TALLINN UNIVERSITY OF TECHNOLOGY
School of Business and Governance
Ragnar Nurkse Department of Innovation and Governance

Martha Chukwu

ALGORITHMIC GOVERNANCE IN NIGERIA'S HEALTH CARE INDUSTRY – UBENWA AI SOLUTION

Master's thesis

Technology Governance and Digital Transformation

Supervisor: Anu Masso, Prof.

Co-Supervisor: Stefano Calzati

Tallinn 2020

I hereby declare that I have compiled the thesis independently
and all works, important standpoints, and data by other authors
have been properly referenced and the same paper
has not been previously presented for grading.
The document length is 13,853 words from the introduction to the end of the conclusion.

Martha Chukwu ……………………………
(signature, date)
Student code: 184529HAGM
Student e-mail address: marthachukwu1@gmail.com; machuk@taltech.ee

Supervisor: Anu Masso, Prof.
The paper conforms to requirements in force.

……………………………………………
(signature, date)

Co-supervisor: - Stefano Calzati
The paper conforms to requirements in force
……………………………………………
(signature, date)

Chairman of the Defence Committee:
Permitted to the defense
…………………………………
(name, signature, date)

TABLE OF CONTENTS

ABSTRACT

Big data has revolutionized human activities and their impact across many sectors, including healthcare. Privacy, safety, security, and human rights concerns remain heightened despite the many values it brings. Different jurisdictions are addressing these concerns differently, notably by developing guidelines and regulations. Social datafication, governance through data, data colonialism, and interoperability governance are emerging areas of research interest. Though behind, Africa is increasingly datafied with regulatory gaps whose potential can result in re-enforced inequalities and biases.

Localized data are often unavailable/ almost non-existent in Africa. Hence, digital solutions deployed in the continent are developed with data from a different social context, causing social inequality and bias. Ubenwa, a use case AI solution that uses the cry of a baby to detect birth Asphyxia, was analyzed to understand how discrimination and inequality can occur from its design and interoperability, how to avoid unintended social harm and how social context plays a role in algorithmic governance. Data governance frameworks were evaluated. In-depth qualitative interviews were conducted with a cross-section of data subjects, who will be impacted by such an AI solution; and health data experts who will use such a solution.

This research also adopted an online quantitative survey method to ascertain the state of data and AI governance in Nigeria, and understand the level of awareness and compliance in addressing algorithmic governance, data colonialism, and interoperability governance. The literature shows that while there are international frameworks for ethical AI, none exists in Nigeria. The interview and survey analysis results show that algorithmic governance is timely. The study found that its successful implementation is dependent on contextualizing initial algorithmic dataset, and regulating and enforcing its use with frameworks/ guidelines. The study made recommendations for improving and supporting algorithmic decision making in the country.

Keywords: Social datafication, Data colonialism, Algorithmic governance, Data-driven / Data-informed governance, Data ethics, Interoperability, Artificial Intelligence, Big Data.

1. Introduction

The proliferation of emerging technologies and their uses have ushered in the era of Big Data and Machine Learning. Digital solutions are now designed with knowledge and intelligence that equate and surpass those of the human brain. Automation and knowledge systems are becoming ubiquitous (Wirén, Mäntymäki and Najmul Islam, 2019). Collected data are used, stored, shared, and analyzed for business or data-subjects' benefits and improved decision making. Artificial Intelligence (AI) algorithms are now part of our everyday lives with increasing intrusive data sources. Collected information are sometimes shared and linked by data custodians, and this can happen in environments with little or no data governance regulation or operational guideline. In cases where the frameworks or regulations exist, there are no enforcement mechanisms. Concerns around privacy, security, data governance, ethics, and transparency remain a topical and touchy point between data subjects and data processors. These concerns birthed the European Union General Data Protection Regulation (EU GDPR) and similar jurisdictional data regulations (Pandit *et al.*, 2018).

Predicting disease outbreaks and modelling critical risk interventions are some of the advantages of Big Data in public health (Mayer-Schönberger and Cukier, 2013). Despite the inadequate Information Communication Technology (ICT) infrastructure, these data enabled digital health services are gradually being adopted in Africa (Akinnagbe, Peiris and Akinloye, 2018). Most intelligent healthcare solutions in Africa are developed in the western world and deployed for Africa. Often the deployment process in different socio-cultural contexts does not factor local input in its design. The risk of unethical breaches, including bias reinforcement, is on the rise. Measures to manage data access, usage, and management are limited. In order to mitigate these challenges in Africa, data governance frameworks are crucial. Though African countries are developing equivalent data protection regulations as GDPR, there is limited capacity and resources to domesticate them via enforcement. African countries will fully benefit from intelligent systems if they are designed and used responsibly and ethically.

This research aims to highlight the challenges of algorithmic governance systems and their effects on African society, with a focus on a use case AI solution deployed in Nigeria. As African governments embrace intelligent healthcare systems, this study will feature data subjects' readiness as these solutions and their societal consequence directly impact them. This study will also aid policy direction for the regulation of such systems. The recommendations will help tackle AI and data governance concerns in the country, as well as the African continent. There are ongoing researches focused on the data-intensive phenomenon. However, they are mostly focused on its quantitative aspects (amount of data) and interoperability. There is little focus on its social perspective (data exploitation and governance). The original focus of this thesis is to feature the social interoperability of datafied solutions.

This dissertation's structure is as follows: Chapter 1 introduces and maps the research study, why it is needed, and its focus. Chapter 2 presents the problem settings (the empirical case under consideration) and the research questions. Chapter 3 discussed the theoretical framework of the study, key themes relevant to the research. The research methodology used in the study is outlined in Chapter 4, including the strategic sampling and analysis methods. Chapter 5 presents the research findings. The study discussion section, including recommendations and limitations, are captured in Chapter 6. Chapter 7 sets out the conclusions to the research.

2. Problem setting and research questions

African countries have underperforming, underfunded, overstretched, and understaffed healthcare systems. The continent lacks the infrastructure to deliver quality health services and lacks medical supplies while also battling the devastating effects of severe healthcare professionals shortage (Novak and Bidwell, 2019). Her health sector is majorly dependent on foreign aid. Poor accountability of public office holders, long term wars, and corruption result in weak healthcare systems (Deaton and Tortora, 2015). The continuous rise in the use of the Internet and mobile phones has positioned the continent for technological leapfrogging, especially in rural communities. For example, the finance sector is contributing significantly to the economy through the introduction of various mobile financing services. In the healthcare sector, both international and locally developed digital solutions are utilized to improve healthcare and manage epidemics, like the Ebola outbreak (Akinnagbe, Peiris and Akinloye, 2018). The region is experiencing a steady growth of digitally enabled systems to support decision making and efficiency in healthcare services, hence facing increased bias or discriminatory results by algorithms. It is likely for algorithms to play a vital role in Africa's healthcare sector's decision-making due to its largely overstretched and understaffed workforce. Nevertheless, the continent currently lacks strong regulatory guidelines for algorithmic accountability and transparency to ensure these digital tools are used for social good. Its populace is non-literate concerning the societal effects of algorithms. This study contributes to discussions on governance through algorithms, how adopting a governance framework can reduce the effects of automated inequalities, biased, discriminatory or unfair results; algorithmic literacy would also be introduced as a standard, giving people the choice to use datafied services.

2.1 Empirical case under consideration

The empirical case under review discusses an AI technology being used in saving newborn babies from birth asphyxia in Nigeria, and the concerns around the datafication of same. Birth Asphyxia is a leading cause of death in infants and under-5 globally, accounting for an estimated 900,000 neonatal deaths annually (Spector, 2008). Birth Asphyxia is a medical condition that results from the lack of blood and oxygen flow to the brain of an infant; if not properly diagnosed and managed, it could lead to neurological conditions like epilepsy, cerebral palsy, development delays, and other defects (Aslam, Saleem and et al, 2014). In its early stages, asphyxia is difficult to detect through physical examination, except through blood gas analysis. The early detection of this condition can be a game-changer in reducing the global infant mortality rate, especially in resource-poor settings that lack adequate infrastructure, required skills, and equipment in the early diagnosis of this condition (Onu *et al.*, 2017).

Ubenwa AI solution

With a population size of about 200 million citizens spread across its four regions, the urgent need to deploy cost-effective and accessible AI tools are crucial to ease the overburdened health workforce in Nigeria. A Nigerian mHealth start-up developed an AI solution for the early detection of birth asphyxia and other related child health diseases through the cry of a child; this solution was birthed from firsthand experience by its principal innovator to the problems of birth asphyxia while volunteering in Nigeria (Onu, 2019). Ubenwa is interpreted as the 'cry of a child' in the Igbo language of Nigeria; it is a cry-based diagnostics mobile app that uniquely detects asphyxia in a neonate. The Ubenwa enabled AI algorithm audio processes an infant's cry with its computational capabilities using smart mobile devices. It then helps provide a qualitative assessment of whether or not the newborn has or is at risk of asphyxia. It generates an instant diagnosis by first recording the child's cry and then analyzing its amplitude and frequency pattern against its deep learning models, as shown in Figure 1 below. An initial dataset of about 1,400 infant crying samples of healthy and asphyxiated newborns was obtained from a Mexican databank. A machine-learning algorithm was applied to detect and identify the slightest presence of asphyxia. Test results have demonstrated an 85% sensitivity and 89% specificity (Onu *et al.*, 2019). The Ubenwa AI solution is cost-effective, non-invasive (no blood is taken, only needs the baby's cry), delivers results instantly, and requires no professional skills. In line with the SDG3's aim at ending preventable deaths of newborns and children under-5 by the year 2030 (*Children: reducing mortality*, 2018), Ubenwa is a timely response to the pressing need for affordable and sustainable tools to help tackle this challenge; it has the potential to be 95% cheaper than current clinical solutions (Onu, 2014).

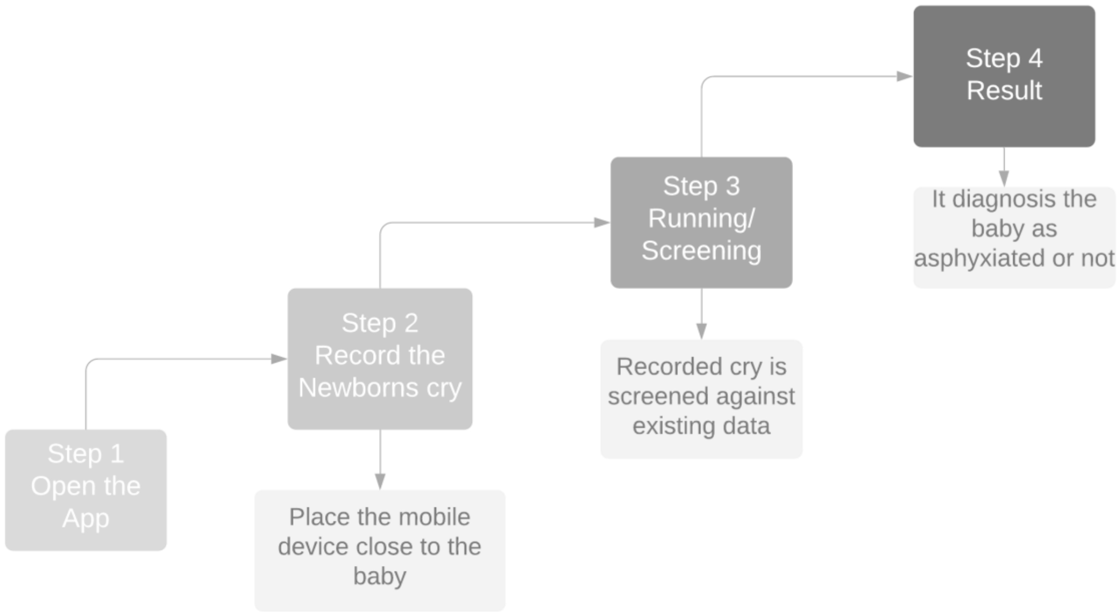


Figure 1: How Ubenwa works - a step-by-step process (Source: author's interpretation)

In the last quarter of 2019, the clinical study of Ubenwa kicked-off in Nigeria at the Enugu State University Teaching Hospital (ESUTH), and the Montreal Children's Hospital (MCH) in Canada (Onu, 2020). The app is being tested with real-life patients, to gather the 'largest' annotated infant cries.

Apgar Scoring method

In the face of lack of proper medical facilities to detect distress in newborns, medical personnel in the Nigerian healthcare system are trained to apply the Apgar scoring technique in the first and fifth minutes of an infant's birth, and later if necessary. The Apgar score is used to determine early signs of distress in the newborn, that could lead to further complications. The Apgar scoring method is a globally recognized standard physical assessment method for detecting birth asphyxia amongst other conditions; it was developed in the 1950s by Dr. Virginia Apgar (Leuthner and Das, 2004). The method uses five indicators, as shown in figure 3 below. Each indicator scored on a scale of 0 to 2, with 2 being the best score and a maximum total of 10.

The score is then used to determine if a baby is mildly or severely asphyxiated. Of the five indicators in the Apgar scoring system, the Ubenwa application addresses only one indicator.

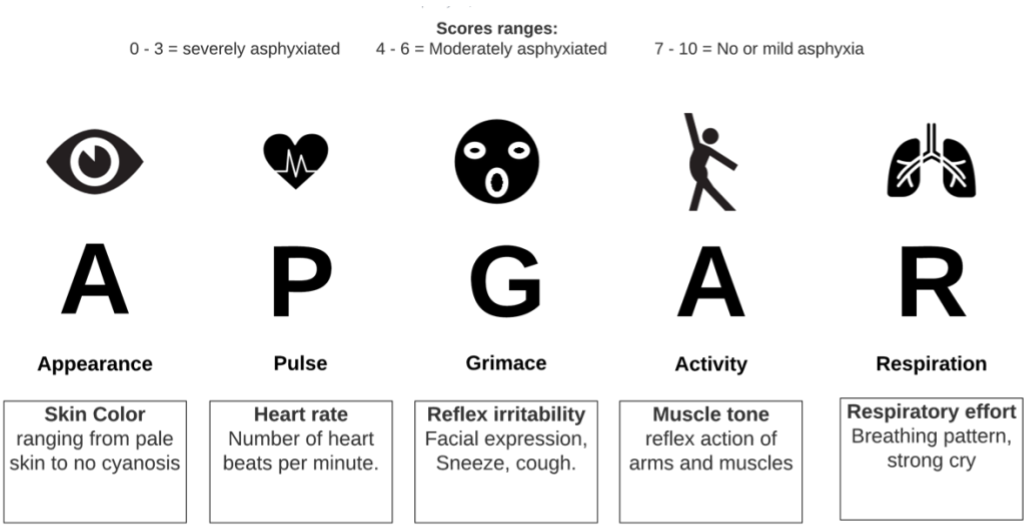


Figure 2: The Apgar scoring method (Source: author's own illustration)

Research Problem

AI systems require large data to train its algorithm, and data is not easily accessible or non-existent in Nigeria. Most AI systems currently deployed in Africa are developed in a different social context, lacking the input of localized data in its algorithm's design. The Ubenwa design lacks a relationship between the technology and the people because the initial dataset and development are not inclusive of the target population group. Ubenwa is utilizing an initial dataset from Mexico to train the algorithm developed in Canada, for testing in Nigeria. Mexico, a central American country, has a different diet, lifestyle, and possibly audio characteristics from a Nigerian. The developers in Canada might not consider cultural diversity, environment, and other factors while developing and testing the algorithm. Hence, the solution is not diverse and representative enough for Nigeria's implementation because limitations can stem from the used data composition.

Ubenwa's diagnosis is based only on an infant's cry, unlike the Apgar scoring method that considers other factors for detecting asphyxia. In an overstretched healthcare system, Ubenwa could play a key decision-making role in detecting asphyxia in newborns, thus exposing the babies to the risks of bias and unfair results, even though the healthcare worker makes the final decision. In comparison to the Apgar method, decisions are not dependent on algorithms developed with human assumptions. Ubenwa addresses only one out of five Apgar indicators, hence it could be referred to as not technically robust as the Apgar method. For instance, babies with weak or no cry could be wrongly diagnosed as asphyxiated. Either because their weak cries could fit into the asphyxiated samples designed into the algorithm, or no cry means they are automatically biased against since Ubenwa depends on only an infant's cry. Ubenwa is unable to apply any other human techniques in its diagnosis except through cry samples, unlike the Apgar method that scores different indicators for diagnosis. The research questions aim to uncover the concerns of the proposed end-users of Ubenwa, that is, Nigerian parents and experts, and further, understand how the available data governance and interoperability frameworks address these concerns.

2.2 Research questions

The research questions aim to measure the Nigerian populace's readiness for the introduction of datafied systems, while also uncovering how available governance frameworks in the country address the concerns of data subjects in using such solutions. The research questions are:

1. Based on the Ubenwa use case, what are the existing concerns of data subjects with the introduction of AI diagnostics tools for the development of healthcare systems in Nigeria?
    1.1. How does the different socio-cultural context in the development of Ubenwa affect interviewees' perception of its use or any other AI solution in Nigeria?
    1.2. In what ways can Ubenwa be improved for effective use in Nigeria, and globally?
2. How ready is Nigeria for algorithmic governance in its health sector?
    2.1 What are the processes/ procedures to ensure proper implementation of and protect citizens from datafied systems like Ubenwa?
    2.2 What level of awareness do the citizens have on algorithmic solutions and their socio-cultural effects?

3. Theoretical Framework

Big Data

The rapid growth of technology is gradually erasing the traditional human activities and interactions, translating, to massive volumes of electronic information. Overwhelmingly complex data is collected from diverse sources, making it increasingly difficult to apply traditional managing, storing, processing, or analysis approaches. The ability to collect and analyze these data deluge can be referred to as Big data. Researchers' understanding of the term Big data is wide and varied. Sam madden defines Big data as sets of data that is too big, too fast, or too hard for traditional data processing methods (Madden, 2012). Kaisler et al described it as the amount of data beyond the capability of a technology to efficiently analyze and manage it (Kaisler *et al.*, 2013). It can be described as large volume of data, both structured and unstructured, that is complex to process using traditional data processing and analysing methods (Urbinati *et al.*, 2019)(Philip Chen and Zhang, 2014). It can also be seen as the act of collecting large data sets from traditional and new digital sources to identify trends and patterns. From the diversely available definition of big data, (Taylor and Broeders, 2015) identified similarities of this phenomena as illustrated in figure 3 below; these similarities are referred to as the 5 V's of big data – Volume, Variety, Velocity, Veracity, and Value (Ishwarappa and Anuradha, 2015). The 5 V’s represent the main characteristics of the big data phenomenon, which over time the size of the data is no longer the focus for defining big data but the inclusion of these basic features.

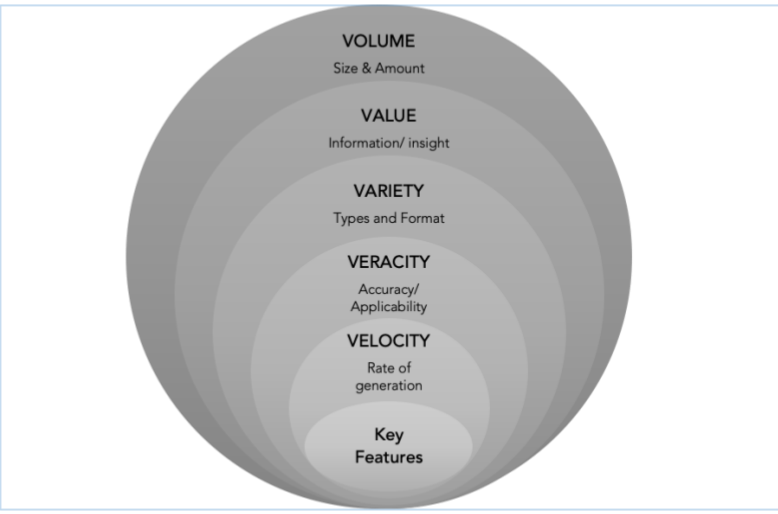


Figure 3: Characteristics of big data (Source: author’s own illustration)

Big data analyzes huge data sets (quantity) into valuable insights (Quality) in real-time (Resnyansky, 2019). These insights arise from event trends and social behavior. For example, big data in addition to other measures, was instrumental in curbing the spread of Covid19 in Taiwan, a country expected to have the second highest number of cases of the virus due to its proximity and relations with China (Wang, Ng and Brook, 2020). To aid case identification, the Taiwanese government integrated its national health insurance database with its immigration and customs database to generate real-time alerts during a clinical visit based on travel history and clinical symptoms. In addition to recent travel history to affected areas, high-risk citizens were tracked and monitored electronically through their mobile phones. This shows that data-based systems have contributed significantly to today's knowledge-driven economy; it maximizes improved decision and profit-making. Though large data processing comes with big promises, its platform has inbuilt constraints and limitations. It is limited to network and algorithmic restrictions, which influences its ability to process all necessary data needed for improved decision making. Hence, algorithms play a key role in shaping decision making processes, similar to the data limitations of the Ubenwa algorithm.

Algorithms

It has been shown over time that irrespective of its size, big data is only useful when it can be analyzed for benefits and improved decision making (George, Haas and Pentland, 2014). Algorithms uncover hidden patterns and trends in given data sets. An algorithm can be seen as a set of rules or functions that the computer uses to process data (Kitchin, 2017); they are socially constructed and deployed. According to Robin Hill's definition of an algorithm, he noted that algorithms could only 'accomplish a given purpose under given provisions' (Hill, 2016). Najafabadi et al. also inferred that an algorithmic process is dependent on the data represented (Najafabadi *et al.*, 2015). These definitions and descriptions share similar opinions on the effects of utilized data sets on the algorithmic decision process, hence fitting into the case study and this research's objectives.

Biased Algorithms

Data systems could be misconstrued as perfect due to their ability to draw up insights quickly. Platform users and few designers have little or no knowledge of algorithms having the capacity to be biased due to pre-existing cultural or social expectations. Limitations in its design could cause this biasness. It could be introduced when used by audiences not considered in its initial software design. Unintended bias in using data from different contexts may also arise, just as the Ubenwa app developed using initial datasets from Mexico and implemented in Nigeria. Concerning this research, Rob Kitchin (Kitchin, 2014) faults inaccuracies in data processes and scientific research, criticizing modern methods of lacking ethics and transparency for refining data collection processes to identify data patterns. He addressed the need for proper guidelines to ensure transparency in computational methods for data collection, especially for datafied systems leveraging on AI to improve user experience. In 2015, an African American was tagged as a Gorilla on his google photo app, while dogs have been mistaken for horses (Mullen, 2015). These examples show how social discrimination is just one out of so many challenges datafied platforms can create. This research will focus on the social perspective of algorithmic governance, which is its exploitation and governance. Three perspectives of data governance directly related to the dissertation will be discussed – Algorithmic governance, Interoperability governance, and Data colonialism.

3.1. Algorithmic governance

Datafication

Datafication is the process of transforming human activities into data used for different purposes(Flyverbom and Murray, 2018). Mayer and Cukier (Mayer-Schönberger and Cukier, 2013)) describe datafication as "unearthing data from materials that no one thought held any value"; this simply means data generated for one purpose can reflect valuable insights when applied for other purposes. Hence, datafication is an integral component of data-intensive systems. Digitization has opened opportunities for people profiling through networks and sensors; everyday activities like banking, education, healthcare amongst others are datafied for economic growth (OECD, 2013). Businesses and governments enjoy tremendous benefits from datafication, thereby translating to more datafication of social life and activities in the future (Dijck, 2017). Social science research is not left out as sampling techniques are usually not utilized with the advent of datafication (Chang, Kauffman and Kwon, 2014)(Vestoso, 2018). Over time, questions have been raised over the datafication of activities - who has access to these data and when? How and what is it being used/ used for? The consequences of the current wave of social datafication have the potential to disrupt socio-cultural activities. It could lead to discrimination, social bias, and injustice as with the Google photo app example above. These anomalies could result from the availability of too few, or as noted by Boyd and Crawford (Boyd and Crawford, 2012), in their big data review, excess or biased raw data during analysis.

Data-driven decision and bias

Bigger data does not automatically translate into better results. Once a dataset is unrepresented during the algorithm training, the outcome is flawed, thereby causing a bias or discrimination (Williams, Brooks and Shmargad, 2018). Koen Leurs et al. described datafied systems as being 'inherently discriminatory against already marginalized subjects' (Leurs and Shepherds, 2017), arguing that the current data discriminatory practices being experienced was requisite even before introducing intelligent data-based systems. These descriptions support this research's objectives and are further illustrated with the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) criminal risk assessment tool used to assess offenders in the US. The system is designed to use offenders' past criminal records/ data; psychological and biographical information are analyzed to rank an offender on different score levels (Freeman, 2016). In the case between Wisconsin vs Loomis, Eric Loomis was arrested and sentenced to prison for 7 years for a drive-by shooting in Wisconsin; the judge's sentence was based on him being scored as a high risk for recidivism by the COMPAS assessment report (Liu, Lin and Chen, 2019). Asides the high chance that Loomis, an African American, could have been unrepresented by biased samples in the dataset used in training the algorithm, the algorithm cannot determine whether the offender has become clean. It carries out evaluations based on previous criminal convictions. This bias against African Americans may have been introduced or reinforced via prejudices long held by platform developers. At the time of writing, hundreds of thousands of people protested in solidarity of the recent murder of George Floyd caught on camera (Taylor, 2020) and believed to be a symbol of widespread bias. These examples show the importance of data representativeness in the design of an algorithm. Ubenwa is utilizing data from a different country to build its algorithm for implementation in another country. Its initial dataset is unrepresentative of the social context it is meant for, hence leading to possible bias or discrimination against newborns in Nigeria. At this point, critical questions arise bordering around the social impact of technology and big data. 1.) When is it ethical to use technology? 2.) When is it appropriate to collect and use personal information? 3.) How do we embed sensitivity into these technologies to be considerate of individuals, society, and the environment? 4.) Who is accountable for using these technologies, as much of big data is often linked to personal information? 5.) What are the ethical responsibilities of data users and coders? 6.) What is the place of human-in-the-loop of data-intensive systems? These questions are derived from the concerns in using algorithmic solutions and are important in the context of this thesis. Machines/ technologies are built to be smart, but they lack the innate ability to learn morals, principles, ethics, or virtues; they lack human discretion in dealing with such challenges.

Governance by and of algorithms

Datafication and algorithmisation make governance more powerful (Mejias and Couldry, 2019). Governments are increasingly dependent on these data-driven systems to provide efficient public service delivery (Dencik *et al.*, 2019). Pascal Konig describes algorithmic governance as an avenue to create diversity, social inclusiveness, participation, and democratic responsiveness (König, 2019), while also noting its apolitical features - undisputed authority, designed to be highly responsive, datafication, amongst others. Over the last decade, studies have understood algorithmic governance from different perspectives. Some studies interpret this phenomenon to intentionally regulate or structure social contexts, particularly to suit preconceived goals. At the same time, other studies show non-intentional forms of social ordering through and with this phenomenon (Yeung, 2018). Researchers have identified this phenomenon to encourage ethical, social, and legal issues, including privacy, violations, bias and discrimination, and lack of transparency. Lepri et al. also describes the dark and good sides of data-driven decision making for social good (Lepri *et al.*, 2017), outlining its opportunities and positive effects. Yeung (2018) and Lepri et al. (2017) share similar opinions of algorithmic governance with Konig (2019) and proffered human-centric requirements to ensure positive data-driven disruptions. In line with this, Latzer and Just highlights four different approaches to tackle this decision-based model (Latzer and Just, 2020), as illustrated in figure 4 below. The described approaches focus on public interests, individual/ user rights, epistemic and normative concerns, and algorithmic systems' regulations. These approaches are relevant to this thesis and would be compared in the discussion section, with results from collected qualitative data.

Risk-based approaches

This approach calls for systematic risk identification and risk assessments, and appropriate choice of governance modes.

The focus is on the control of risks associated with certain types of algorithms e.g. social discrimination, violation of intellectual property rights, amongst others.

Ethics-based approaches

This approaches addresses the quality of evidence, the actions and traceability of algorithmic governance.

The focus is on bias, discrimination, unjustified actions, etc.

Human-rights-based approaches

This approaches tackles concerns bordering around privacy and data protection, freedoms of expression, enjoyment of human rights without discrimination, amongst others.

The focus is on the rights of an individual in algorithmic decision making as health diagnostics, social scoring, amongst others.

Principles-based approaches

This approach reflects on the need for appropriate legal frameworks and policies to regulate the effects of algorithmic governance.

The focus is on ensuring accountability, transparency, liability, etc. in algorithmic governance

Figure 4: Approaches to tackle algorithmic governance (Source: author's own illustration adapted from Latzer and Just's proposed framework (Latzer and Just, 2020)*).*

Data governance country case studies

Self and statutory regulations are currently in place to address the identified risks and concerns of algorithmic governance, e.g. the EU-GDPR. Unlike the GDPR that governs the protection of citizens data in the EU, the African continent lacks a unified data protection regulation. The individual nations have their policies based on their national needs, and mostly borrowing its contents from the EU-GDPR - rights of the data subject, definition of personal data, data processing requirements, consent, and penalties for non-compliance. In comparison to the EU-GDPR, these regulations lack depth, thereby leaving loopholes for misinterpretation. In Nigeria, for example, data subjects have the right to ownership of their data, but, in practice, it can easily be misinterpreted to suit a case by case scenario (Elebeke, 2020). In addition, weak disciplinary measures for data offenses and an almost authoritarian style of leadership is stalling the process of enforcing developed policies. In supporting the discussions, it could be concluded that an unregulated data-driven governance system in Africa risks having unhealthy consequences for its citizenry as inaccurate data can lead to misinformed decisions. Ubenwa poses potential damage and harm to end-users in Nigeria; it was designed utilizing Mexican data. Mexico and Nigeria are in 2 different continents with distinct features ranging from the environment to culture, lifestyle, amongst other characteristics. This practice is generally common in Africa, where datafied solutions are developed without localized data and imported into the continent for implementation, with the available governance frameworks not capturing its cause and effect on its populace.

3.2. Interoperability governance

Interoperability is fundamental to the growth and success of modern technology. Platforms are able to access, exchange, and share information efficiently, irrespective of their technology, location, or institution. According to the European Interoperability Framework (2004), interoperability is the ability to exchange data amongst ICT tools and enable the sharing of knowledge and information (EuropeanCommunities, 2004). For example, Phones are able to connect and communicate with each other automatically, irrespective of the brand, call mode (either landline or wireless), and in most cases, the network provider. In healthcare systems, interoperability supports the exchange, interpretation, and use of patient’s data within and across organizational boundaries, regardless of the application, vendor, or technology. Medical services providers can easily share patient information, thereby advancing the effective delivery of healthcare for individuals and communities. In the context of this thesis, the focus will be on the social perspectives of an interoperability system, its effects on society and the governance of same.

Socio-cultural interoperability

Interoperability is a key component in recording the success of data-driven solutions in a digitized society (Davies *et al.*, 2020). According to Das et al. (Das and Mahapatra, 2012), interoperability supports the exchange and use of certain geographic data between communicating entities, based on specific agreements, and Lesh et al. (Lesh *et al.*, 2007) describe it as the communication between information systems and humans, and vice versa. These definitions describe socio-cultural interoperability as the connection between the environment and the human processes involved in information exchange. Hence, supporting the objectives of this research, because achieving a successful interoperability system calls for an existing relationship or sync between the technology and the people. Algorithms need not only imitate the human learning process but equally understand the social contexts and personal/ individual experiences to function better (Masso and Kasapoglu, 2020). The experiences of the local populace are crucial for every interoperability system, irrespective of the institutional structures regulating them (Rubinstein, 2014). In the case of Ubenwa, the local Nigerian data was not included in the design of the system, asides that the idea was founded by a Nigerian. This concept is illustrated better with the AI facial recognition system currently deployed in Zimbabwe.

The Zimbabwean government is collaborating with a Chinese start-up company to provide mass facial recognition programs in their country (Raji *et al.*, 2020). Local data would be captured through CCTV cameras, smart financial systems, amongst other national databases, to retrain the algorithm. Recent studies from the Massachusetts Institute of Technology's (MIT) Media lab shows that facial recognition AI in China has been trained with predominantly white and male faces (Buolamwini and Gebru, 2018). Therefore, the initial algorithm utilized has no pre-existing relationship with the Zimbabwean populace. The researchers observed the accuracy of three major facial recognition software providers in identifying the gender of a person from a picture and discovered the error rate for identifying a lighter-skinned man was 1% compared to 35% for darker-skinned women. In summary and support of previous descriptions, input data plays a significant role in the accuracy of an AI system. Otherwise, it is bound to fail like the facial recognition system in Zimbabwe.

The initial Ubenwa algorithm was not contextualized to the deployed locality, just as the Zimbabwean face recognition system. Existing research does not support the use of data from one social context in another due to socio-cultural differences. Studies have shown that social determinants, environment, and behavioral factors account for 80% to 90% of health outcomes in a population, with medical care only accounting for 10% to 20% (Hood *et al.*, 2016). These non-medical factors are described by the World Health Organization (WHO) as the Social Determinants of Health (SDoH), that is, the wider conditions of the daily life of the populace, e.g. where they are born, live, grow, work, amongst others (WHO, 2012). As much as Ubenwa was developed with the cause to improve on health systems delivery in Nigeria, environmental, behavioral, and socio-economic differences between an average Mexican and a Nigerian play a significant role in encouraging bias and discrimination in its diagnosis. Based on the identified concerns and gaps, there is a need for governance frameworks or guidelines to ensure that algorithmic systems consider socio-cultural interoperability by design. These frameworks would support algorithmic systems to contextualize solutions considering social and cultural factors for effective and auditable implementation.

Interoperability governance

Interoperability governance was defined by Wimmer and his / her colleagues as a system that provides the enabling framework, guidelines and policies for decision making; it supports the growing usage of interoperable platforms (Wimmer, Boneva and di Giacomo, 2018). These standards engage the technology, services, and implementers, ensuring they comply with the right processes to data quality. The established governing body consists of broad representatives to define set standards on data management practices – what, how, to whom, and when data will be shared. The developed framework defines the participants, policies, procedures (technical, operational, and legal requirements) and data sharing agreements for the protection and disclosure of personal information (Kelley, Feldman and Gravely, 2016). The European Interoperability Framework (EIF) is an example of an open standard that interconnects many online public services in Europe. It links all stakeholders in one single digital network and guides the use and sharing of information across different European territories (Kouroubali and Katehakis, 2019). Unlike the EIF, intraregional connectivity in Africa is almost non-existent due to weak infrastructure and connectivity challenges and a lack of policy coordination between governments. This has led to the short duration or general failure of digital health initiatives implemented in the continent.

Standardization and interoperability are significant challenges faced in the region, and algorithmic solutions like Ubenwa require regulatory frameworks and guidelines to function effectively. In Nigeria, there are established guidelines to regulate and manage the use of personal health information and address other challenges related to health data exchange. Some of these include the Data Interoperability Standards 2016, National Health Act 2014, the National Data Protection Regulation (NDPR), amongst others. However, Nigeria, and other African countries, do not have frameworks to regulate AI systems in their countries. With no guidelines to ensure that deployed AI solutions in the country are designed to fit local purposes, the tendency for failure is imminent. There is widespread introduction of alternative digital solutions in Nigeria's healthcare industry but limited by the inadequate capacity to handle the implications of algorithmic systems without laid down guidelines and regulations. The development of AI regulations like the EU-GDPR is crucial to the success of algorithmic governance systems in Nigeria and Africa.

3.3. Data inequalities and colonialism

The concept of data colonialism and inequalities is a global issue. African data is largely unavailable and untapped. Technology companies like Google and Facebook are moving into the African market to gain access to localized data sets for profit-making (Coleman, 2019), by building network connectivity infrastructures to support their data collection drive. Algorithmic solutions like Ubenwa would also be collecting the largest database of annotated infant cry during its pilot testing in Nigeria (Onu, 2020). Personal information is collected to utilize such systems hence supporting the new social order run by capitalists. Datafication supports the notion of data being king in today's digital society, as the power lies in the hands of those who have access to data. With the high percentage of data illiteracy, unsuspecting end users are unknowingly exposed to social inequalities and bias or discrimination by using these systems. With weak data protection laws to protect its citizenry from data colonialism, the continent is yet to grasp the implications of accessing personal information by algorithms.

Data Colonialism and surveillance

History records that early colonizers built their empires by acquiring landed territories and exploiting their natural resources. Unlike historical colonialism, where physical invasion was used to force acceptance and non-resistance, data relations are willingly entered into. Quintillions of behavioral data are being harvested from personal lives and experiences for economic benefits (Thatcher, O'Sullivan and Mahmoudi, 2016). Data remains the driver of growth and change in the digital era. The benefits of staying connected and accessing faster and easier services make control of activities through databases and networks seamless. Terms of agreement to use digitized solutions are mostly accepted without reading or understanding its implications as they usually contain lengthy and ambiguous legal jargons difficult to understand (Brunon-Ernst, 2015)(Borgesuis, 2015). Governments are indirectly empowered to dominate and control. Asides government, large tech companies are equally empowered because they have access to personal data e.g., Facebook, Amazon, Airbnb, amongst others. These service providers have been alleged to have more information about our personal lives than our families and loved ones (Sorescu, 2017).

Couldry and Mejias (Couldry and Mejias, 2019b) defined data colonialism as "an emerging order for appropriating and extracting social resources for profit through data", a process that normalizes the exploitation of human beings through data. People think they have lost the benefits or presumed benefits of participating in the social network site whenever they refuse to give up their data or participate on the network (Mejias, 2013). Millions of data are harvested via social media platforms for capital, creating opportunities for social discrimination and behavioural influence (Couldry and Mejias, 2019a). Personal data could be transferred to corporations to generate profit, and influence consumerist decisions, thereby transforming data into money and power. Case in point is the Cambridge Analytica allegations on how behavioral data was monetized; the data of millions of American Facebook users was accessed to influence political decisions of susceptible voters through microtargeted messaging (Isaak and Hanna, 2018). Machine learning and AI was adopted to predict and influence individual behavior and psychologically motivate their political decisions (Ward, 2018). These kinds of platforms have also been alleged to have incited the ethnic cleansing against the Rohingya minority in Myanmar (Venier, 2019). These examples support the notion that data colonialism is built on an existing stage of capitalism, sharing similar features with historical colonialism like the appropriation of resources, value extraction, amongst others.

Just like historical colonialism, data are not stored locally; data centers are located mainly in Europe and the US, a hindrance to digital industrialization in the global south. Local cloud companies are unable to compete and develop links to their local IT sector. In the context of this thesis, the Ubenwa solution is being pilot tested in Nigeria, with the aim of gathering the largest annotated infant cries. For a continent where data collection is a major challenge, this opportunity could give rise to colonialism and surveillance. Data is king and the driver of growth. These collected sample data could be linked to parent's data, which could in turn, be used for other purposes than intended. This is a cause for concern as the patient's privacy could be breached.

Data Inequalities

There is growing concern that people in the global south are subject to data colonialism, exploitation from data, and information. According to Virginia Eubanks (Eubanks, 2017), the vulnerable and poor are the worst hit in a system of continuous surveillance, i.e. dataveillance. They are unable to fight back when being victimized, stereotyped, and excluded by data-driven decisions made by the government, service providers, and credit raters. She illustrated the failures of such automated systems using case studies from the healthcare and social services systems in the US, and how they automate class and racial inequalities. Linnet Taylor’s (Taylor,

2017) views supports Virginia Eubanks writings on discriminatory data systems and the exclusion/ marginalization of the disadvantaged, advocating for people to be justly treated and having the right to appeal an algorithm-based decision. Hence, the need to implement a data justice system to regulate and ensure ethical compliance through dataveillance with the rise in data-driven policymaking. He further proposed a framework for data justice to allow opting out of data collection processes, preserve autonomy with technologies, and provide protection from data-driven discrimination. The proposed framework is built on three pillars – 1.) Visibility, considering privacy, and representation. 2.) Engagement with technology, focus on autonomy with technologies, and sharing in data's benefits. 3.) Non-discrimination, the power to challenge bias and discrimination. Taylor's proposed framework seeks to deal with data technologies related to human needs and suggests it operates as a part of core governance principles. Implementing data justice as a core of global governance principles would address the challenges of discrimination and privacy, especially in Africa, where people lack basic rights to privacy and protection.

In Nigeria, as in most African nations, data is mostly collected and used without the consent of data subjects. China has been severally accused of exploiting the African continent and extracting her data (Daouda, 2012) (Heffron, 2017). This is likened to the example of the facial recognition system being implemented in Zimbabwe, China is collecting African data to improve on its facial recognition database, with the privacy and freedom of the people at risk (Sachikonye, 2019). Digital tools are imposed as 'innovation' for informational capitalism without the consent of the people whose problems are being solved (Taylor and Broeders, 2015), supporting the exclusion by inclusion principles by social platforms. The lack of appropriate data regulations to activate datafied solutions for social good is a major challenge (Carman and Rosman, 2020). The lack of local data in the design of algorithmic systems could lead to unfair bias and discrimination, thereby promoting inequalities against the people. People are unable to opt-out of using these systems mainly because they are unaware of its effects on society. The continent has a high percentage of data illiterates, further widening the uneven power structure between developers (the capitalists) and the end-users. This does not meet ethical and sustainable requirements. There is a growing concern for algorithmic systems to be understood and held accountable for their decision-making processes, therefore the urgent need to develop and ensure compliance to strong ethical frameworks for the region. For the effective performance of AI solutions, a part of these regulations would consider mandating localization of personal data in Nigeria, and Africa generally.

4. Research Methodology

This study utilizes a mixed research method as it aims to gain an understanding of the phenomena through the unpacking of participants' opinion about datafied solutions like the Ubenwa app. Trends of thoughts and opinions of the participants were uncovered during the research. This exploratory research would adopt a non-experimental research design, as the study intends to understand a social phenomenon. An online quantitative survey of data subjects and qualitative in-depth interviews of data subjects and key stakeholders relevant to the study in Nigeria was conducted.

4.1 Qualitative in-depth interview method

A qualitative in-depth interview method was adopted to obtain detailed information on the interviewee's behavior, experiences, and thoughts about datafied solutions like Ubenwa. This method was chosen because the research involves a small number of participants and is focused on a particular product and situation. The core of this method is to explore the perceptions of key data experts in Nigeria on critical AI and data governance issues, using the Ubenwa use case. The interview will capture the concerns, perspectives, and reactions of parents on the Ubenwa solution and existing data malpractices, while assessing their data awareness and literacy level. One-on-one direct engagement with the participants was carried out face-to-face and via telephone, arrangements was made based on participants' preference due to the current pandemic and enforced social distancing. An interview guide consisting of semi-structured questions was developed to facilitate the interview and understand participants' complete views. These series of open-ended questions adopt a conversational tone to allow for flexibility during the interview; that is, new questions can arise from an interviewee's response for further scrutiny. Unlike most other AI health apps available in Nigeria, which are Patient-facing, Ubenwa is a health facility app that is provider-facing. The app is only available to the developers, so participants have no prior knowledge of it. Hence the walk-through method was adopted during the interview. This method is well suited to this study as a critical analysis of an app is being conducted (Light, Burgess and Duguay, 2018). For a better understanding of its functionalities, picture posters were used in describing and walking through the different interface of the app during the discussions. The posters contain the step-by-step documentation of the Ubenwa app screen features and process flows. The posters are available as appendix 1 to appendix 3 at the end of this thesis.

Due to the Covid19 pandemic and the enforced social distancing, an online interview method was adopted as an option where direct engagement is not feasible. For the qualitative interview, body language, facial expression, and other nonverbal signs are important for contextualizing the interviewee. To enable real-time discussions, the Skype video conferencing software was used to allow for an interview that closely resembles the face to face communication; participants can be seen while conversing. This interview method could be subjective due to technological requirements, but it will help to eliminate geographical barriers; it is convenient and provides an increased level of privacy (Nehls, Smith and Schneider, 2015). All participants were contacted ahead via telephone calls to inform them of the purpose of the study and inviting them to participate in the interview. Those who expressed their interest in participating were provided an informed consent (participant information sheet) ahead of the interview, explaining the objectives of the study, confidentiality, reference person, and the interview procedure. Interviews were conducted anonymously to protect participants' identities, especially the data experts who directly engage with the government as employees or consultants. All interviewees provided verbal consent to participate in the interview. An online survey will be conducted to measure the general opinion on the background information of healthcare governance, but the main method and focus of analysis rely on the qualitative data.

4.2 Strategic sampling method

To get a background information on the general awareness and readiness to use datafied solutions for healthcare in Nigeria, an online survey was conducted. The survey form was designed using the google docs tool, with questions on the Ubenwa AI solutions, data governance, and tech awareness on algorithmic governance in Nigeria. The target populace are Nigerian parents (both mothers and fathers) and the form was shared mainly via the WhatsApp social media platform. A large number of Nigerians have smart phones with access to the Internet; social media platforms are the commonly used communication channels in the country and would give a wider reach. The form was shared via personal contacts on WhatsApp and emails, with a request to forward to other parents. The minimum expected response is a hundred parents using these proposed channels. At the end of the survey form, respondents interested in participating in future studies on similar topics were requested to share their email addresses.

The purposeful sampling method was adopted for the in-depth interviews, to ensure the variation of the perspectives on algorithmic governance (Suri, 2011). Nigeria has a population size of about 200 million citizens spread across its four regions. It is a diverse country made up of 250 ethnic groups and different religious beliefs. These ethnic groups are spread across the regions, with some sharing similar cultural beliefs. Of the total population, about 62% is literate, over 50% live in the urban areas, and 40% live below the poverty line. With the diversity of the Nigerian populace, the purposeful sampling method was beneficial in capturing the diverse perspectives of the people. A total of 12 participants (stakeholders) was interviewed for this research – six data experts and six data subjects.

The data experts are key stakeholders in digital health in Nigeria, consisting of health officials and individual researchers. They represent potential decision-makers who will later use the Ubenwa app in health decisions, with background experience in Nursing, Medicine, and informatics, and App development. The data experts are made up of four males and two female interviewees, all possessing over five years of relevant professional experience in Nigeria. Their experience and expertise in Nigeria's healthcare sector will help in assessing the Ubenwa application for diagnosis.

The parents, representing the end-users of the Ubenwa app, will be interviewed solely to assess user experience, which is very crucial for every algorithmic solution. The direct target group of the Ubenwa app will not be studied as the aim of this thesis is not to conduct medical research, but social research about the perceptions on the app. For instance, parents whose children belong/ have belonged to the direct risk group of birth asphyxia will be excluded from this interview. The parents are selected based on their familiarities with the use of apps and mobile technologies. Only parents with smart-phones and children less than 12 months were interviewed for a better reflection of the Ubenwa app. Their educational and professional experience was equally considered for an easy flow of discussion during the interviews. In Nigeria, caring for a child is a major responsibility of the mother hence the parent sample size was narrowed to include four mothers and two fathers; the fathers were selected based on their experience in caring for their babies. All parents are selected to represent the three major ethnic groups.

4.3 Analysis method

The interview data are collected by combining open-ended questions with projective techniques and walk-through method (Light, Burgess and Duguay, 2018) (Soley and Smith, 2008). The conducted in-depth interviews were audio-recorded and transcribed verbatim i.e., word for word, to reflect and affirm theories (Davidson, 2009). After which a thematic analysis method was applied to the transcripts using a combination of manual techniques and computer-aided analysis. The thematic analysis method is used in identifying and analyzing qualitative data patterns (Braun and Clarke, 2013). The flexibility of this method will help examine the different perspectives of the interviewees and generate insights by stressing similarities and differences in the data sets. Individual data was analyzed against interview questions and participants' responses and then compared with other data based on specific themes. The research adopts the six-phase guide proposed by Braun and Clark in conducting thematic analysis, as illustrated in figure 4 below (Braun and Clarke, 2006). Themes were identified with a deductive approach, ideas generated from the collected data was analyzed against the theoretical framework represented in a data analysis software.

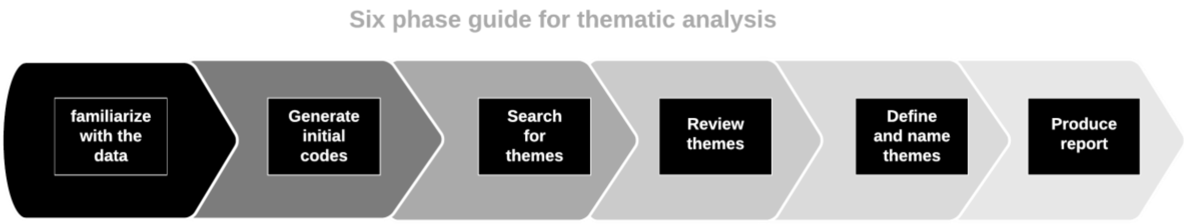


Figure 5: The six-phase guide for thematic analysis (Source: author's own illustration of Braun and Clarks analysis guide)

A Qualitative Data Analysis Software (QDAS) was used to aid sorting and organizing of collected data set. MAXQDA software is utilized for qualitative and mixed-method research. It aids meaningful interpretations through sorting and arranging information, examining relationships, and combining analysis. The software can code different types of data (e.g., text, audio, video, images, surveys, and others), and would simplify the analysis of the transcribed in-depth interview texts and other data formats (Kuckartz and Rädiker, 2019). Codes were applied to identify and categorize data into common themes, patterns and relationships. This software would aid in creating simplicity in the data analysis process and improve efficiency (Woolf and Silver, 2017).

5. Findings

5.1 Survey result

The aim of the online survey is to get background information on the general awareness and readiness of datafied solutions for healthcare in Nigeria. The designed form was made available online from the 14th to 30th June 2020, via email and majorly through the WhatsApp social platform commonly used in Nigeria. A total of 174 responses was received, exceeding the minimum expected number of 100 participants. Of the total respondents, 170 currently resident in Nigeria. The socio-demographic details of all participants are shown in figure 6 below.

Socio-Demographic information of survey participants

Gender	No. of participants	Age range				Educational qualification			
		19 - 24	25 - 35	36 - 45	46 and above	Secondary	Bachelors	Masters	PhD/ Doctorate
Male	86	0	16	46	24	1	36	46	2
Female	88	1	32	40	15	2	43	41	3
Total	**174**	**1**	**48**	**86**	**39**	**3**	**79**	**87**	**5**

Figure 6: Background information of survey participants

Over 80% of the total respondents were unaware of any AI solution currently deployed for healthcare management in Nigeria; figure 7 below shows the highest percentage of total respondents are willing to trust or accept the use of AI for healthcare diagnosis in Nigeria. A high percentage of the respondent were equally willing to accept the use of AI for diagnosis of their babies; this interprets the readiness of the Nigerian populace to utilize AI systems for diagnosis.

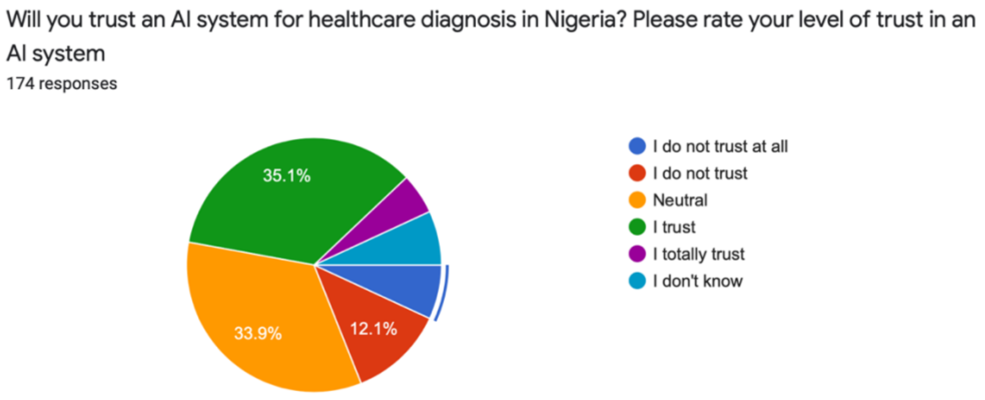


Figure 7: Trusting an AI system for diagnosis

Despite the different socio-cultural context in the development of the Ubenwa solution, a high number of respondents accept its use for diagnosis of asphyxia in their baby, as shown in figure 8 below. They indicated their trust in the Ubenwa tool, irrespective of its initial dataset. The survey also shows that 46% of the total respondents have no idea of data governing policies in Nigeria; they are unaware of current data protection laws in place to protect them from data malpractices. Only 9 of the total respondents have excellent knowledge of these laws; this could be linked to non-enforcement of these policies; hence citizens viewing data breaches as common practice.

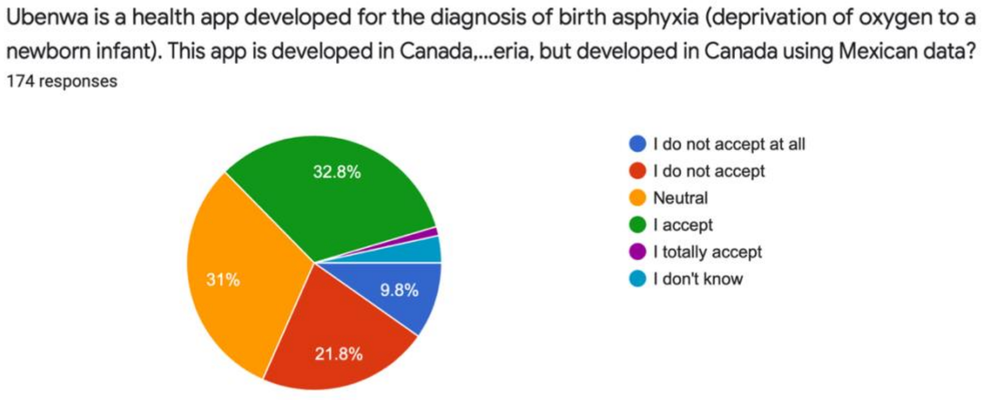


Figure 8: Accepting the use of Ubenwa

In view of data awareness gaps, most respondents (160 persons) recommended data awareness programs (training, educational events, amongst others) on AI systems and its impact on Nigerians. Of the total responses, 46 respondents suggested all the areas of focus be considered for the awareness creation. Figure 9 shows data protection rating the highest amongst the focus areas for awareness. A participant also recommended AI apps as a source of quick self-help, as another focus area for the awareness programs.

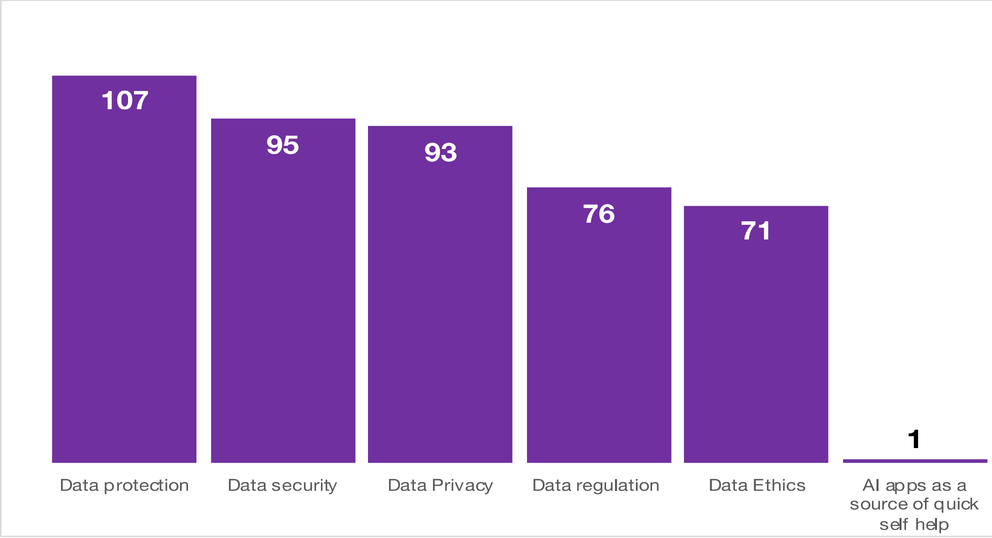


Figure 9: Focus areas for data awareness creation

5.2 Qualitative analysis result

Background information

The in-depth interviews with the 12 participants were conducted both face-to-face and online; 6 parents and 4 data experts were interviewed in their homes, while 2 data experts were interviewed via skype due to proximity challenges at the time of the interview. On average, interviews lasted about 45 minutes, with some lasting over an hour. All participants are educated and professionals in their various fields, with years of practice in Nigeria spanning over four years; they all have experience in the use of smart-phones and apps. Figure 10 below gives a brief background description of all interview participants, providing details on the age and number of their children and other socio-demographic information.

Participants Background Information			
Parents			
Participant	**Number of children**	**Age of children**	**Professional experience**
Parent 1	2	3 years; 9 months	Nurse with 6 years experience
Parent 2	1	12 months	Communications specialist for 8 years
Parent 3	3	9 years; 7 years; 12 months	Business consultant for 15 years
Parent 4	2	3 years; 9 months	HR and administrator, 5 years experience
Parent 5	1	10 months	Office adminsitrative personnel with 4 years experience
Parent 6	2	4 years; 11 months	Business woman with 6 years experience
Data experts			
Participant	**Professional experience**	**Data governance experience**	**AI experience**
Data expert 1	Digital health consultant	Yes; 11 to 12 years supporting use of technology in the health sector	Only research based; There is not a lot of AI experience or tools in Nigeria
Data expert 2	Nurse with 8 years experience	Yes; in the health insurance sector	Non
Data expert 3	Nurse/ Midwife for 10 years	Non; only Clinical experience	Non
Data expert 4	Head of FCT eHealth desk	Yes; Overseeing health data governance in the FCT	Only from research but no professional expereince
Data expert 5	Health data analyst for over 5 years	Yes; at both national and state level	Non
Datat expert 6	Medical doctor for over 10 years	Non; only Clinical experience	Non

Figure 10: Socio-demographic information of all interview participants

The transcribed audio recordings were imported into the MAXQDA version 2020 for reading, coding, and analysis. The unit of analysis selected was relevant to the study's research questions. The researcher read through the transcript thoroughly to ensure complete immersion of data collected; this helped to create open coding, categories, and themes. Memos were created while comments were added where necessary. Data were coded by tagging keywords within the transcripts in line with research objectives while assigned codes were labelled and categorized. Codes were grouped into themes to represent concepts; themes were interpreted in line with research objectives that form the basis for explaining the study's findings. Frequency of occurring responses with emphasis from the respondents formed bases for the categorization that eventually developed into themes. The MAXQDA software for analysis helped in linking relationships among codes, and also visualization of the data. The results of the analysis generated from the MAXQDA were reported in line with the objectives of the study, and deduced from the identified codes and themes, only figures relevant to the research are presented.

Report of interview findings

On the use of initial datasets from Mexico, parents and data experts shared similar opinions, mostly negative. As shown in figure 11, Participants did not support the use of Mexican data to train the Ubenwa algorithm, sharing similar opinions on the app being used in the country where initial data was gotten.

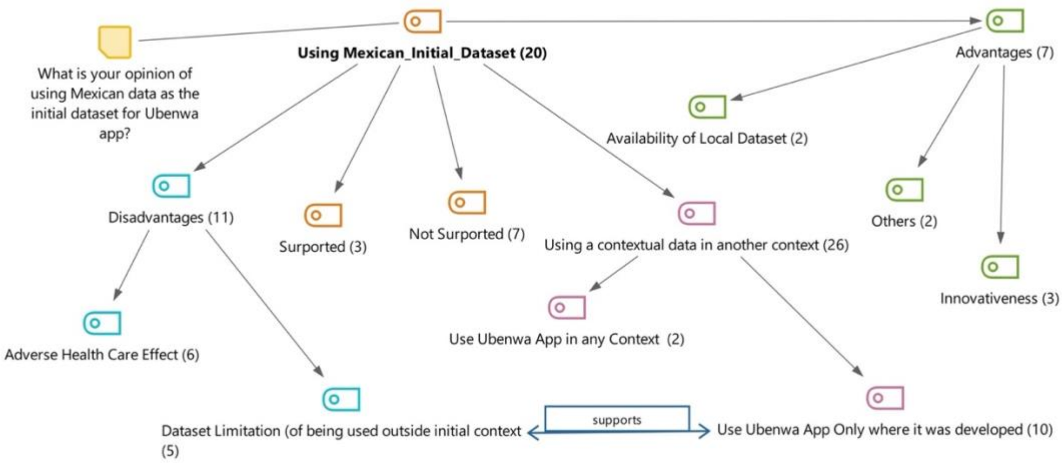


Figure 11: Using Initial dataset from Mexico

The data experts were disappointed with the Ubenwa solution utilizing the initial dataset from Mexico. They expressed their concerns around bias and misdiagnosis against Nigerian babies, explaining how wrong it is to utilize data set from one context to implement in another. The challenges of Nigeria being a socially and culturally diverse country was made mention of, stating that data from one region might not be accurate for use in another. Data expert 1, a digital health consultant with vast experience in Nigeria, and other African countries said:

> *I object strongly against that, no matter what the reason was, the application should have been designed within the Nigerian context, if you have a data from one region, it is not definitive that that data would be valid and would be able to generate accurate data for another region; (Data expert 1 interview transcript, Pos. 44)*

Another data expert made mention of an ongoing research that shows babies cry in different languages, which could increase the chances of bias and misdiagnosis against Nigerian babies. In her words, she said:

> *as much as I love the app and its purpose, the Mexican data would not be suitable for use on a Nigerian baby. Babies cry in different languages, and the last I checked, Mexicans speak Spanish while Nigeria has over 200 different languages. So, you do the math. (Data expert 6 interview transcripts, Pos. 43)*

Going by the research of babies crying in different languages, there is high risk of bias or misdiagnosis with the use of Ubenwa, as the country has over 250 ethnic languages. In disagreement, data expert 5 did not see anything wrong with Ubenwa utilizing data from another context to train its algorithm. He stated that such practices are okay if there is no readily available data to train the algorithm, as in the case of Ubenwa where Nigerian data is unavailable. In his words, he said:

> *Using Mexican data to develop the application is not a problem as far as the data is good; if the data is good then it's good. Especially since Nigeria does not have quality data for such application system, so using datasets from other places where we trust the data that is good. (Data expert 5 interview transcripts, Pos. 43)*

Contrary to other data experts, he further encouraged such practices in regions where data is not available clearly alluding that:

> *half bread is better than none. (Data expert 5 interview transcripts, Pos. 45)*

Sharing similar opinions with the experts, parents discouraged the use of data from one social context in another. Their major concerns bordered around misdiagnosis of their babies due to differences between a Mexican and a Nigerian. Though the parents have little knowledge about Mexico as a country, their knowledge and perception of the south American country stemmed from Mexican movies they have watched previously, and not personal experiences. Parent 2 clearly stated she was paranoid knowing Mexican data was used for Ubenwa's development. In her words, she said:

I'm actually watching a movie from there right now and it's leaving me very paranoid; so I don't know if I would rely on anything from Mexico (Parent 2 interview transcript, Pos. 30)

Also based on his perceptions from watched movies, another parent made a strong insinuation about Mexican babies. In his words, he said:

most of those children in Mexico, over 15% of them would have hard drugs in their blood line because of what happens in their environment. (Parent 3 interview transcript, Pos. 57)

In support to data expert 1, a parent shared the same views on Nigeria being socio-culturally diverse and its adverse effects on the use of Ubenwa. In his words, the uniqueness in the Nigerian environment, culture and other factors would affect result output with collected data within Nigeria itself. In his exact words:

the climate is not same, the weather, the region totally varies talk more of a different continent totally; (Parent 4 interview transcript, Pos. 50)

These views show that the parents are concerned that the utilized data might not fit into the Nigerian cultural space. Another parent stated she had no concerns about the dataset used for Ubenwa, her worry was focused on Ubenwa's effectiveness. She shared same views with other parents on data collected in one and used in another social context; giving an example, she says:

In most cases, I don't think it turns out well. It's just like going to a well-formed economy, collecting data of standard of living and then using that to form a policy in another country that the standard of living is low. (parent 5 interview transcript, Pos. 52)

Contrary to the initial dataset utilized for Ubenwa, both parents and data experts were pleased with the development of Ubenwa in Canada. Their reactions were more positive and welcoming, with figure 12 showing their openness and support to it. According to them, Canada is a technologically advanced country, with expertise and standard processes, which they all pointed out as an advantage. According to one of the parents and a data expert, Ubenwa being developed in Canada is an advantage for Nigerians, sharing the same thoughts with other participants; they stated:

> *I believe in Canada they have all the technology they need to really work on the app and give it an almost 100 if not 100 percent accuracy in diagnosis. (parent 5 interview transcript, Pos. 56)*

> *I will trust applications developed in Canada more than those developed in Nigeria. (Data expert 5 interview transcripts, Pos. 48)*

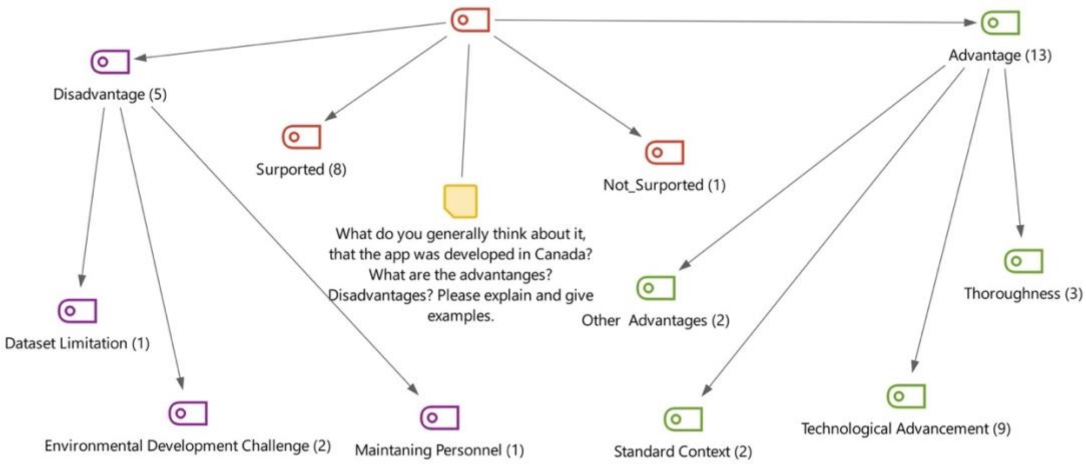


Figure 12: Development of Ubenwa in Canada

Data limitation still stood out as a barrier to participants for Ubenwa's successful use in Nigeria, despite its development in Canada. Though the Ubenwa solution encourages innovation to support the challenges currently facing the Nigerian healthcare industry, its development with non-local data was frowned upon by most participants. Hence introducing the Apgar method as an alternative to Ubenwa recorded a mixed reaction, as shown in figure 13 below, especially from the non-clinicians among the participants. The clinicians were quick to point out the Apgar method from the onset of the interview, advising for Ubenwa to adopt other indicators for detecting asphyxia as diagnosis based on only an infant's cry is not sufficient. Though it was noted by some parents that Ubenwa is fast for diagnosis compared to Apgar and would, therefore trust Ubenwa solely for diagnosis, others asked for the use of both systems together to complement each other, seeing it as an opportunity for Ubenwa to collect local data and improve on its software. In their opinion, using both systems together is an advantage, especially for Ubenwa. But some experts did not share the same views.

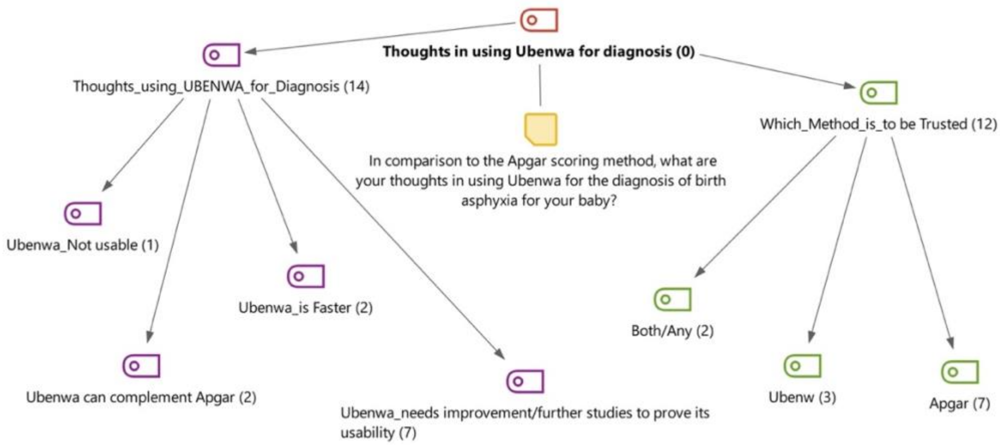


Figure 13: Comparing Ubenwa to the Apgar method

Most of the data experts advised for the Ubenwa app to be developed using the components of the Apgar score and disagreed with the use of Ubenwa and Apgar together, seeing it as overburdening the healthcare workers. In their words:

> *I feel for Ubenwa to be really effective, then other elements should be picked form Apgar score and implemented into it. (Data expert 2 interview transcript, Pos. 78)*

> *Using them at the same time will increase the time spent in diagnosing a baby. However, it will ensure that there are no mistakes in the diagnosis. (Data expert 5 interview transcripts, Pos. 73)*

> *that would now be double work for the health worker who will do the cry of the baby and then not capture the Apgar score. And that will worsen the healthcare situation in the country already because we do not have health workers and now you're asking them to do double work. (Data expert 1 interview transcript, Pos. 75)*

While some parents agreed to trust Ubenwa solely due to its fast diagnosis, all data experts except one who was indifferent, trust the Apgar method solely for diagnosing asphyxia. For them, Ubenwa's diagnosis is based on a machine which is unable to take into consideration other human factors in decision making, so they feel safe with the Apgar method; even though it takes a longer time to diagnose, and could be subjective, it utilizes 5 different indicators which makes them trust its use for diagnosis over Ubenwa. One of the data experts was curious to know the rationale behind the development of Ubenwa, was the deciding factor for its development inconsistencies detected from the use of Apgar? A parent was also curious to know what the app developers intended to gain in return for designing the solution. Prior to the questions around what the Ubenwa developers stand to gain, the parent had previously expressed concerns around the data collected by the Ubenwa app, and how it would be used. Her thoughts could be interlinked as the trophy for the app developers and the reason for Ubenwa's development. Could it be because the app developers are truly interested in saving the lives of infant children in Nigeria, or there were other ulterior motives? In their words, they asked:

It would be nice to know the justification or the rationale behind the ubenwa app, were there studies that showed there were inadequacies in the current Apgar system? (Data expert 4 interview transcript., Pos. 103)… It would be nice to look at what justifies this generally. (Data expert 4 interview transcript., Pos. 103)

my worry is, because these guys developed this app, are we not giving them a lot of information? Because you can't say that they won't have access to our information and we know data is King, so how much of our data are we giving out to these guys? They're in their comfort spaces and getting this trailer load of data that they can use for whatever. (Parent 2 interview transcript, Pos. 38)

So, I would also love to know the people behind the Ubenwa app and what the catch is, like I said why are they doing this? Is it because they want to advance out health system or because there's a reason? I would really genuinely want to know why they're doing it (Parent 2 interview transcript, Pos. 79)

But in view of utilizing local dataset to improve the Ubenwa algorithm, Figure 14 illustrates the diverse reactions of the participants. Most participants agreed on the use of local datasets to improve and encourage the use of Ubenwa. They pointed out acceptance/ ownership, diversity, improved diagnosis, elimination/ reduction of bias, amongst other reasons as an advantage of doing this.

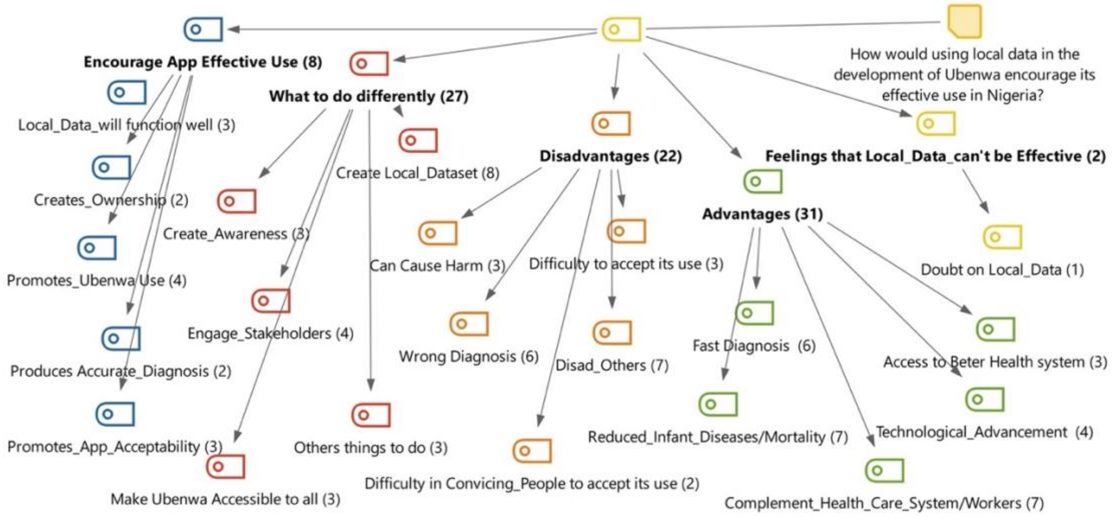


Figure 14: Local datasets to improve Ubenwa

Data expert 1 explained how AI systems works based on the data used in their design, pointing out the unethical use of Mexican data in the design of Ubenwa and his non-acceptance of the solution on his child based on that. In his opinion, local data would improve transparency and trust in the use of Ubenwa, and possibly eliminate bias and wrong diagnosis. In his words:

> *if you want to measure the reaction of a human and you use the barks of a dog, in as much as you're using data, if it's not the relevant data you will get different things; or you tag an AI picture of cats and books and stuff, and tag them as human, if anything happens, then it would be generating paper as a human; it would give you wrong result. So, using the right data gives you the right results. (Data expert 1 interview transcript, Pos. 62)*

Though using local data could be seen as an enhancer of the Ubenwa solution, the disadvantages of the same were equally pointed out by some participants. One of the data experts was quick to express his doubts on local data making any change in improving Ubenwa, alleging falsification of data in Nigeria as his reason and would not consider using Ubenwa if developed with Nigerian data. This supports his earlier statement on Nigeria not having quality data hence approving the use of data from another country for Ubenwa's development. In his exact words, he stated:

> *Nigeria does not have quality data for such an application system (Data expert 5 interview transcripts, Pos. 43) … I do not trust data from Nigeria because they are not*

accurate and full of errors. I will prefer data from other countries, because I trust the quality of the data. (Data expert 5 interview transcripts, Pos. 60)

In support to this data expert's allegations, other participants have made mention to falsified data in Nigeria at different points during their interviews – both parents and data experts. Some statements include:

Well if they want to depend on results based on Nigeria, I'm sorry because the results would be falsified. (Parent 3 interview transcript, Pos. 57) … When you talk about data in Nigeria, they don't work together. It's always falsified (Parent 3 interview transcript, Pos. 93) … Nigeria and data are not friendly. (Parent 3 interview transcript, Pos. 93)

this systems has been fraught with high level of data quality problems that people cannot rely on the data coming in because they do not believe in the data (Data expert 1 interview transcript, Pos. 91)… Personal data is highly political in Nigeria; (Data expert 1 interview transcript, Pos. 95)

And don't forget that in Nigeria, people want to be seen a certain way, so they'll always falsify data, people are always lying about data. (Parent 2 interview transcript, Pos. 69)

This new perspective brings on data falsification as a twist to localized data being the solution to improving the effective use of Ubenwa in Nigeria. In this new light, using Mexican data or Nigerian data for Ubenwa's development might not record any significant change in eliminating or reducing bias, discrimination or misdiagnosis by the Ubenwa solution in Nigeria. As earlier stated, data is not readily available in Nigeria, and when it is, the data source is usually questionable. In general, corrupt practices have significantly affected data practices in Nigeria; the need to hide or change information is an everyday practice. Data management practices in Nigeria is very poor, hence the reluctance of the participants to trust localized data.

So, how can Ubenwa be improved for local use in the face of data falsification? As proposed by one of the parents, Nigerian cry samples can be collected while Ubenwa is being used alongside the Apgar method. The accuracy of diagnosis can be determined during such a process, and the algorithm can be retrained to fit into the Nigerian context. In agreement, other participants supported the use of Ubenwa alongside the Apgar method to improve the data utilized in its design. In addition to the collection of Nigerian cry samples, participants made other suggestions for improving the Ubenwa solution; this is highlighted further in the discussion section.

6. Discussion

The research revealed the enthusiasm of interviewees, who were clearly elated at the development of Ubenwa for diagnosing birth asphyxia. Its fast diagnosing feature would support the reduction of infant mortality through prompt intervention, which all participants found really useful to this part of the world. For them, Ubenwa would help to promote accessible and affordable healthcare, bring technological advancement into and encourage innovative research in the health industry, while complementing understaffed and overworked medical facilities. The participants were fascinated at the thought of quickly carrying out diagnosis with their mobile devices and seek further medical attention based on the results received, and not have to depend on the overstretched healthcare workers. Ubenwa and other AI solutions were further described as 'healthcare on our fingertips'. The concerns on utilizing datasets from Mexico came to bear; though it would reduce development costs and time, fears bordered around causing harm through failed diagnosis, especially when the decision is based solely on the technology. The research questions are addressed through the research findings.

Research Question 1 - Based on the Ubenwa use case, what are the existing concerns of data subjects with the introduction of AI diagnostics tools for the development of healthcare systems in Nigeria?

The research findings flagged issues bordering around bias and discrimination, privacy, damage, or harm through failed/ wrong diagnosis, and the system not being robust enough in comparison to the Apgar method being used for detecting asphyxia in the country. As previously stated, Ubenwa was developed in Canada using Mexican data to design its algorithm, which is currently being pilot tested in Nigeria. From the participants' point of view, the implementation of datafied solutions developed with data from another social context is wrong; based on this, the authenticity of Ubenwa's result output was questioned, with some participants rejecting its use based on this. Bias and misdiagnosis were the major challenges highlighted with the use of Ubenwa owing to the differences in culture, lifestyle, physiology, and environment between a Nigerian and a Mexican. The use of wrong datasets in Ubenwa's design could cause damage to parents and babies in the event of misdiagnosis, with the possibility of babies suffering from permanent disabilities and even death. Hence, Ubenwa is at risk of not delivering on its intended purpose of reducing infant mortality in Nigeria; based on its design. It was concluded that the app would be more useful in Mexico than in Nigeria. A participant raised concerns around data collected by the Ubenwa app and its intended use, making mention of privacy and utilization of data for activities outside its original purpose.

From the interviews, it was deduced that socio-cultural interoperability plays a vital role in the success of algorithmic governance tools. The use of contextualized data would strengthen people's trust in the system, as with Ubenwa, where most interviewees were willing to accept its use only when the algorithm is improved on with local content. But this situation is quite peculiar as it was also deduced that Nigeria lacks credible localized data for the design of algorithmic tools like Ubenwa, based on interviewees alleging to ongoing data falsification practices in the country. Based on this, suggestions were made on improving the Ubenwa tool with the right data, which includes unanimously collecting local cry samples, both asphyxiated and non-asphyxiated, from different hospitals in the different regions across the country, over a specific period of time. To ensure Ubenwa's effectiveness, it was also suggested for a comparative analysis to be carried out, that is the comparison of results from the Ubenwa solution with other diagnosis methods like the Apgar score. This would ensure the machine functions better by understanding the diverse contexts of individuals' experiences. Asides local context, other suggestions gathered includes:

a. Awareness creation: Based on survey results, over 80% of respondents are unaware of any AI tool for healthcare management in Nigeria; interviewees were equally unaware except for one data expert. Behaviour change could play a crucial role in boosting the acceptance of datafied solutions like Ubenwa, hence the need to create awareness of these solutions.
b. Stakeholder involvement: Stakeholder involvement is crucial to the success of datafied solution. In the case of Ubenwa, the stakeholders would include medical and technological experts, academia, regulatory agencies, amongst others, to support the acceptance and robustness of the system through the perspectives and experiences of all involved.
c. Regulatory framework: Regulatory frameworks would address all concerns related to system design, data usage, and end user’s safety; hence it is important for Nigeria to develop an AI regulatory framework for solutions like Ubenwa.
d. Use of local language: The adoption of a common language, asides the official English language by Ubenwa or any other AI solution. For example, pidgin English, which is understood by the majority of the populace.

e. Robustness: Ubenwa's diagnosis is solely dependent on its learning algorithm, which could be biased; it is unable to apply human factors in its decision making. Therefore, adopting other features from the Apgar method to its design will reduce machine error as cry is not universal, as stated by many participants. The new features would consider physical assessment.

Further suggestions included creating a feedback system, enabling a more precise diagnostic result for better understanding, and enabling alternative power solutions to address poor infrastructure issues for its use, especially in rural communities. With these listed suggestions, the diversity of the Nigerian populace would be captured, further improving on the effectiveness of Ubenwa.

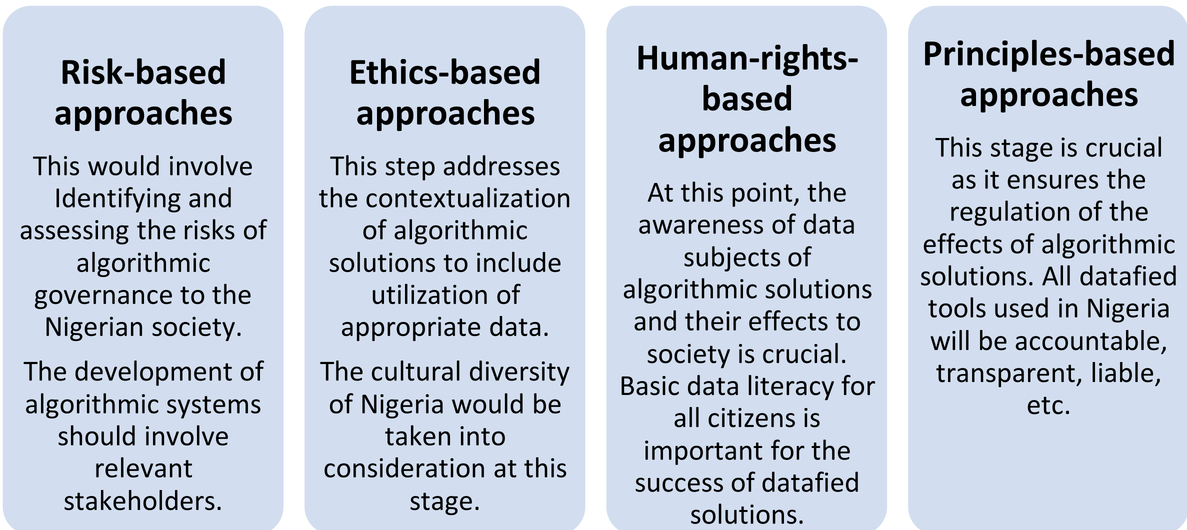
Research question 2 – How ready is Nigeria for algorithmic governance in its health sector?

One of the ways in determining Nigeria's readiness for algorithmic governance is the existence of regulations to manage the use of the same. Compliance with the same would encourage AI for social good. As mentioned in the theoretical framework, Nigeria, as well as other African countries, lack regulatory frameworks for AI systems. This view was supported by the interviews. At the moment, there is no identifiable regulatory body to assess Ubenwa and other algorithmic tools in the country; for the effective use of datafied solutions in the country, these regulations have to be developed and enforced to protect citizens from bias and social inequalities. Suggestions for assessing such tools were made by data experts; they include having standard regulations to govern the use of algorithmic solutions in the country, stakeholder involvement to determine opportunities and risks of algorithmic systems before deployment, monitoring and evaluating such systems, and the automation of public health systems to better interoperate with datafied tools. In addition, citizens should have the right to accept or reject the use of these solutions, likewise the collection of their personal information.

Nigerians are largely unaware of their rights as data subjects; the research findings support this view. Data malpractice experiences were shared during interview sessions; for instance, healthcare workers carelessly divulging or hoarding patients' personal information. In practice, access to patient's healthcare data is solely dependent on a health worker's discretion, totally contradicting available data protection regulations in the country, which gives sole ownership of personal data to individuals. These malpractices could pave the way for data colonialism and inequality. Previous research has shown that data colonialism is built on an existing system of capitalism; in today's digital society, access to personal data is the driver of capitalism. The Ubenwa solution intends to gather the largest annotated infant cry samples during its pilot testing in Nigeria, for a country where data regulations are not enforced for citizens protection, patient's privacy could be breached. Collected data samples could be exploited for other purposes than intended for profit-making, without the knowledge of data owners due to a large percentage of Nigerians being unaware of their personal data rights. Social inequalities could also arise due to their unawareness of algorithmic solutions and their effects on society.

There were no suggestions for improving the data governance systems in Nigeria. Participants do not believe the system could be improved. They only requested for data and AI awareness programs for the entire populace, to promote the success of Ubenwa and other algorithmic tools in the country. Awareness should address data protection, privacy, security, ethics and regulations, and should be carried out through training and sensitization programs, community outreaches, introduced into the school curriculum, amongst others.

The researcher is proposing the adoption of the Just and Latzer's approaches (Latzer and Just, 2020) to address existing concerns of algorithmic governance, raised from research findings. This proposed framework would be compared with research findings to advise on best practices for improving the use of algorithmic solutions like Ubenwa in Nigeria. Based on the interview findings, participants' perceptions aligned with the Latzer and Just approaches for addressing algorithmic governance concerns. Figure 4 above illustrating the approach has been modified/ revised below to better suit the concerns of the Nigerian context.



The first step for the success of algorithmic solutions in Nigeria is to identify and assess the risks of such tools. This is best handled by stakeholders familiar with the terrain of the industry concerned and the country. Their professional experiences would capture the publics' perspective and different social strata in designing such solutions, as Nigeria is culturally and socially diverse. Risks to look out for would include social discrimination, intellectual property violations, amongst others. Using the Ubenwa case study scenario, stakeholders would include doctors, midwives/ nurses, academia, professional bodies, and relevant government agencies. The first step sets the foundation for the other processes to follow. The ethics-based approach would follow suit in ensuring the traceability of algorithms for the elimination and reduction of bias and discrimination. Social interoperability comes to play here; algorithms would be designed and trained using representative data sets. With a focus on Ubenwa, this step would ensure the Ubenwa solution utilizes cry samples of Nigerian babies as against Mexican samples. This would promote acceptance, build trust, as well as eliminate bias and wrong diagnosis. With the initial step already in play, stakeholder engagement would support quality data collection across different hospitals in the different regions for inclusiveness and representation. With the first two steps in motion, issues of privacy and data protection would begin to arise. Hence the introduction of data awareness programs for citizens. Individuals should have basic knowledge of their rights as data subjects in order to address human rights concerns; they should be free to express themselves without being discriminated, understanding their rights to participate in the use of datafied solutions like Ubenwa. Lastly, the need for appropriate legal frameworks and policies for regulating algorithmic solutions cannot be overemphasized. The initial three steps – proper system design, utilization of appropriate data, and end user's safety, is enclosed in this final stage. This stage would ensure accountability, transparency, liability, amongst others. This stage supports Rob Kitchin’s call for proper guidelines to improve users' experience in utilizing algorithmic governance solutions (Kitchin, 2014).

At the moment, Nigeria lacks any AI regulation; the researcher advises that the country adopt the EU AI-HLEG guidelines for trustworthy AI until it is able to develop one that suits its needs. Developing an AI framework that is contextualized to the Nigerian tradition and history would support a just algorithmic governance process. This would form a solid foundation for the success of algorithmic governance in Nigeria. The guideline would address the need for fairness and prevention of harm by ensuring equal opportunity in the development, deployment, and use of an Algorithmic tool; and the need to promote an ethical mindset through education and awareness on basic AI literacy. This framework would also ensure the Nigerian authorities are accountable for data governance, thereby reducing data falsification in the country.

The researcher would also like to propose further research to study the perspectives of healthcare and data experts from Canada and Mexico. In the course of conducting the interviews with data experts and parents in Nigeria, many questions were left unanswered; hence the study would propose that these additional studies be extended to the Ubenwa app developers to find answers and gain clarity to the existing concerns raised by participants.

Limitations of the study

Conducting the face-to-face interviews with the covid19 pandemic was a challenge as precautionary measures had to be taken. Ensuring strict adherence to the social distancing and compulsory face masks guidelines imposed by the government throughout the interview sessions caused initial discomfort with participants. Conducting the interviews online would have been a safer option, but the erratic internet service in Nigeria limited this opportunity, as some participants had poor network connections in their homes. Time limitations were also experienced as some of the interviewees, especially involving the clinicians, had to be rescheduled oftentimes due to their involvement in the fight against the covid19 pandemic. This slowed the research process as results analysis had to be delayed until all interviews had been done.

CONCLUSION

In our current datafied society, algorithms play a major role in uncovering new insights from massive data sets. Both public and private sector actors are utilizing algorithmic decision-making tools for various purposes. It can automate simple and complex decision-making processes, but the general public poorly understands its processes. Algorithms are socially constructed and deployed, hence the possibility of bias that can originate from incomplete or unrepresentative training data. The decisions based on biased algorithms have grave impacts on society's segments, with platform users and designers usually unaware of these biases. This research focuses on algorithmic governance's social perspectives, that is, its exploitation and governance, and the suggested approaches for addressing identified concerns.

Algorithmic solutions mostly deployed in Africa are designed using non-representative data, largely because there is little or no available local data in the continent. Previous research has shown that data representativeness is crucial in designing an algorithm that does not bias and discriminate. However, for a continent that lacks local data, this could prove difficult. Its healthcare industry is currently understaffed and overstretched and could benefit from algorithmic solutions to improve efficiency. Digitally enabled systems are currently being deployed for healthcare, thereby increasing the risks of discriminatory or unfair diagnosis. The continent lacks strong regulatory guidelines to ensure the accountability and transparency of these algorithmic systems for social good; its populace is mostly unaware of the effects of algorithms on their society. Hence, this research is vital in helping to identify, mitigate, and remedy the end-user impacts of algorithmic solutions in Nigeria, and Africa at large. The study highlights the challenges and effects of algorithmic governance systems in Africa, using the Ubenwa AI solution deployed in Nigeria as a case study.

Ubenwa is an AI solution leveraging the cry of neonates to detect birth asphyxia. In-depth interviews and an online survey were conducted to capture the concerns of the proposed end-users of the Ubenwa solution. Nigerian data experts and parents were interviewed to understand their perceptions about datafied solutions like Ubenwa, while the online survey captured the general awareness and readiness to use datafied solutions for healthcare in Nigeria. Manual techniques, in combination with the MAXQDA software, were utilized in analyzing the qualitative data.

This study has shown that using data from a different context in other contexts is ethically wrong and could lead to bias and may reinforce social inequality against end-users of such solutions. The Ubenwa solution highlights a missing relationship between the technology and the people because its development is not inclusive of the target population group. Compared to the manual Apgar method currently used for birth asphyxia detection in Nigeria, the Ubenwa solution is not technically robust as it addresses only one out of the five indicators of the Apgar scoring method. Ubenwa, like every other automation system, is built to be smart, but it lacks the innate ability to apply any other human techniques in diagnosing asphyxia except through cry samples.

From the results, participants shared similar reactions to the Ubenwa solution creating efficiency in the healthcare industry in Nigeria; but they questioned the authenticity of and dependence on Ubenwa due to its originating dataset. Their preferred choice of diagnosis was the Apgar method because it does not depend on a biased diagnosis algorithm. This social experiment supports ongoing research on the importance of including the local populace's experiences in the design of algorithmic solutions. This study also found that social interoperability is crucial to the success of datafied solutions. The study's future work area will include designing a framework for awareness creation for algorithm developers, AI tools users, and the citizens to better strengthen the AI regulatory system.

LIST OF REFERENCES

Akinnagbe, A., Peiris, K. D. A. and Akinloye, O. (2018) ‘Prospects of Big Data Analytics in Africa Healthcare System’, *Global Journal of Health Science*, 10(6), p. 114. doi: 10.5539/gjhs.v10n6p114.

Aslam, H. M., Saleem, S. and et al (2014) ‘Risk factors of Birth Asphyxia’, *Italian journal of Paediatrics*, 14(94). doi: 10.1186/s13052-014-0094-2.

Borgesuis, F. Z. (2015) ‘Behavioural Sciences and the Regulation of Privacy on the Internet’, in *Nudge and the Law : A European Perspective*. Hart Publishing, p. 40. doi: 10.5040/9781474203463.ch-008.

Boyd, D. and Crawford, K. (2012) ‘Critical Questions for Big Data’, *Information, Communication & Society*, 15(5), pp. 662–679. doi: 10.1080/1369118X.2012.678878.

Braun, V. and Clarke, V. (2006) ‘Using thematic analysis in psychology’, *Qualitative Research in Psychology*, 3(2), pp. 77–101. doi: 10.1191/1478088706qp063oa.

Braun, V. and Clarke, V. (2013) ‘Teaching thematic analysis : Overcoming challenges and developing strategies for effective learning’, *The psychologist*, 26(2), pp. 120–123.

Brunon-Ernst, A. (2015) ‘The Fallacy of Informed Consent: Linguistic Markers of Assent and Contractual Design in Some E-User Agreements’, *Revista Alicantina de Estudios Ingleses*, (28). doi: 10.14198/raei.2015.28.03.

Buolamwini, J. and Gebru, T. (2018) ‘Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification’, in Friedler, S. A. and Wilson, C. (eds) *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. New York, New York, USA, pp. 77–91.

Carman, M. and Rosman, B. (2020) ‘Applying a principle of explicability to AI research in Africa: should we do it?’, *Ethics and Information Technology*. doi: 10.1007/s10676-020-09534-2.

Chang, R. M., Kauffman, R. J. and Kwon, Y. (2014) ‘Understanding the paradigm shift to computational social science in the presence of big data’, *Decision Support Systems*, 63, pp. 67–80. doi: 10.1016/j.dss.2013.08.008.

Children: reducing mortality (2018) *World Health Organisation*.

Coleman, D. (2019) ‘Digital Colonialism: The 21st Century Scramble for Africa through the Extraction and Control of User Data and the Limitations of Data Protection Laws’, *Michigan Journal of Race and Law*, 24, pp. 418–440.

Couldry, N. and Mejias, U. A. (2019a) ‘Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject’, *Television & New Media*, 20(4), pp. 336–349. doi: 10.1177/1527476418796632.

Couldry, N. and Mejias, U. A. (2019b) *The costs of connection: how data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

Daouda, C. (2012) 'GOING GLOBAL" IN GROWTH MARKETS — CHINESE INVESTMENTS IN TELECOMMUNICATIONS IN AFRICA', *Sun Scholar Research Repository*. Stellenbosch University Library and Information Services, pp. 1–4.

Das, T. K. and Mahapatra, D. K. (2012) 'Addressing the Influencing Factors for Interoperability in a Global Domain', *IOSR Journal of Computer Engineering*, 7(2), pp. 41–48. doi: 10.9790/0661-0724148.

Davidson, C. (2009) 'Transcription: Imperatives for Qualitative Research', *International Journal of Qualitative Methods*, 8(2), pp. 35–52. doi: 10.1177/160940690900800206.

Davies, J. *et al.* (2020) 'A formal, scalable approach to semantic interoperability', *Science of Computer Programming*, 192, p. 102426. doi: 10.1016/j.scico.2020.102426.

Deaton, A. S. and Tortora, R. (2015) 'People In Sub-Saharan Africa Rate Their Health And Health Care Among The Lowest In The World', *Health Affairs*, 34(3), pp. 519–527. doi: 10.1377/hlthaff.2014.0798.

Dencik, L. *et al.* (2019) 'The "golden view": data-driven governance in the scoring society', *Internet Policy Review*, 8(2). doi: 10.14763/2019.2.1413.

Dijck, J. Van (2017) 'In data we trust? The implications of datafication for social monitoring', *MATRIZes*, 11(1), p. 39. doi: 10.11606/issn.1982-8160.v11i1p39-59.

Elebeke, E. (2020) 'NITDA Promises to Align COVID-19 Data Collection Strategies with NDPR Guidelines', *Vanguard Ngr*, 29 March. Available at: https://www.vanguardngr.com/2020/03/nitda-promises-to-align-covid-19-data-collection-strategies-with-ndpr-guidelines/.

Eubanks, V. (2017) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. First. New York, New York, USA: St. Martin's Press.

EuropeanCommunities (2004) *EUROPEAN INTEROPERABILITY FRAMEWORK FOR PAN-EUROPEAN eGOVERNMENT SERVICES*. Belgium.

Flyverbom, M. and Murray, J. (2018) 'Datastructuring—Organizing and curating digital traces into action', *Big Data & Society*, 5(2), p. 205395171879911. doi: 10.1177/2053951718799114.

Freeman, K. (2016) 'Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis', *NORTH CAROLINA JOURNAL OF LAW & TECHNOLOGY*, 18(5), pp. 75–106.

George, G., Haas, M. R. and Pentland, A. (2014) 'Big Data and Management', *Academy of Management Journal*, 57(2), pp. 321–326. doi: 10.5465/amj.2014.4002.

Heffron, C. (2017) 'China builds "the world's most powerful" facial recognition system that can identify any one of its 1.4 billion citizens within three seconds', *The Daily Mail*, October.

Hill, R. K. (2016) 'What an Algorithm Is', *Philosophy & Technology*, 29(1), pp. 35–59. doi: 10.1007/s13347-014-0184-5.

Hood, C. M. *et al.* (2016) 'County Health Rankings', *American Journal of Preventive Medicine*, 50(2), pp. 129–135. doi: 10.1016/j.amepre.2015.08.024.

Isaak, J. and Hanna, M. J. (2018) 'User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection', *Computer*, 51(8), pp. 56–59. doi: 10.1109/MC.2018.3191268.

Ishwarappa and Anuradha, J. (2015) 'A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology', *Procedia Computer Science*, 48, pp. 319–324. doi: 10.1016/j.procs.2015.04.188.

Kaisler, S. *et al.* (2013) 'Big Data: Issues and Challenges Moving Forward', in *2013 46th Hawaii International Conference on System Sciences*. Wailea, Maui, HI, USA: IEEE, pp. 995–1004. doi: 10.1109/HICSS.2013.645.

Kelley, K. W., Feldman, S. S. and Gravely, S. D. (2016) 'Engaging and Sustaining Stakeholders: Towards Governance', in Dixon, B. (ed.) *Health Information Exchange - Navigating and Managing a Network of Health Information Systems*. 1st edn. Indianapolis: Elsevier, pp. 59–76. doi: 10.1016/B978-0-12-803135-3.00004-9.

Kitchin, R. (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. First, *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. First. Edited by R. Rojek. 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom: SAGE Publications Ltd. doi: 10.4135/9781473909472.

Kitchin, R. (2017) 'Thinking critically about and researching algorithms', *Information, Communication & Society*, 20(1), pp. 14–29. doi: 10.1080/1369118X.2016.1154087.

König, P. D. (2019) 'Dissecting the Algorithmic Leviathan: On the Socio-Political Anatomy of Algorithmic Governance', *Philosophy & Technology*. doi: 10.1007/s13347-019-00363-w.

Kouroubali, A. and Katehakis, D. G. (2019) 'The new European interoperability framework as a facilitator of digital transformation for citizen empowerment', *Journal of Biomedical Informatics*, 94, p. 103166. doi: 10.1016/j.jbi.2019.103166.

Kuckartz, U. and Rädiker, S. (2019) 'Introduction: Analyzing Qualitative Data with Software', in *Analyzing Qualitative Data with MAXQDA*. Cham: Springer International Publishing, pp. 1–11. doi: 10.1007/978-3-030-15671-8_1.

Latzer, M. and Just, N. (2020) 'Governance by and of Algorithms on the Internet: Impact and Consequences', in *Oxford Research Encyclopedia of Communication*. Oxford University Press. doi: 10.1093/acrefore/9780190228613.013.904.

Lepri, B. *et al.* (2017) 'The Tyranny of Data? The Bright and Dark Sides of Data-Driven Decision-Making for Social Good', in Cerquitelli, T., Quercia, D., and Pasquale, F. (eds) *Transparent Data Mining for Big and Small Data*. Springer International Publishing, pp. 3–24.

doi: 10.1007/978-3-319-54024-5_1.

Lesh, K. *et al.* (2007) ‘Medical Device Interoperability-Assessing the Environment’, in *2007 Joint Workshop on High Confidence Medical Devices, Software, and Systems and Medical Device Plug-and-Play Interoperability (HCMDSS-MDPnP 2007)*. Boston, MA: IEEE, pp. 3–12. doi: 10.1109/HCMDSS-MDPnP.2007.22.

Leurs, K. and Shepherds, T. (2017) ‘Datafication & Discrimination’, in Schafer, M. T. and Es, K. van (eds) *The Datafied Society: Studying Culture Through Data*. Amsterdam: Amsterdam University Press, pp. 211–233. doi: 10.5117/9789462981362.

Leuthner, S. R. and Das, U. (“Shonu”) G. (2004) ‘Low Apgar scores and the definition of birth asphyxia’, *Pediatric Clinics of North America*, 51(3), pp. 737–745. doi: 10.1016/j.pcl.2004.01.016.

Light, B., Burgess, J. and Duguay, S. (2018) ‘The walkthrough method: An approach to the study of apps’, *New Media & Society*, 20(3), pp. 881–900. doi: 10.1177/1461444816675438.

Liu, H.-W., Lin, C.-F. and Chen, Y.-J. (2019) ‘Beyond State v Loomis: artificial intelligence, government algorithmization and accountability’, *International Journal of Law and Information Technology*, 27(2), pp. 122–141. doi: 10.1093/ijlit/eaz001.
Madden, S. (2012) ‘From Databases to Big Data’, *IEEE Internet Computing*, 16(3), pp. 4–6. doi: 10.1109/MIC.2012.50.

Masso, A. and Kasapoglu, T. (2020) ‘Understanding power positions in a new digital landscape: perceptions of Syrian refugees and data experts on relocation algorithm’, *Information, Communication & Society*, 23(8), pp. 1203–1219. doi: 10.1080/1369118X.2020.1739731.

Mayer-Schönberger, V. and Cukier, K. (2013) *Big Data: A Revolution that Will Transform How We Live, Work and Think*, *Zhurnal Eksperimental’noi i Teoreticheskoi Fiziki*. Boston, Massachusetts: Houghton Mifflin Co., Boston.

Mejias, U. A. (2013) *Off the Network: Disrupting the Digital World*, *Electronic Mediations*. London: University of Minnesota Press. doi: 10.5749/minnesota/9780816678990.001.0001.

Mejias, U. A. and Couldry, N. (2019) ‘Datafication’, *Internet Policy Review*. doi: 10.14763/2019.4.1428.

Mullen, J. (2015) ‘Google rushes to fix software that tagged photo with racial slur’, *CNN*, July.
Najafabadi, M. M. *et al.* (2015) ‘Deep learning applications and challenges in big data analytics’, *Journal of Big Data*, 2(1), p. 1. doi: 10.1186/s40537-014-0007-7.

Nehls, K., Smith, B. D. and Schneider, H. A. (2015) ‘Video-Conferencing Interviews in Qualitative Research’, in Hai-Jew, S. (ed.) *Enhancing Qualitative and Mixed Methods Research with Technology*. IGI Global, pp. 140–157. doi: 10.4018/978-1-4666-6493-7.ch006.

Novak, K. and Bidwell, L. (2019) ‘The Future of Healthcare in Africa’, *American Society for Competitiveness*, 17(2), pp. 431–438.

OECD (2013) ‘Exploring the Economics of Personal Data: A Survey of Methodologies for

Measuring Monetary Value', *OECD Digital Economy Papers*, 220, p. 39. doi: 10.1787/5k486qtxldmq-en.

Onu, C. C. (2014) 'Harnessing infant cry for swift, cost-effective diagnosis of Perinatal Asphyxia in low-resource settings', in *2014 IEEE Canada International Humanitarian Technology Conference - (IHTC)*. IEEE, pp. 1–4. doi: 10.1109/IHTC.2014.7147559.

Onu, C. C. *et al.* (2017) 'Ubenwa: Cry-based Diagnosis of Birth Asphyxia', *Cornell Univeristy*.

Onu, C. C. *et al.* (2019) 'Neural Transfer Learning for Cry-based Diagnosis of Perinatal Asphyxia'.

Onu, C. C. (2019) *UBENWA - 2018 IN REVIEW*, *Ubenwa.ai*.

Onu, C. C. (2020) *UBENWA - HIGHLIGHTS OF 2019*, *Ubenwa.ai*.

Pandit, H. J. *et al.* (2018) 'Investigating Conditional Data Value Under GDPR', *Procedia Computer Science*, p. 5.

Philip Chen, C. L. and Zhang, C.-Y. (2014) 'Data-intensive applications, challenges, techniques and technologies: A survey on Big Data', *Information Sciences*, 275, pp. 314–347. doi: 10.1016/j.ins.2014.01.015.

Raji, I. D. *et al.* (2020) 'Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing', in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. New York, NY, USA: ACM, pp. 145–151. doi: 10.1145/3375627.3375820.

van Reisen, M. *et al.* (2020) 'FAIR Practices in Africa', *Data Intelligence*, 2(1–2), pp. 246–256. doi: 10.1162/dint_a_00047.

Resnyansky, L. (2019) 'Conceptual frameworks for social and cultural Big Data analytics: Answering the epistemological challenge', *Big Data & Society*, 6(1), p. 205395171882381. doi: 10.1177/2053951718823815.

Rubinstein, R. A. (2014) 'Humanitarian-Military Collaboration: Social and Cultural Aspects of Interoperability.', in Albro, R. and Ivey, B. (eds) *Cultural Awareness in the Military*. 1st edn. Palgrave Macmillan, pp. 57–72. doi: 10.1057/9781137409423.0008.

Sachikonye, L. (2019) 'Zimbabwe and China: An all-weather relationship for development and stability?', in Hartmann, C. and Noesselt, N. (eds) *China's New Role in African Politics: From Non-Intervention towards Stabilization*. First edit. London: Routledge. doi: 10.4324/9780429422393.

Soley, L. C. and Smith, A. L. (2008) *projective techniques for social science and business research*. Southshore Press.

Sorescu, A. (2017) 'Data-Driven Business Model Innovation', *Journal of Product Innovation Management*, 34(5), pp. 691–696. doi: 10.1111/jpim.12398.

Spector, J. (2008) 'Preventing those so-called stillbirths', *Bulletin of the World Health*

Organization, 86(4), pp. 315–316. doi: 10.2471/BLT.07.049924.

Suri, H. (2011) 'Purposeful Sampling in Qualitative Research Synthesis', *Qualitative Research Journal*, 11(2), pp. 63–75. doi: 10.3316/QRJ1102063.

Taylor, D. B. (2020) 'George Floyd Protests: A Timeline', *The New York Times*, June.

Taylor, L. (2017) 'What is data justice? The case for connecting digital rights and freedoms globally', *Big Data & Society*, 4(2), pp. 1–14. doi: 10.1177/2053951717736335.

Taylor, L. and Broeders, D. (2015) 'In the name of Development: Power, profit and the datafication of the global South', *Geoforum*. New York, New York, USA: ACM Press, 64(8), pp. 229–237. doi: 10.1016/j.geoforum.2015.07.002.

Thatcher, J., O'Sullivan, D. and Mahmoudi, D. (2016) 'Data colonialism through accumulation by dispossession: New metaphors for daily data', *Environment and Planning D: Society and Space*, 34(6), pp. 990–1006. doi: 10.1177/0263775816633195.

Urbinati, A. *et al.* (2019) 'Creating and capturing value from Big Data: A multiple-case study analysis of provider companies', *Technovation*, 84–85, pp. 21–36. doi: 10.1016/j.technovation.2018.07.004.

Venier, S. (2019) 'The Role of Facebook in the Persecution of the Rohingya Minority in Myanmar: Issues of Accountability Under International Law', *The Italian Yearbook of International Law Online*, 28(1), pp. 231–248. doi: 10.1163/22116133_02801014.

Vestoso, M. (2018) 'The GDPR beyond Privacy: Data-Driven Challenges for Social Scientists, Legislators and Policy-Makers', *Future Internet*, 10(7), p. 62. doi: 10.3390/fi10070062.

Wang, C. J., Ng, C. Y. and Brook, R. H. (2020) 'Response to COVID-19 in Taiwan', *JAMA*, 323(14), pp. 1341–1342. doi: 10.1001/jama.2020.3151.

Ward, K. (2018) 'Social networks, the 2016 US presidential election, and Kantian ethics: applying the categorical imperative to Cambridge Analytica's behavioral microtargeting', *Journal of Media Ethics*, 33(3), pp. 133–148. doi: 10.1080/23736992.2018.1477047.

WHO (2012) *Social Determinants of Health, World Health Organisation*.

Williams, Brooks and Shmargad (2018) 'How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications', *Journal of Information Policy*, 8, pp. 78–115. doi: 10.5325/jinfopoli.8.2018.0078.

Wimmer, M. A., Boneva, R. and di Giacomo, D. (2018) 'Interoperability governance: A Definition and Insights from Case Studies in Europe', in *Proceedings of the 19th Annual International Conference on Digital Government Research Governance in the Data Age - dgo '18*. New York, New York, USA: ACM Press, pp. 1–11. doi: 10.1145/3209281.3209306.

Wirén, M., Mäntymäki, M. and Najmul Islam, A. K. M. (2019) 'Big Data Value Chain: Making Sense of the Challenge', in Pappas, I. O. et al. (eds) *Digital Transformation for a Sustainable Society in the 21st Century*. Cham: Springer International Publishing (Lecture Notes in

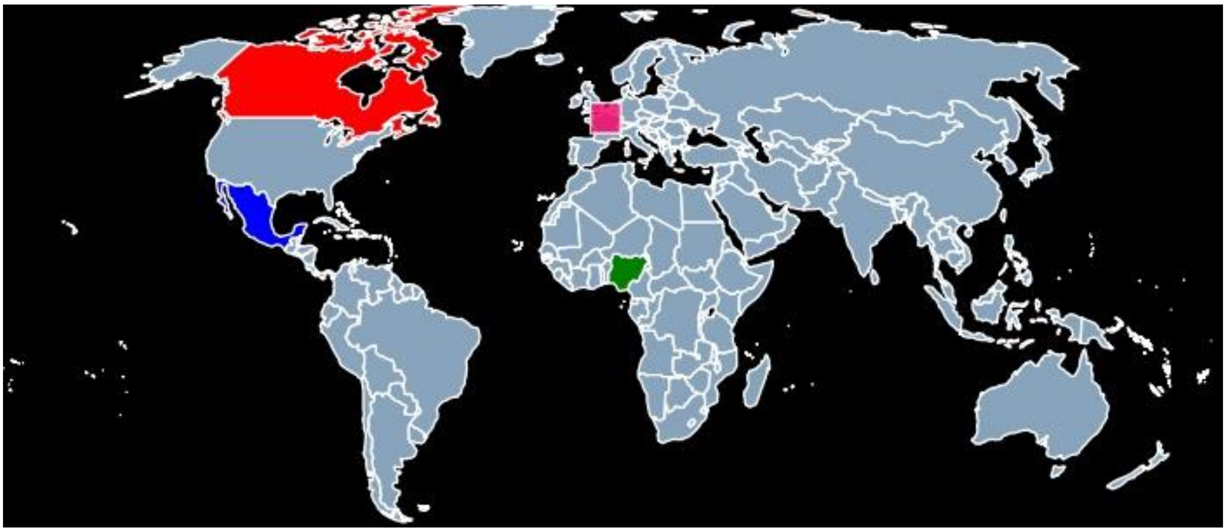
Computer Science), pp. 125–137. doi: 10.1007/978-3-030-29374-1_11.

Woolf, N. H. and Silver, C. (2017) *Qualitative Analysis Using MAXQDA*, *Qualitative Analysis Using MAXQDA*. Edited by N. H. Woolf and C. Silver. Abingdon, Oxon ; New York, NY : Routledge, 2017. | Series: Developing qualitative inquiry: Routledge. doi: 10.4324/9781315268569.

Yeung, K. (2018) 'Algorithmic regulation: A critical interrogation', *Regulation & Governance*, 12(4), pp. 505–523. doi: 10.1111/rego.12158.

APPENDICES

Appendix 1. Poster - Illustrative Map for Ubenwa



Canada is marked in Red, Mexico is marked in Blue and Nigeria is marked in Green.

Appendix 2. Poster – Ubenwa Diagnosis

Ubenwa - Cry-based diagnosis of birth asphyxia

10 seconds to diagnosis

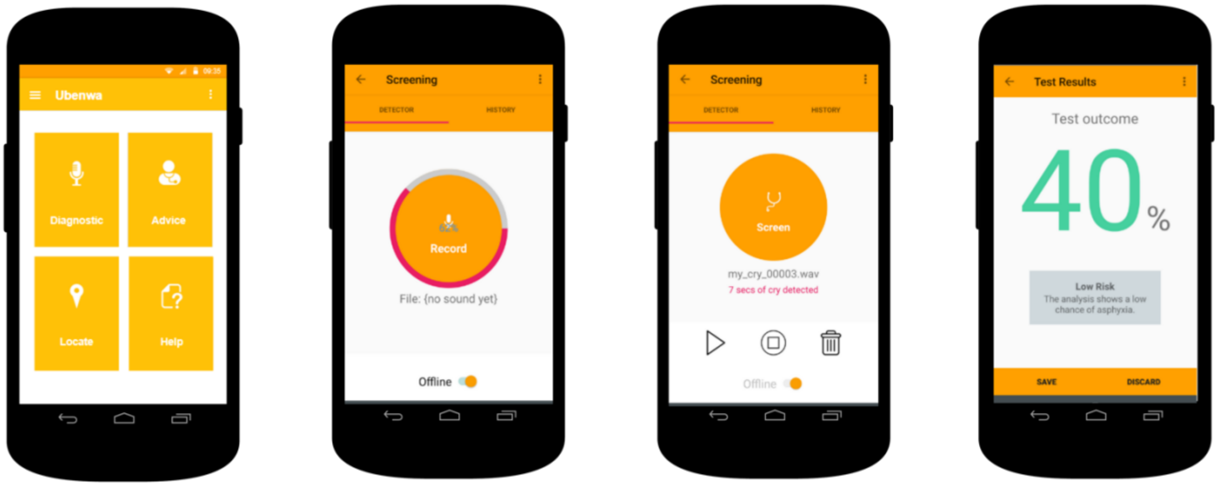
Non-invasive method - cry based rather than blood

95% cheaper than other alternatives

No expertise needed

Appendix 3. Poster – How Ubenwa works

How Ubenwa works



Appendix 4. Non-exclusive licence

A non-exclusive licence for reproduction and for granting public access to the graduation thesis[1]

I _________Martha Chukwu_______ (*author's name*)

1. Give Tallinn University of Technology a permission (non-exclusive licence) to use free of charge my creation

ALGORITHMIC GOVERNANCE IN NIGERIA'S HEALTH CARE INDUSTRY –

UBENWA AI SOLUTION

____,

(*title of the graduation thesis*)

supervised by________ Prof. Anu Masso ________________________________,

(*name of the supervisor*)

Co-supervised by________ Stefano Calzati ________________________________,

(*name of the co-supervisor*)

1.1. to reproduce with the purpose of keeping and publishing electronically, including for the purpose of supplementing the digital collection of TalTech library until the copyright expires;

1.2. to make available to the public through the web environment of Tallinn University of Technology, including through the digital collection of TalTech library until the copyright expires.

2. I am aware that the author will also retain the rights provided in Section 1.

3. I confirm that by granting the non-exclusive licence no infringement is committed to the third persons' intellectual property rights or to the rights arising from the personal data protection act and other legislation.

---

[1] *The non-exclusive licence is not valid during the access restriction period with the exception of the right of the university to reproduce the graduation thesis only for the purposes of preservation.*