt, Libya and Nigeria in sub-Saharan Africa. They possess 91.5 percent of the continent’s proven reserves. This brings Nigeria’s undeveloped natural gas reserves to limelight and at such a target of IOC giants in the sector (AfDB & AU., 2009). In Nigeria, about 2.5 billion cubic feet of proven natural gas is flared per day, which represents 82.8 percent of the net gas produced. This is one of the common problems of the oil producing operations in Niger Delta that occurs regularly and leaves a negative impact on the health and environment of the host communities (Stephen, 2011). The oil and gas industry is widely seen as an industry with high negative impacts on the environment, despite the fact that it has improved its environmental performance (Greenpeace Russia, 2014). Hence Social Impact Assessment (SIA) focuses on managing the social issues associated with planned interventions and ensuring that the objectives of the proposed benefits of the project are achieved (Frank & Ana, 2011).



Figure 1: The Niger Delta’s landscape is scarred by oils spills, and clean-up operations are often inadequate

# 3.0 Aims and Objectives of the research

The main aim of this research is to investigate the impact of the policies and operations of IOC in sub-Saharan Africa using unsupervised data mining techniques. To this end, the objectives that will lead to achieving the goal are as follows:

* Investigate and analyse the operational standards of IOCs in sub-Saharan Africa
* Analyze the appropriate legislations of host countries
* Compare and analyze data collected from the public, government and IOC corporations in the sub-Saharan Africa
* Critically review the relation between the IOC operational standards and legislations of host countries
* To develop social impact assessment framework by using mongo dB (database) to manipulate the collected data
* To implement proposed social impact framework for adoption by stakeholders

Figure 2 shows the data summarization and visualization processes involved in data mining.

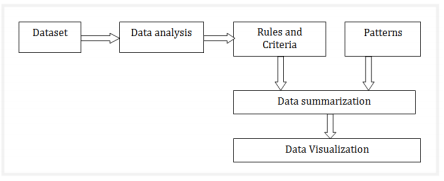


Fig. 2: Process of data analysis and visualization

## 3.1 Research Motivation

Social assessment of the operational issues of IOCs in Sub-Saharan African are begging for solutions from the right minds. With this research, a very reliable social impact assessment frame work can be presented to all stakeholders that can forecast to shape government policies as well as that of the international operators, and in so doing can became a novel if not a very relevant body of expertise providing enduring and fruit producing solutions to social issues arising from oil and gas operations. Such analytic framework as proposed in this research can bring the social vices in Nigeria for instance to a halt.

# 4.0 Research Questions

The following are the likely research questions:

* How is data mining related to the sustainable policies and operations of IOC in sub-Saharan Africa
* What data are available and can be sourced
* What is the significance of data mining to social issues (population, health, environment, politics, financing, poverty, wealth creation etc.)
* What are the identified problems relating to social science (I don’t mean Engineering problems), the social problems arising from the operations and guiding policies of oil companies in sub-Saharan Africa

# 5.0 Research approach and Methodology

This research will be conducted following the Design Science Research Methodology (DSRM). Although several methodologies exist for digital research in Information System (IS) project, but DSRM will be used because it has great potentials to support this research work. It is a very important methodology that supports the creation of artefacts in IS research (Shrestha, 2018). DSRM equally meet these three objectives: consistency with prior literature, provides a process-based model for carrying out DS research, and can provide a mental model for presentation and evaluation of DS research in IS (Vaishnavi, 2004/17).

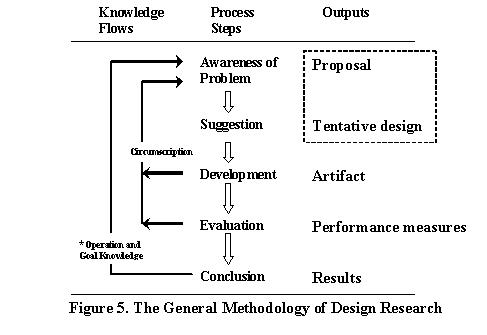


Fig 3: The General Methodology of Design Research

The research will be based on Machine learning techniques for data mining. Due to its advances in data analysis research, growth in the database industry and the resulting market needs for methods that can extract valuable knowledge from large data stores. Also, since the research is focused on extraction of implicit, previously unknown and potentially useful information from oil and gas industry data, government data and host community data.

The NoSQL database, which is a current approach for large and distributed data management and database design will be used in this research. The mainstream of our system logic will be entirely Big Data platforms and as such we are adopting NoSQL to break and transcend the rigidity of normalized relational database management system (RDBMS) schemas. There are several NoSQL database management technologies, but we are implementing our system in this research with the MongoDB NoSQL database management technology.

# 6.0 Research Plan

|  |  |  |
| --- | --- | --- |
| Year | Phase | Task |
| One | Exploration and Identification of research project | * Presentation of working prototype * Literature review * Identify data collection approaches * Attend scheduled meeting with supervisor * Gathering industry related data |
| Two | 1. Writing/Development of research content  2. Data Analysis  3. Development of framework  4. Presentation of working Tool | * Literature review * Collection of more data * Attend scheduled meeting with supervisor * Writing and structuring the paper into chapter titles * Investigation and Analysis of collected data * Development of proposed framework * Presentation of working tool |
| Three | Evaluation | * Testing and evaluation * Continuous writing, review and corrections of research content. * Submission of research paper to supervisor * Final amendments and submission of PhD thesis. |

# References

AfDB & AU., 2009. *Oil and Gas in Africa,* Chennai: Oxford University Press.

Christian Aid, 2004. *Behind the Mask: The Real Face of Corporate Social Responsibility,* London: Christian Aid.

Daniel, T. L. & Chantel, D. L., 2014. *Discovering Knowledge in Data: An Introduction to Data Mining.* 2nd ed. Hoboken: Wiley.

Frank, V. & Ana, M., 2011. *New Directions in Social Impact Assessment: Conceptual and Methodological Advances.* Cheltenham: Edward Elgar Publishing Limited.

Greenpeace Russia, 2014. *A Brief Overview of the Oil Spill Problem in Russia,* Russia: Greenpeace.

Julia, B., 2017. *Mixing Methods: Qualitative and Quantitative Research.* London: Routledge.

Katharina, M., Kanishka, B. & Hillol, K., 2012. Introduction to Data Mining for Sustainability. *Data Mining and Knowledge Discovery,* 24(2), pp. 311 - 324.

Michael, J. B. &. G. L., 1997. *Data Mining Techniques.* New York: John Wiley & Sons.

Roy, D. &. T. F., 2016. *Ohio Short Histories of Africa: ken Saro Wiwa.* Ohio: Ohio University Press.

Shrestha, A. A. C.-S. M. T. &. T. R., 2018. Benefits and relevance of International Standards in a design science research project for process assessments. *Computer Standards and Interfaces.*

Stephen, C. N., 2011. Climate change and energy implications of gas flaring for Nigeria. *Internation Journal of Low-Carbon Technologies,* 6(3), pp. 193 - 199.

Stibu, S. R. M. C., 2016. A Graph Data Summarization and Data Visualization Technique. *International Journal of Advanced Research in Computer and Communication Engineering,* 5(4).

Tomlinson, K., 2017. *Oil and gas companies and the management of social and environmental impacts and issues.* [Online]   
Available at: https://www.wider.unu.edu/sites/default/files/wp2017-22.pdf  
[Accessed 14 July 2018].

Vaishnavi, V. B. K. S. P., 2004/17. *Design Science Research in Information Systems.* [Online]   
Available at: http://www.desrist.org/design-research-in-information-systems/  
[Accessed 20 July 2018].