A Sense-making Perspective of Algorithmic Accountability during Infodemics

BY

RAMARU ROFHIWA RODNEY

(11573950)

A DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF COMMERCE IN BUSINESS INFORMATION SYSTEM

IN THE

DEPARTMENT OF BUSINESS INFORMATION SYSTEMS

FACULTY OF MANAGEMENT, COMMERCE & LAW

SUPERVISOR: DR. WILLARD MUNYOKA

CO-SUPERVISOR: PROF. NIXON MUGANDA OCHARA

DECLARATION

I, Ramaru Rofhiwa Rodney, hereby declare that this dissertation titled - **SENSEMAKING PERSPECTIVE OF ALGORITHMIC ACCOUNTABILITY DURING INFODEMICS** - submitted to the University of Venda, has not been submitted previously for any degree at this or any other university. It is original in design and execution, and all reference materials contained therein have been duly acknowledged.

STUDENT Ramaru R.R ____________________________ **DATE** **09/03/2023**…………………….

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to those who have supported and contributed to the completion of my master's research project.

My sincere appreciation goes to my dedicated supervisors, Dr William Munyoka and Prof Muganda Ochara, and English editor Dr Kaburise whose guidance and expertise were instrumental throughout every stage of this research. Their insightful feedback, encouragement, and commitment to academic excellence have been invaluable.

I extend my thanks to the DEPARTMENT OF BUSINESS INFORMATION SYSTEMS FACULTY OF MANAGEMENT, COMMERCE & LAW for providing the necessary resources, research facilities, and a stimulating academic environment that fostered my intellectual growth.

Dedication

I dedicate this master's research work to the individuals who have been the cornerstone of my academic journey, providing unwavering support, encouragement, and inspiration.

To my siblings Tshinanne Cosmic Ramaru, Fhatuwani Sylvia Ramaru, and Livhuwani Ramaru, finally, my nephew Khumbudzo Brandan Ramaru, whose love and sacrifices have been my motivation. Your belief in my potential has been a guiding light, and I am grateful for your continuous support.

In conclusion, I dedicate this dissertation to my late parents, my father, Ailwei George Ramaru and my mother Naledzadzi Elizabeth Ramaru.

ABSTRACT

During the pandemic era, enormous amounts of data were available to decision-makers. Much of the data was generated using algorithms in which people did not want to be held accountable for any wrong/false information (misinformation) that was being provided to the public, either on internet websites, social media or television and radios. Algorithmic Accountability examines the process of assigning responsibility for harm when algorithmic decision-making results in discriminatory and inequitable outcomes. Mis-and disinformation about science, technology, and health is neither new nor unique to the COVID-19 era. Amid an unprecedented global health crisis, many journalists, policy makers, and academics have echoed what World Health Organisation has stressed that misinformation about the pandemic presents a serious risk to public health and public action. The biggest challenge is that this misinformation or fake news is drowning official public health advice on COVID-19, making it extremely problematic for the voices of healthcare professionals to be heard; the implications of this may be enormous as it may cause the virus and other pandemics to spread more rapidly within diverse populations. The purpose of this study was to establish the extent of Misinformation during Infodemics, to address the lack of Algorithmic Accountability. A mixed-methods approach was used as the topic required a purposeful mixing of methods in data collection, data analysis, and interpretation of the evidence. The key word is 'mixed'; this explains that an essential step in this approach is data linkage or integration at appropriate stages in the research process. The researcher used Spyder Python for extracting data from the Twitter API and then used Feedly which is a news aggregator application for various web browsers and mobile devices running iOS and Android; this procedure is also available as a cloud-based service. Python was used to compile news feeds from a variety of online sources for the researcher to customize and share with others; after filtering for the Covid-19 Tweets there were 21 of 508 records. The analysis suggests that misinformation about COVID-19 comes in many different forms, from many different sources, and makes many different claims. The process of misinformation often rearranges existing or accurate content relatively than formulated on a large scale, and where it is manipulated, it is edited with simple tools. Given the breadth of the pandemic, independent media and actions by platforms and others played a vital role in addressing virus-related misinformation. Government websites and the World Health Organization (WHO) can help sort false from true material, and accurate from misleading assertations. With running topic modelling for the data consisting of 21 of 508 tweets, topic modelling provided a way to compress the big data qualitative phase of the research.

Key words: Infodemics; Misinformation, Algorithmic Accountability; Sense-making perspective; Covid-19 Pandemic; Algorithmic decision-making systems; algorithmic systems, algorithmic transparency.

Abbreviations

AA – Algorithmic Accountability

AI – Artificial Intelligent

BDA – Big Data Analytics

BDA – Big Data Application

COVID - Coronavirus

EHR - Electronic Healthcare Records

NLP – Natural Language Processing

WHO – World Health Organization.

Contents

Chapter 1: INTRODUCTION

1. Introduction

The focus of this study was to investigate the sense-making perspective of algorithmic accountability during infodemics in South Africa. Detecting the spread of misinformation, such as, rumors, hoaxes, fake news, propaganda, spear phishing, and conspiracy theories, is an important task for Natural Language Processing (NLP) (Hossain, 2020).

The rationale for this study was to introduce the concept of sense-making from the perspective of algorithmic accountability in big data analytics (BDA) and to describe how algorithmic accountability may improve the interpretation of information. Chapter one covers - the background of the study, statement of research problem, research aim, questions, objectives, justification, operational definition of key terms and the entire research structure.

1.1 Background of the Study

Private and public entities around the world, particularly in the healthcare and governance sectors, are developing and deploying a range of artificial intelligence (AI) systems in emergency response to emergencies like COVID-19. Many of these systems are to trail and anticipate the spread of viruses to support medical responses and maintain social control. AI can reduce the pressure on devastated healthcare systems and curb the spread of the virus (Genevieve et al., 2020)

The algorithms which are driving the systems are man-made creations, and as a result, they are subjected to biases that can deepen societal inequalities and pose risks to business and society more broadly. Algorithmic accountability is the process of assigning responsibility for harm when algorithmic decision-making results in discriminatory and inequitable outcomes (Csir,2020).

The rapid global spread of COVID-19 was accompanied by what the WHO (2020) described as a massive infodemics (Fleming, 2020). According to WHO, the COVID-19 outbreak, and the response has been accompanied by a massive infodemics; an overabundance of information –

some accurate and some not – that makes it hard for people to find trustworthy sources and reliable guidance when they need it. WHO (2020) describes infodemics as a large increase in the volume of information associated with a specific topic whose growth can occur exponentially in a short period due to a specific incident, such as the COVID-19 pandemic (Saúde,2020).

Currently, there are few consumers' or civil rights protections that limit the types of data used to build data profiles or that require the auditing of algorithmic decision-making. Even though algorithmic systems can make decisions based on protected attributes like race, income, or gender, even when those attributes are not referenced explicitly, because there are many effective proxies for the same information. This study focused on the sense-making perspective on algorithmic accountability during infodemics on the pandemic regarding the misinformation and the analyzing of data and suggest ways in which public and private entities in the healthcare and governance sectors can help ensure that they develop, manage, and use algorithms equitably and responsibly.

Algorithmic accountability examines the process of assigning responsibility for harm when algorithmic decision-making results in discriminatory and inequitable outcomes. Africa was besieged with so many misconceptions and/or misinformation about COVID-19 via the media, astute politicians, social commentators as well as social media (Frenkel et al., 2020; Russonello, 2020).

Mukherjee et al., (2016) defines 'BDA' as a strategy used to analyze colossal information sets containing assorted qualities of information. Nowadays every organization wants to concentrate on extracting key values and information from huge datasets to achieve objectives of organization. BDA comprises of tools, algorithms and architecture that analyze and transform large and massive volumes of data (Michael et al., 2012). Applications of BDA can - improve public and private services in detecting the spreading of diseases earlier, generate new insights into disease mechanisms, monitor the quality of the medical and healthcare institutions as well as provide better treatment methods (Agarwal et al., 2015).

1.2 Statement of the Research Problem

The huge demand for information on pandemics, their toll on health-care systems, and the many unanswered questions about them created the perfect breeding ground for myths, fake news, and conspiracy. Some could be dismissed as ludicrous and largely harmless, but others were life-threatening (Fleming, 2020). There was too much misinformation that people received daily. In a 24-hour period, people could receive different information for one thing which could lead to misinformation then to confusion. Mis- and disinformation about science, technology, and health was neither new nor unique to COVID-19. Amid any unprecedented global health crisis, many journalists, policymakers, and academics echoed what WHO had stressed, that misinformation about pandemics presented a serious risk to public health and public action (Brennen et al., 2020). People were desperate for information on how the coronavirus evolved. Where did it come from? Would there be a cure or a vaccine? How could we keep staying safe? Was it here to stay forever or would it pass? Would life get back to normal again? There was a lot of information as well as misinformation about COVID-19 which made it challenging for ordinary people to know the truth. In the era of infodemics, the challenges posed by misinformation are complex and multifaceted, impacting various aspects of society, communication, and technology. Misinformation can spread rapidly and be amplified through social media platforms, online forums, and other digital channels. The viral nature of misinformation can lead to its widespread dissemination, making it challenging to control or counteract (Vosoughi.,2018). Misinformation often exploits gaps in trust and credibility, particularly in institutions, authorities, and traditional media sources. When individuals lack confidence in official sources, they may turn to alternative sources of information, including unreliable or misleading sources (Edelman., 2021). The biggest challenge was that misinformation or fake news was drowning official public health advice on COVID-19, making it extremely problematic for the voices of healthcare professionals to be heard; the implications would have been enormous as the virus was spreading between and within diverse populations (Ahinkorah et al., 2020). Many individuals lack the necessary skills and knowledge to critically evaluate information sources, assess credibility, and identify misinformation. Improving media literacy and digital literacy education is essential for empowering individuals to navigate the information landscape effectively (Hobbs.,2019). Algorithms used by social media platforms and other online services can unintentionally amplify misinformation by promoting sensational or controversial content. These algorithms may prioritize engagement metrics, such as likes, shares, and comments, over the accuracy or credibility of the information. In this pandemic, there was a lot of information that was not being published or given to the public (Guess.,2019). In algorithms, there was transparency of data, which was useful to consider if people were mindful of its bonds and limitations. The objective of any transparency policy was to clearly disclose information related to a consequence or decision made by the public; whether it was buying a product, making votes, selecting people if they qualified or not, or using a particular algorithm, people made more informed decisions. (Diakopoulos, 2014).

1.3 Research Aim

The aim of this study was to establish the extent of misinformation during infodemics to address the lack of algorithmic accountability.

1.4. Research questions

1.4.1. What determines the Influence of Algorithmic Accountability on Infodemics?

1.4.2. How does Algorithmic Accountability problematize transparency?

1.4.3. How can Algorithmic Accountability be realized in practice?

1.4.4. How can coronavirus and other pandemics' misinformation best be confronted?

1.5. Research Objectives

1.5.1 To Determine the Influence of Algorithmic Accountability on Infodemics

1.5.2. To examine if Algorithmic Accountability can problematize transparency.

1.5.3. To examine if Algorithmic accountability can be realized in practice.

1.5.4. To propose strategies on how COVID-19 and other pandemics' misinformation can be confronted.

1.6. Justification of the research

The COVID-19 infodemics stimulated a rapid global response that was in the early stages of an overall process of emergency management. As the situation unravels and new strategies of working were tried, lessons were learnt. The initial take-away was that infodemics was unprecedented in its magnitude and pace; that swift forms of misinformation were surfacing daily and that there was no global consensus on how to classify the types of false information being encountered.

Moreover, there were likely bigger issues surfacing because much of the misinformation being spread had not been unreliable health advice but was state-disseminated disinformation designed for destabilizing political scenes. With regards to infodemics, governments, academics, digital platforms, fact checkers, concerned institutions and the public, crucially demanded a shared knowledge of what was meant by the terms 'disinformation' and 'misinformation'. Without a scientific understanding of these concepts, there was a danger of significant socialist consequences from global health crises, like COVID 19.

Algorithms operated within the contexts of specific datasets, interfaces, situations, business models, and cultural expectations. By focusing uniquely on algorithms as the source of both pros and cons in a data-driven world, it was possible that algorithms might have been biased. The information-systems research community could have vigorously worked to help avoid this. Determining algorithmic accountability had real consequences for understanding and regulating who or which entities-controlled flows of information in public and private entities. What would governing, policing, or even designing ethical algorithms have looked like? The social, cultural, and political impact of an algorithmic solution had consequences that played out far beyond the technical innovation behind the restructuring of information. Algorithms broke down information into certain constituent parts and reconfigured it into a new production of information to achieve specific objectives, which had huge societal benefits. Hence, there was a need to develop quantitative science of such infodemiology to guide any adaptation of policies from epidemiology (Johnson et al., 2020).

1.7. Operational Definitions

1.8.1 INFODEMICS: Infodemics is defined as an overabundance of information – some accurate and some not – that makes it hard for people to find trustworthy sources and reliable guidance when they need it.

1.8.2. MISINFORMATION: Information or a story usually propagated through social media which is ultimately established as false or inaccurate. This definition covers different kinds of false information classified by their intention or motivation, such as profit, political intervention, deliberate deception, and unintentional misrepresentation (Brasoveanu et al., 2020).

1.8.3 ALGORITHMIC ACCOUNTABILITY: This refers to the assignment of responsibility for how an algorithm is created and its impact on society; if harm occurs, accountable systems include a mechanism for redress (Csir, 2020).

1.9 Structure of the Thesis

Chapter One is an introduction that describes all the key elements of the study, such as, the background, problem statement, research questions, research objectives, research design, and delimitation. The chapter also demonstrates how the research questions drawn from the research problem were to be answered by carefully following the research objectives.

Chapter Two discusses algorithmic accountability with reference to the coronavirus pandemic, the role of BDA and its impact on the algorithm systems. Based on the factors identified from the literature review, a theoretical framework which underpinned the study is presented. The framework is crucial for data collection as it clearly indicates what type of data is to be collected and from whom.

Chapter Three presents the research methodology exploited in this study. In this chapter, is outlined the steps and procedures on how data was gathered, hence, it gives the methodological approach to the research.

Chapter Four of the study presents the results obtained from the analysis of the collected data. This is done by generating the descriptive analysis, correlation, and regression analysis together with the Chi-Square test analysis for testing the hypotheses developed.

Chapter Five presents the discussion of the results, by narrating the significance of the results in relation to the set hypotheses. This chapter also makes a descriptive comparison of what has been obtained from this research in relation to the literature and actual practice. The findings set grounds for the research's recommendations and direction for future studies.

1.10 Summary

This chapter gave a brief introduction to this study and provides reasons to justify the study being conducted. The benefits of this study were highlighted in this chapter to give the reader a clear view of what to expect. This chapter also gave a clear explanation of what to expect from each chapter. The next chapter is a literature review of the study.

Chapter 2: Literature Review

This chapter reviews the literature on previously conducted studies based on the sense-making perspective of algorithmic accountability during infodemics. The main aim of the present study, therefore, was to determine the influence of the algorithmic systems on infodemics in order to build a framework that will provide a sense-making perspective, focusing on the misinformation during the coronavirus pandemic. The literature included professional journals, websites, and government documents.

2.1. Algorithmic Accountability

A systematic review of empirical studies conducted in 'algorithmic accountability' has so far been lacking (Wieringa, 2020). Complex algorithms are increasingly responsible for choices, operations, and decisions previously left to human actors. Fleming (2020) explores algorithmic accountability issues or assigning responsibility for harm when algorithmic decision-making results in discriminatory and inequitable outcomes.

Currently, few consumers' or civil rights' protections limit the types of data used to build data profiles or that require the auditing of algorithmic decision-making, even though the systems can make decisions based on protected attributes like, race, income, or gender, even when those attributes are not referenced explicitly, because there are many effective proxies for the same information.
Accountability is fundamentally about checks and balances to power. In theory, governments and corporations are held accountable through social, economic, and political mechanisms; journalism and public advocacy are additional tools to hold powerful institutions and individuals accountable. In a world of data and algorithms, however, accountability is often murky. Beyond questions about whether the market is sufficient or governmental regulations are necessary (Veeder et al., 2017), algorithms can be hugely beneficial in sorting through vast troves of information to determine the most useful sort. Automated algorithms can use well-defined steps and instructions to generate categories for filtering information based on a combination of either motives or the desirable outcomes.

In the final combination, the elements of - uncertainty, subjective interpretation, arbitrary choice, accidents, and other ingredients in the mix - are rendered invisible, and what is displayed to the end-user who interfaces with the algorithm's product, is just the functionality of the technology. For instance, Google, Yahoo, and other search engines can effectively create "filter bubbles" for the results people see when they query items, which can be problematic (Diakopoulos, 2015). Some information is more visible to one individual versus another, based on the user profiles that the search engine has of them, and how its algorithms predict what might be most relevant to the users according to their profiles.

Questions that can be asked at this junction are: *How might algorithms affect the flow of coronavirus data or other types of information? Who or which networks of stakeholders are the arbiters of algorithmic power that strongly influence information flows?*

The ever-increasing application of algorithms to decision-making in a range of social contexts has prompted demands for algorithmic accountability. Accountable decision-makers must provide their consumers with justifications for their automated system's outputs; but what kinds of broader principles should we expect such justifications to appeal to? Drawing from political philosophy, I present an account of algorithmic accountability in terms of the democratic ideal of 'public reason'. I argue that situating demands for algorithmic accountability within this justificatory framework enables us to better articulate their purpose and assess the adequacy of efforts toward them.

Algorithmic technologies are by no means objective, and examples littering the news, show that algorithmic technologies are not value-free; they can embed and perpetuate biases. Not knowing how inputs lead to outputs is becoming an urgent concern — claiming that "it was the algorithm" is unlikely to be accepted as a response to growing calls for accountability when there are concerns (Raymond, 2017: 215).

Various local governments and media workers should use public-health experts, especially, those from the Centers for Disease Control, to precisely provide relevant and accurate information to avoid fear among the public during pandemics like COVID 19 (Mian et al., 2020). As the use of intelligent algorithms and analytics is becoming more involved in how decisions are made in private and public life, the societal values of fairness, accountability, and transparency as well as the multidimensional value of human well-being are being discussed in the context of addressing potential negative and positive impacts of AI (Ashofteh et al., 2020).

2.2. Algorithmic accountability applied in operation.

In designing and developing algorithmic systems, the private sector plays a vital role because it owns large datasets (Mahdi, et al., 2021). Models for understanding and holding systems accountable have long rested upon ideals and logic of transparency. Seeing a system is sometimes equated with knowing how it works and being able to govern it; this is a pattern that recurs in recent work about transparency and computational systems (Ananny, et al., 2018). Table 2.1 shows the comparison of algorithmic fair data practices and unfair data practices for socio-legal fairness against unfair data practices.

	Legally fair data practices	Legally unfair data practices
Algorithmically fair data practices	Individuals fully understand and accept what data about them is being collected and how it will be used; the collection and usage occur in a way ensuring that data is not biased (or de-biased) and trained algorithms are employed in a way ensuring algorithmic fairness.	Individuals do not understand or do not accept what data about them is being collected and how it will be used, yet the collection and usage occur in a way ensuring that data is not biased (or de-biased) and trained algorithms deployed in a way

		ensuring algorithmic unfairness
Algorithmically unfair data practices	Individuals fully understand and accept what data about them is being collected and how it will be used; the collection and usage occur in a way that replicates or creates bias, and trained algorithms are used in an unfair, discriminatory manner.	Individuals do not understand or do not accept what data about them is being collected and how it will be used; on top of that the collection and usage occurs in a way that replicate or create bias, and trained algorithms are used in an unfair, discriminatory manner

Table 2.1: Algorithmic and socio-legal fairness and unfairness data practices (Pałka.,2018).

As the data used to train consumer-facing algorithms stems from consumer activity online, to ensure the overall fairness and transparency of these systems, we need to guarantee socio-legal fairness on top of the algorithmic one. These two, although in practice often occur jointly, in theory, are independent of one another; machine learning can be deployed to combat socio-legal unfairness (Lippi, et al.,2020).

2.3. Artificial Intelligence Algorithms Accountability

Research on algorithms and their impact proliferates, and so does calls for accountability of algorithms. A systematic review of the work that has been done in the field of 'algorithmic accountability' has so far been lacking (Wieringa.,2020). Algorithms also govern the operation of our internet search engines; they retrieve and prioritize search information and ultimately, therefore, what information is available to users and social media. High-frequency trading algorithms run our electronic markets; algorithms inform automated individual credit decisions and key health decisions and govern semi-autonomous driving (Pasquale, 2011). AI algorithms govern in subtle yet fundamental ways the way we live and are transforming our societies. The promise of efficient, low-cost, or 'neutral' solutions harnessing the potential of big data has led public bodies to adopt algorithmic systems in the provision of public services (Busuioc, 2021). Defining algorithmic systems, as noted above, algorithms lily instructions feed to a computer (Fuller., 2008); these are technical constructs that are simultaneously deeply social and cultural (Nick, 2017). Appreciating this 'entanglement' of various perspectives and enactments of algorithms, sees algorithms not as solely technical objects, but rather as socio-technical systems, which are embedded in culture(s) and can be viewed, used, and approached from different perspectives such as legal, technological, cultural, and social. This rich set of algorithmic multiples can enhance accountability rather than limit it.

The interdisciplinary systematic literature review presented in the remainder of these discussions, bundles knowledge and insight from a broad range of disciplines and appreciates the entanglements and multiples that are invariably a characteristic of algorithmic system interaction (Wieringa, 2020). As algorithmic systems have become increasingly ubiquitous in the public sector, they raise important concerns about meaningful oversight and accountability (Bullock, 2019). AI algorithms govern in more subtle yet fundamental ways how to live and are transforming societies. Tremendous technological advances brought on by data, computation, and the growing power of machine pattern recognition - relying on a range of methods referred to as 'deep learning' or 'neural networks'- have led to the ubiquity of AI algorithms in structuring not only technological but also human interactions (Wieringa, 2020).

2.4. Algorithmic fairness

Algorithms are often deployed to correct a source of bias in decisions made by humans, however, many algorithmic systems either codify existing sources of bias or introduce new ones. Additionally, bias can exist in multiple places within one algorithm (Csir, 2020). The figure below represents the combating of fake news during the COVID-19 pandemic.

2.5. Transparency

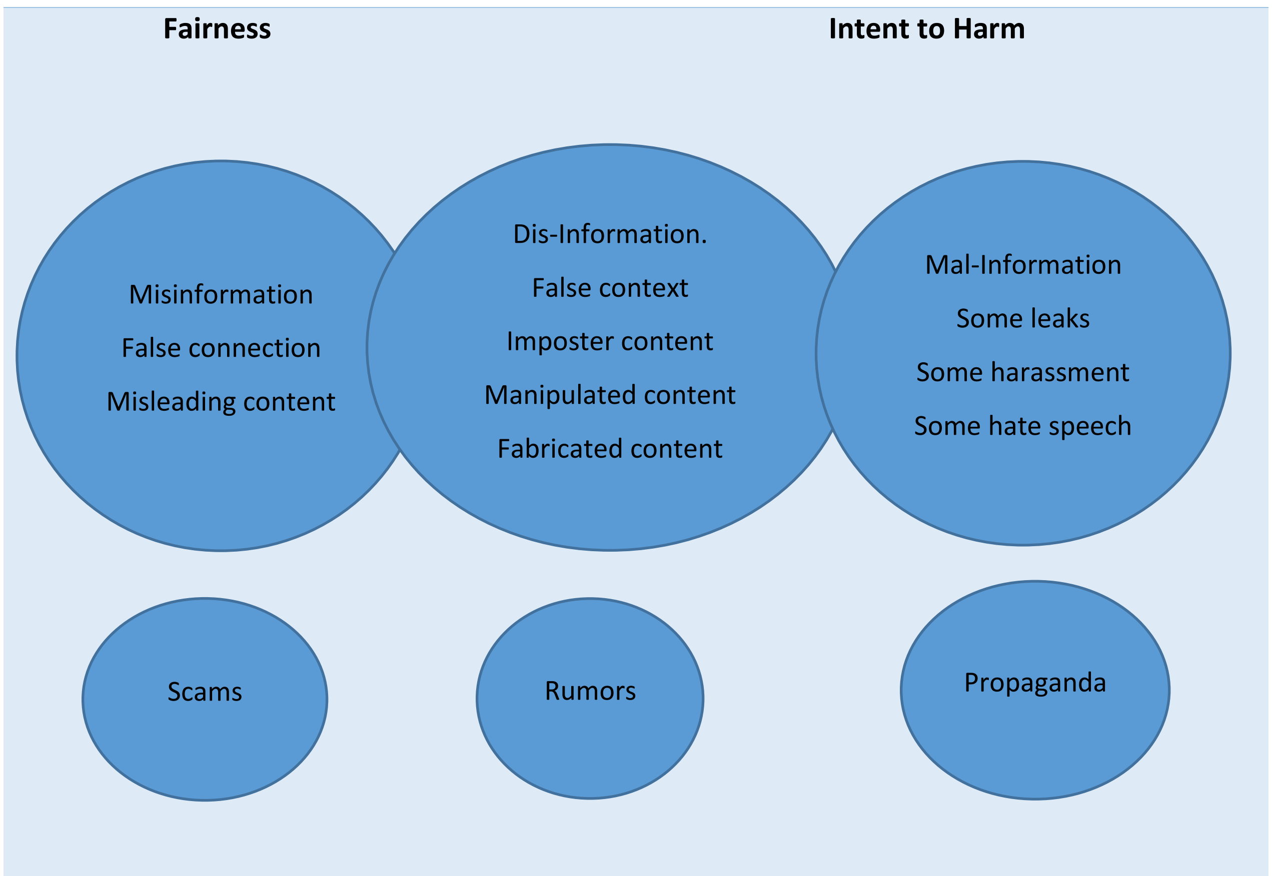


Figure 2.1. Transparency of the algorithmic fairness (Combating fake news during the COVID-19 pandemic presentation outline, CSIR, 2020)

According to Eprs, et al., (2019) depending on the type and use of an algorithmic decision system, the desire for algorithmic transparency may refer to one, or more of the following aspects - code, logic, model, goals decision variables, or some other aspect - that are considered to provide insight into the way the algorithm performs. Algorithmic system transparency can be global, seeking insight into the system behavior for any kind of input, or local, seeking to explain a specific input-output relationship.

Transparency implies the most basic conception of accountability, if individuals cannot know what an organization is doing, we cannot hold it accountable, and cannot regulate it; the demand for transparency in algorithmic systems goes beyond the simple and assumed sense. Transparent algorithmic systems have become so central to our lives and economies, and yet some of them use models and algorithms the workings of which are extremely complex for the human mind to follow. While the 'black box' metaphor by Pasquale, et al (2015) is clearly indicative of this sense of impenetrable mystery of the systems to us, it is important to consider making the box transparent, so that we can see the ‘gears’ within; this is truly what is needed to satisfy our concerns with these systems. Irrespective of which aspect of an algorithmic system is in question, that is usually not what the calls for transparency really aim at (Ziewitz et al., 2016).

2.5.1. Data- the transparency of the data used by the algorithmic system - in particular by machine learning and deep learning algorithms refers to the raw data, to the data’s sources, to how the data were preprocessed, to the methods by which it was verified as unbiased and representative. This includes looking at the features that are proxies for information about protected classes or to the processes by which the data are updated and the system is retrained on them. The transparency of the data used by algorithmic systems during infodemics refers to the extent to which the sources, quality, biases, and processing methods of the data are disclosed and accessible to stakeholders and users, particularly in the context of managing and mitigating misinformation (Taddeo et al.,2018).

2.5.2 Algorithms

The transparency of the systems' algorithms can refer to testing its output against inputs for which the proper output is known; reducing the variables to the most significant so we can validate them, and testing the system with counterfactuals to see if prejudicial data is infecting the output. The Algorithm transparency in algorithm accountability during infodemics is the extent to which the inner workings, decision-making processes, and outcomes of algorithms are visible, understandable, and accessible to stakeholders, particularly in the context of managing and combating misinformation during infodemics (Diakopoulos ,2016). A third-party code review provides an analysis of how the algorithms work, an inspection of internal and external bug reports, or an assurance that the software development processes are sound.

2.5.3 Goals - Algorithmic systems can also be transparent about their goals. When a system has multiple goals, this would mean being transparent about its relative priorities. Algorithmic systems can indeed enhance transparency by clearly articulating their goals and objectives. Such transparency is crucial for fostering trust, accountability, and ethical use of algorithms, particularly in contexts like managing misinformation during infodemics, Algorithmic systems should provide clear and unambiguous statements outlining their goals and intended outcomes. By explicitly defining their objectives, these systems offer transparency regarding the purpose and mission that guide their operations (Diakopoulos, 2016). This clarity ensures that stakeholders understand the intended focus and scope of the algorithmic system. Transparency in goal setting involves ensuring that the objectives of algorithmic systems align with the expectations and interests of various stakeholders, including users, policymakers, and the society (Floridi et al., 2018).

2.5.4. Outcomes - Manufacturers or operators could be required to be transparent about the outcomes of the deployment of their algorithmic systems. Requiring manufacturers or operators to be transparent about the outcomes of deploying their algorithmic systems is essential for accountability, consumer protection, and ethical oversight. Transparent reporting of outcomes empowers clients to make informed decisions about the products or services they engage with. When manufacturers or operators disclose information about how their algorithmic systems perform in real-world contexts, consumers can assess the risks, benefits, and potential impacts on their lives (Powles et al., 2018). Transparency allows clients to exercise agency and hold manufacturers accountable for the consequences of their algorithms.

2.5.5. Compliance - Manufacturers or operators may be required to be transparent about their overall compliance with whatever transparency requirements have been imposed upon them. Transparent reporting of overall compliance ensures that manufacturers or operators adhere to the transparency requirements mandated by regulatory authorities or industry standards. By providing detailed accounts of their adherence to transparency guidelines, manufacturers demonstrate their commitment to regulatory compliance and accountability (Wachter et al., 2017). Transparency allows regulators to monitor and enforce compliance effectively, thereby upholding the integrity of algorithmic systems. In many instances, authorities may insist that these compliance reports are backed by data that is inspectable by regulators or the public. Transparency in compliance reporting encourages manufacturers or operators to continuously evaluate and improve their transparency measures over time. By soliciting feedback, conducting audits, and implementing best practices, manufacturers can enhance the effectiveness and robustness of their compliance efforts (Floridi et al., 2018). The iterative process of improvement promotes a culture of accountability and innovation within the organization.

2.5.6. Influence - Transparency plays a crucial role in ensuring algorithmic accountability during infodemics, influencing various aspects of algorithmic systems design, implementation, and operation. Just as the public has an interest in knowing if an article in a newspaper was paid for by an interested party, the public may have an interest in knowing if any element of the AI process was purposefully structured to favor a particular outcome; this may be, for example, if a trusted search platform is artificially boosting some results because it was paid to, and if it is not flagging that fact to users, users, are then being manipulated. Regulators might want to insist that such influence be conspicuously acknowledged. Transparency in data collection and processing practices promotes accountability by disclosing the sources, types, and use of data within algorithmic systems. By providing visibility into data acquisition, storage, and processing methods, operators can demonstrate accountability in ensuring data quality, relevance, and fairness (Mittelstadt et al., 2016). Transparent data practices mitigate the risk of bias, misinformation propagation, and privacy violations during infodemics, thereby enhancing algorithmic accountability.

2.5.7. Usage - Users may want to know what personal data a system is using, either to personalize outcomes or as data that can train the system to refine or update it. Knowing what personal data is being used may then enable them to control that usage, perhaps to make their personalized results more accurate or, more urgently, because they feel that the usage violates their privacy. This may be the case even though the data in question may already be a desired part of the system, such as a purchase or search history. There are grey areas here as well, such as collecting anonymized, highly detailed information about trips made by autonomous vehicles and how often the car breaks or swerves, as for example, it could be important to optimize traffic for safety or fuel efficiency. Regulators may face some difficult decisions as well as drawing relatively obvious lines.

2.6. Data Mining in social media

Data mining is a process of discovering useful or actionable knowledge in large-scale data. The process also means knowledge discovery from data (KDD), which describes the typical activity of extracting useful information from raw data (Han, et al., 2011). Data mining algorithms are classified into - supervised, unsupervised, and semi-supervised - learning algorithms. Classification is a common example of a supervised learning approach; this involves a given data set being typically divided into training and testing data sets, with known class labels (Barbier, 2011). Supervised algorithms build classification models from the training data and use the learned models for prediction (Barbier, 2011). Supervised algorithms build classification models from the training data and use the learned models for prediction. To evaluate a classification model's performance, the model is applied to the test data to obtain classification accuracy. Unsupervised learning algorithms are designed for data without class labels; clustering is a common example of unsupervised learning. For a given task, unsupervised learning algorithms build the model based on the similarity or dissimilarity between data objects (Barbier, 2011).

Data mining has processes involved in knowledge discovery, which are as follows:

- Data cleaning for removing noise or irrelevant data.

- Data integration where multiple data sources may be integrated or combined.
- Data selection is where data relevant to the user's request or analysis task is retrieved from the database.
- Data transformation is where data is transformed or consolidated into suitable forms for mining, by performing summary or aggregation operations, for instance.
- Pattern discovery is where patterns are discovered by applying intelligent methods.
- Pattern evaluation identifies the truly interesting patterns representing knowledge, based on interesting measures.
- Knowledge presentation to visualize the patterns in different forms.

(Pushpam et al., 2017).

Data available in social networks are user-generated content, which is vast, noisy, distributed, unstructured, and dynamic. These inadvertent characteristics pose challenges to data mining tasks to develop efficient algorithms and techniques (Nandi et al., 2013). Table 2.2 shows the data mining techniques and discussions on how those techniques are used.

Techniques	Discussion
Characterization	Characterization is used to generalize, summarize, and differentiate data characteristics.
Classification	Data classification is a process in which given data is categorized into different classes according to a model.

Regression	This process is like classification; however, the major difference is that the object to be predicted is continuous rather than discrete.
Association	In this process, an association between objects is found. It discovers the relationship between various databases and between the attributes of a single database.
Clustering	Clustering involves grouping data into several new classes such that they describe the data. The procedure involves breaking large data sets into smaller groups to make the designing and implementation process simpler. The task of clustering is to maximize the similarity between the objects in the classes, thereby, reduce the similarity between the classes
Change Detection	This method identifies any significant changes in the data from the previously measured values.
Deviation Detection	Deviation detection focuses on the major deviations between the actual values of the objects and their expected values. This method finds out the deviation according to the time, as well as, among different subsets of data.

Link Analysis	It traces the connections between objects to develop models based on the patterns in the relationships, by applying graph theory techniques.
Sequential Pattern Mining	This method involves the discovery of frequently-occurring patterns in the data.

Table 2.2: Data mining techniques (Jin, et al., 2015).

2.7. Social Media Data

Social media allows people to share and exchange content in online networks (Kaplan et al., 2010). Social media platforms enable users to share posts containing text, images, and videos; users can then like, share, and comment on each other's posts, forming a network of users and content. Personal profiles and posts can be private or public, depending on the platform and the user's settings (Lange, 2007). Most social media platforms allow users to geotag their posts, which makes social media data comparable to other types of geographic information. While 'social media' is a broad concept, in this study, the focus is specifically on social networking sites and content communities such as Twitter and other websites that may contain COVID-19 misinformation. Social media also plays a significant role in mid- and post-disaster management. This is achieved through listening, integrating, collaborative networking, creating cohesion, fundraising, and providing monitoring tools and research media (Machmud, 2021), therefore, the government is empowered with the ability to obtain accurate information and take rapid action in unexpected situations. Online data collection takes two main forms: conducting 'traditional' methods online and using 'naturally occurring' online data. Traditional methods, such as interviews and group discussions can be conducted online; secondly, social media is also a rich source of ‘naturally occurring’ data. Researchers can use textual, photographic, and video data created and shared on social media to answer research questions. Over recent years, researchers have been able to collate vast quantities of information from social network websites known as ‘data mining’. Data mining involves “examining large sets of pre-existing data to produce new information” (Beninger et al., 2014). The use of data mining of tweets to better understand complex social issues has been growing in popularity, amongst researchers (Beninger et al., 2014).

There are different approaches to acquiring social media data ranging from manual searches to programmatic access to data (Batrinca et al., 2015). The approaches vary in the required skill level of the analyst and the volume of data acquired. Manually browsing through social media groups is time-consuming but requires little computational skill, while obtaining data using automated tools typically requires technical knowledge about the interface, as well as programming skills, and appropriate computing infrastructure for continuous data storage. The following are data acquisitions approaches:

Application Programming Interfaces (APIs) provide a defined set of methods for programmatically interacting with social media platforms. Social media platforms often provide an API through which third parties, such as application developers and researchers, so they can interact with the platform automatically, which makes APIs an efficient tool for researchers (Lomborg et al., 2014). Streaming APIs continuously deliver newly posted messages, for example, on a specific topic in a read-only format. RESTful API interfaces are used by requesting specific data from the API and allowing more flexible queries back in time.

Purchase of data from an authorized data vendor has several advantages, such as little to no manual work and programming effort, and the opportunity to access time series of historical data.

Web scraping, or web crawling, is an approach for downloading and extracting data from web pages using an automated script. Compared to APIs, web crawlers can only access the public web, while APIs may provide access to content requiring authentication (Toivonen et al., 2019). Regardless of the data acquisition method, a researcher is responsible for storing and analyzing the data ethically and responsibly because each social media platform has its 'Terms of Service' that define how data acquired from these platforms can be stored and used by third parties. These conditions should be acknowledged when retrieving data for research purposes (Batrinca, et al., 2015). Social media data are generally referred to as 'big data' (Crampton, et al., 2013), and the volume of user-generated data worldwide is overwhelming, however, the amount of data needed in the end-analysis can be significantly reduced if one is interested only in a specific topic or region. The data are generally rich in visual and textual content and variable across topics and languages and despite the evident noisiness of these data, social media content has been found to contain relevant information for studying human-nature interactions from a relatively local to a global scale. Overall, social media data seems most useful for studying relatively broad areas, frequently visited by people or popular topics among social media users (Toivonen et al., 2019).

2.8. Information Cascade

The developments of AI based on machine learning techniques through Big Data, have raised multiple ethical and legal concerns, all of which ultimately hinge on the issues of responsibility, which is increasingly invoked, not as a remedy but as a character that should shape the whole

development process of AI, as well as its functioning (Gordoni, 2020). Some aspects of AI, taken in its scientific and communal act, threaten some established ideas related to human agency. Currently, there is evidence pointing to the adverse effects of algorithmic formation on human datasets, as often, such algorithms mirror and exacerbate existing inequalities in the input data (Stoica, 2020).
The perception of accountability is related to the necessity to report and justify automated decision-making. This implies knowing how a system operates; in other words, being able to audit a system. Auditing is the process whereby an external third-party, with no conflict of interest examines and evaluates the performance of a system to ensure no harm is derived from its use (Unceta et al., 2020). Accountability should facilitate action; actionable accountability can, therefore, be understood as a mechanism that involves auditing and mitigation.

The societal concerns posed by machine-learning have been discussed for decades, but the contemporary development of automated decision-making represents a major societal concern, given the changed societal context, as regards the ethical, legal, economic, and wide societal implications of AI (Gorgoni, 2020). The world is rapidly moving from the times of the internet to the algorithmic society, and soon, we will look back on the digital age as the forerunner to the algorithmic society. While the role of technology is under debate in the inequality or the bias of algorithms, this study aims to unravel the power of computational tools in analyzing instances of social bias on a large scale, as well as evaluating the effects of certain algorithms that are determined from biased data on the social bias. It is of utmost importance to understand the disparate effects of algorithmic accountability on social inequality and formulate a fair and explainable framework in designing such tools for prediction (Stoica, 2020).

This study intertwines theoretical underpinnings for explaining such effects and building interventions for algorithmic accountability, starting from defining - what 'algorithmic accountability' means, what properties of social media lead to differentiated outcomes, and what reasonable means, in different contexts of social media, finally leading to a re-designing of learning algorithms to produce more reasonable results. There is the ubiquity of health-related communications via social media, despite this, no consensus has emerged on whether this medium, on the balance, jeopardizes or promotes public health. During the COVID-19 pandemic, social media were described as either, the source of toxic Infodemics or a valuable tool for public health as no conceptual model exists for examining the roles that social media can play concerning population health (Schillinger et al., 2020).

Proactive communications are usually generated by public health entities that effectively prevent or minimize the spread of misinformation and increase public awareness of accurate information (Southwell, et al., 2020). Injected messages are created and disseminated as a reaction to misinformation. In response to misinformation concerns related to COVID-19, WHO's risk communication team, launched a new information platform called 'WHO Information Network for Epidemics', which uses social media amplifiers to share tailored information with target groups (Pagoto, 2019). Computer-language methods allow the exploitation of big data from social media to identify emerging trends, track behavioral changes, and detect or even predict disease outbreaks (Schillinger, et al., 2020).

An overabundance of information - some accurate and some not- makes it difficult for people to separate false from true information easily. Infodemics and the spread of health misinformation are global and multifaceted phenomena, calling for both specific theoretical frameworks and applied tools for measurement and analysis, as well as for the coordination of public health institutions, policymakers, information professionals, researchers, journalists, information technology experts, digital platforms and the entire civil society in containing their negative effects when utilizing evidence-based joint efforts and policies (Sell , et al., 2020).

2.9. Big Data Analytics

More than 64.5 million people globally were infected by the virus; about 41.5 million recovered and 1.49 million of them have died from this virus (Template: COVID-19 pandemic data – Wikipedia, 2023). This magnitude of cases and their health data have created a vital source of information and knowledge. There is, hence, an urgent requirement to store such a large amount of data, using different data-storage technologies. These data are used to undertake research and development about viruses, pandemics, measures to fight these viruses and their after-effects.

Big data is an innovative technology that can digitally store a large amount of data on patients; t helps to computationally analyze the stored data to reveal patterns, trends, associations, and differences; it can also help in revealing insights into the spread and control of this virus. With detailed data capturing capability, big data can be used gainfully to minimize the risk of spreading viruses (Wang, et al., 2020).

Big data has been employed in a wide range of industrial application domains, including healthcare where electronic healthcare records (EHRs) are exploited by using intelligent analytics for facilitating medical services; for example, health big data potentially supports patient-health analysis, diagnosis assistance, and drug manufacturing (Priyanka, 2014). Big data has the potential for fighting pandemics, like COVID-19, via four main application domains: - outbreak prediction; virus spread tracking; virus diagnosis/ treatment, and vaccine/drug discovery - as shown in Fig. 3 below (Pham, 2020).

China has employed the COVID-19-related data from authoritative sources, such as national, provincial, and municipal health commissions (Data governance, 2019). This big data source helped implement pandemic modeling to interpret the cumulative numbers of infected people, and recovered cases in five different regions, - the Mainland, Hubei, Wuhan, Beijing, and Shanghai. In addition, the process has allowed the performance of simulations to predict the trends in COVID-19 and other outbreaks, through identifying the areas at high risk of pandemics and detecting the population with increasing infected cases, all of which contribute to the success of anti-pandemic campaigns. Big data, hence, enables outbreak prediction on a global scope. From the data analysis point of view, the outbreak was predicted using available data points that put the accuracy of reliable estimation under doubt, due to the lack of comprehensive investigations. Accuracy may depend on the number of contributory factors, ranging from infected cases, population, living conditions, and environments.

Big data technology can store a massive amount of information about the people who were infected with the COVID-19 virus. It helps in understanding, in detail, the nature of this virus. The data obtained can further be trained over again for developing future preventive methods. This technology has been used to store the data of all types of cases - infected, recovered, and expired - affected by COVID-19. This information can be effectively used for case identification and helping to allocate the resources for better protection of public health. Several modalities of digital data including - patient location, proximity, patient-reported travel, co-morbidity, patient physiology, and current symptoms - can be digitized and used for generating actionable insights at both community and demographic levels. Figure 2.2 shows the processes of how Big Data can fight pandemics, like COVID-19.

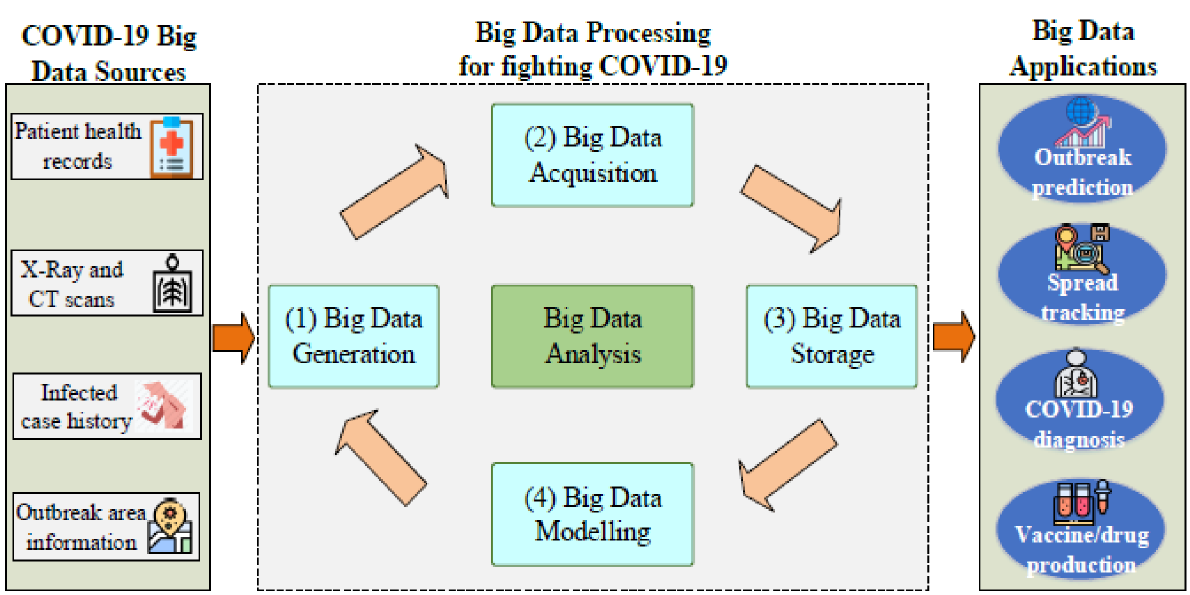


Figure 2.2: Big Data Processing for fighting COVID-19 (Shah, et al.,2022).

Big data provides a massive amount of information to scientists, health workers and epidemiologists to help them to make informed decisions on how to fight viruses, like the COVID-19 one. This data can be used to track viruses on a global basis, continuously and to create innovation in the medical field (Xia, 2020). It can help to forecast the impact of viruses, like COVID-19, in a particular area and the whole population. It helps in the research and development of new treatment procedures. Big data can also provide possible sources and opportunities for people, thus helping to handle stressful situations. Overall, this technology provides data to undertake - analysis of disease transmission, its movement, health monitoring and prevention system.

BDA will act as a medium for tracking, controlling, researching, and preventing the reoccurrence of COVID-19 as a pandemic. It will diversify manufacturing, and enhance vaccine development on a more profound scale, with absolute knowledge. Prevalent modeled data helps in understanding and offers an edge over the other process, like corresponding homology models. These are predicted by fold and function assignment system server for each target protein which was downloaded from Protein Data Bank, for predicting a COVID-19 cure and identifying the symptoms associated with the disease (Amini et al., 2017; Wu, 2020).

Big data provides insights and analyses into the factors leading to better containment of infected persons. China suppressed COVID-19 with the help of data collection and implemented with AI led to a low rate of spread. There are several big data components to this pandemic, where AI can play a significant role, like biomedical research, NLP, social media, and mining scientific literature. Big data provides information to identify any suspected cases of this virus. It helps to provide an efficient way to prevent illness and extract other valuable information. In the future, big data will help the public, doctors, other healthcare professionals, and researchers to track this virus and analyze the infection mechanism of COVID-19. The data provided helps to examine how this infection can be slowed or eventually prevented and helps to optimize the allocation of resources, hence, contributes to taking appropriate and timely decisions. With the assistance of this digital data-storing technology, doctors and scientists can also develop a convenient and efficient method of COVID-19 testing (Haleem et al., 2020).

S. No	Applications	Description
1.	Identification of infected cases	It can store the complete medical history of all patients, due to its capability of storing a massive amount of data. By using the captured data, this technology helps in the identification of infected cases and undertakes further analysis of the people's level of risks.
2.	Travel history	Used to store the travel history of the people to analyze the risk. Helps to identify people who may be in contact with infected patients of this virus.
3.	Fever symptoms	Big data can keep a record of the fever and other symptoms of a patient and suggest if medical attention is required. Helps to identify suspicious cases and other misinformation using the appropriate data
4.	Identification of the virus at an early stage	Quickly helps to identify the infected patient, at an early stage. Helps to analyze and identify persons who can be infected by this virus in future.
5.	Identification and analysis of fast-moving disease	Helps to effectively analyze any fast-moving disease as efficiently as possible.

		Has the potential to handle appropriate information regarding a disease.
6.	Information during lockdown	This technology collects information regarding this virus during any lockdown that may have been imposed. Track and monitor the movement of people and the entire health management process.
7.	People entering or leaving an affected area.	It helps to analyze the number of people entering or leaving an affected city. With these vast amounts of data, health specialists can quickly identify the possible behavior of the virus in those peoples.
8.	Faster development of medical treatments	Assist in fast-tracking the development of new medicines and equipment needed for current and future medicinal needs. Provides previous data on viruses, thus, helps in gaining an advantage over newer pandemics/epidemics using previously analyzed results.

Table 2.3: Applications of big data in the COVID-19 pandemic (Haleem, et al., (n.d.).

2.10. Understanding Infodemics

These days every hour we are inundated by so much data and information, that the WHO coined a word for it 'infodemics'. Infodemics is defined as an overabundance of information – some accurate and some not – that makes it hard for people to find trustworthy sources and reliable guidance when they need them (Culp, 2020). COVID-19 infodemics spans four major thematic areas where people looked for trustworthy information. information and where there were misinformation and rumors. These areas were - the cause and origin of the virus and disease; its symptoms and transmission patterns; available treatments, prophylactics, and cures; and the effectiveness and impact of interventions by health authorities or other institutions (Culp, 2020). The coronavirus marks the first social media infodemics, and the participatory nature of online platforms has allowed disinformation to spread and flourish at unprecedented speed, creating an environment of heightened uncertainty that has fueled anxiety and racism in persons, and online (Sloan, 2020). Increased global access to cell phones with an internet connection, as well as social media, has led to the exponential production of information and the number of possible paths for getting it, creating information infodemics (Saúde, 2020). This simply means that an enormous amount of information is produced and shared all over the world, reaching an unknown number of people.

2.11. Misinformation

The timing is challenging, as platforms are fighting to contain an epidemic of misinformation, with user traffic hitting records. To make up for the absence of human reviewers, platforms largely handed over the role of moderating content to algorithmic systems. As a result, machines currently have more agency over the regulation of our public discourse than ever before. In mid-February, the WHO announced that the new coronavirus pandemic was accompanied by an 'infodemics' of misinformation (WHO, 2020).

Misinformation- and disinformation about science, technology, and health is neither new nor unique to COVID-19. Amid an unprecedented global health crisis, many journalists, policymakers, and academics have echoed their concerns to WHO and stressed that misinformation about the pandemic presents a serious risk to public health and public action (Mihailidis, et al., 2017). Generally, fake news is primarily spread via social media platforms, such as Facebook, Twitter, Reddit, YouTube, and Instagram, among others (Mihailidis, et al, 2017; Pew Research Center, 2018). Fake news is hard for some people to identify and can create confusion about what is true and doubts about accurate information; reliance on falsehoods, and is continuously shared among members of the public, unknowingly (Jang et al, 2018; Rapp et al, 2018). With the rapid dissemination of fake news through social media, users are bombarded with false information, regularly; much of this false information is based on the agendas of partisan news coverage (Vargo et al., 2017). Usually, when consumers learn information via fake news that information is persistent and long-lasting (Jeong-woo et al., 2019). It usually has more impact on consumers than information from true news sources, most likely because of its sensational nature and shock value (Wright et al., 2019).

While research has demonstrated that there are specific consumer characteristics that may make them a little more susceptible to believing fake news, including being White, male, and identifying as evangelical Christian and conservative (Wright et al., 2019), the main way we can combat fake news is by educating ourselves about what it is and what it is not. We must do our fact-checking when we are presented with information from any source, including hard news sources, government agencies, social media, friends, and family. Snopes is a great source for fact-checking news but the best place to get accurate information related to COVID-19 is the Centre for Disease Control and Prevention. The biggest challenge, however, is that this misinformation or fakes news is drowning official public health advice on COVID-19, making it extremely problematic for the voices of healthcare professionals to be heard; the implications of this may be enormous as the virus continues to spread between and within diverse populations (Ahinkorah et al., 2020).

2.12. Covid-19 Misinformation in South Africa

The arrival of the COVID-19 virus in South Africa saw an increase in the dissemination of misinformation about the virus on social media and other platforms. These range from messages minimizing the virus's harm in the country (Davis, 2020), to the propagation of conspiracy theories about government actions to control the virus. Deliberately spreading fake news and other misinformation in South Africa about the virus was declared an offense punishable by a fine, six months imprisonment, or both.

One individual was arrested for posting a video showing himself drinking in public with friends following the national lockdown, whilst stating that there was "nothing called corona here" (Davis, 2020). In another incident, a man claimed that 10000 government officials would be going door-to-door using contaminated test kits to test people for the virus (COVID-19 Test Kits, 2022). A conspiracy theory that Bills Gates wished to test a COVID-19 "vaccine" in Africa or South Africa first caused significant controversy on social media following the publication of a now-retracted story in News24. Fake news that 5G cellular technology was the true cause of COVID-19 symptoms also spread in the country during this period as it did in other countries worldwide (Cilliers, 2020).

2.13. Sense-making perspective on Algorithmics.

Governments, institutions, and citizens of nearly every nation have been compelled to respond to COVID-19. Many measures have been adopted, including contact-tracing and risk-assessment, whereby citizens' whereabouts were constantly monitored to trace contact with other infectious individuals and isolate contagious parties via algorithmic evaluation of their risk status (Liu, 2020) The relationship between algorithms, data, and individuals is dynamic and tempestuous in the era of big data. Algorithms and related technologies have been employed for multiple social purposes, including but not limited to public health efforts to fight COVID-19 (Liu et al., 2020).

Sense-making is essentially something actors do as they engage in practice. Established literature conceptualizes sense-making as an ongoing, situated process that involves the creation of coherent understandings through interlinked, observation, interpretation, and action. As individuals and groups 'enact' order in the chaotic or uncertain situations they encounter through their actions, they generate new observations that trigger reinterpretation (Schildt et al., 2020). Sense-making involves turning factors into a situation that is appreciated explicitly in words and that serves as a starting point for action. The seemingly transient nature of sense-making contradicts its significant role in the determination of human behavior, whether people are acting in formal organizations or otherwise. Sense-making is central because it is the primary site where meanings materialize that inform and constrain identity and action. In terms of 'application', sense-making theorists have often been, especially, interested in individual and collective decision-making and processes of group and organizational change (Brown et al., 2015).

Sense-making enables stakeholders to better understand what is going on within their environments, thus facilitating other relevant activities, such as visioning, relating, and inventing. The move to the complex occurs as new information is collected, and new actions are taken. Then as patterns are identified, and the new information is labeled and categorized, the complex becomes simple once again, albeit with a higher level of understanding. Sense-making is most often needed when our understanding of the world somehow becomes unintelligent. This occurs when the environment is changing rapidly, presenting us with surprises for which we are unprepared or confronting us with adaptive rather than technical problems to solve (Ancona, 2012). We learn the most about events or issues when we view them from a variety of perspectives; while each may have its flaws, when the different modes of analysis reveal the same patterns, we can feel more confident as we converge on an interpretation of what is going on (Weick, 1995).

2.14. Coronavirus Misinformation Tracking systems

As a new strain of the coronavirus spreads across the globe, so does disinformation and misinformation, following the spread of this dangerous information, with My Broadband's Coronavirus Misinformation Tracking Center. Red-Rated news and information sites in South Africa have been identified as publishing materially false information about the virus. These are websites that are notorious for publishing false health content and political sites whose embrace of conspiracy theories extend well beyond politics. Among the hoaxes these sites published were that swallowing bleach or colloidal silver will prevent the coronavirus—when in fact these treatments can be harmful, according to - African News Updates, Blackopinion.co.za, mzasitimes.com, nationalnewsbulletin.com, news24-TV, and Pretoria Live – which are some of the South African-based fake news sites (Bangalee et al., 2021)

2.15. Challenges in Search of Covid- 19 Information

Beyond the challenges of data analysis, there were other challenges related to research on COVID-19, as discussed next.

2.15.1. Searching for relevant information sources

Every day there is a tremendous growth of articles published on the COVID-19 pandemic. For researchers in the field of evidence synthesis, the challenge is searching for relevant information sources. Creating a specialized and publicly accessible collection of studies that specifically focus on COVID-19 will be very helpful. WHO has set up a collection of articles about COVID-19 compiled in a publicly available database (Wolkewitz et al, 2020).

2.15.2. Incorporating information swiftly

In a world where each day brings tons of new articles on a well-known topic, conducting evidence synthesis is particularly challenging. Systematic reviews are considered by many as the highest level of evidence in the hierarchy of evidence in medicine, but their production often takes years (Runjic, et al, 2019), however, multiple systematic reviews about COVID-19 have already been published. What remains to be seen is the quality of these rapidly produced systematic reviews. Producing evidence syntheses on a short-time scale usually requires cutting corners with the methodology, and for this reason, rapid reviews have evolved. Rapid reviews are conducted with a condensed timeline, sacrificing certain aspects of systematic review methodology for speed (Garritty et al, 2019). A pilot study has shown that rapid research that needs appraisal can be conducted within five days in the case of an infectious disease outbreak (Sigfrid et al., 2019), however, it has also been shown that lack of transparency and inadequate reporting are the major limitations of such reviews (Kelly et al., 2016).

2.15.3. Ensure acceptable quality of published research.

According to Wolkewitz et al., (2020), journal editors are currently under pressure to publish relevant articles on COVID-19 quickly, which these editors have described as “rather maddening”. It has been argued, however, that this could also be advantageous in the long run, as it could help journals to become more efficient in the future, although, haste is likely to be detrimental to the quality of publications. Speed is not necessarily a friend of good science. Articles may be assembled too quickly, publishing processes may be hastened, hence, the quality of the peer reviewing may not be adequate.

Anecdotal reports indicate that highly specialized experts in the field may be swamped with requests for peer reviews that they are unable to accommodate, which may lead to inviting less specialized peer-reviewers, to the detriment of manuscript quality (Wolkewitz et al., 2020). There is a need to wait to find out how many corrections and retractions there will be for journals published hastily on the topic of COVID-19, and whether the methodological and reporting quality of those articles will be lower, compared to articles on other topics. In times of emergency, researchers should still pay attention to transparency and adequate reporting of their research, to ensure its quality, reproducibility, and usefulness.

2.15.4. Data Distribution

To enable analysis of data gathered during the COVID-19 pandemic, principles of open science and raw data sharing will be of utmost importance. Modjarrad et al., (2016) proposed global norms for data sharing during global health emergencies, and it remains to be seen whether researchers will be more likely to share their raw data publicly in articles covering COVID-19.

2.16. Data Governance

Data governance refers to the exercise of authority and control over the management of data. The purpose of data governance is to increase the value of data and minimize data-related costs and risks (Data Governance, 2019). Governance refers to what decisions must be made to ensure effective management and use of IT, such as decision domains, and who makes the decisions, such as locus of accountability for decision-making. Governance includes establishing who in the organization holds decision rights for determining standards for data quality. Management involves determining the actual metrics employed for data quality (Vijay et al., 2010), The definition of data governance includes who holds the decision rights and accountability regarding an enterprise's data assets therefore, the decision domains should be identified to assign the right responsibilities and duties. In reviewing the literature relating to data governance frameworks, Khatri et al., (2010) proposed a framework for data that contains five interrelated decision domains. In these domains, the data principles are shown at the top of the framework as they are intended to establish the direction for all other decision domains, hence, the principles set the boundary requirements for the use of data assets, which address the organization's standards for data quality. The data quality then refines the basis for how data are interpreted (metadata) as well as accessed (data access) by users. Lastly, the data life-cycle decision defines the production, retention, and retirement of data assets which play a fundamental role in operationalizing the data principles into the IT infrastructure (Alhassan et al., 2016).

<table>
<tr><td colspan="3">DATA PRINCIPLES</td></tr>
<tr><td rowspan="2">DATA QUALITY</td><td>METADATA</td><td rowspan="2">DATA LIFE CYCLE</td></tr>
<tr><td>DATA ACCESS</td></tr>
</table>

Table 2.4: Decision domains for data governance (Khatri, et al.,2010)

Manene et al., (2023) proposed a framework that has been illustrated using two interacting activity systems, representing the activities set by the two subjects - social media users and regulators. The authors mentioned subjects which refer to all the individuals and groups participating in an activity. These people are the overall focus of a study and will be responsible for implementing the strategies identified in the framework. Figure 2.3 shows the proposed framework.

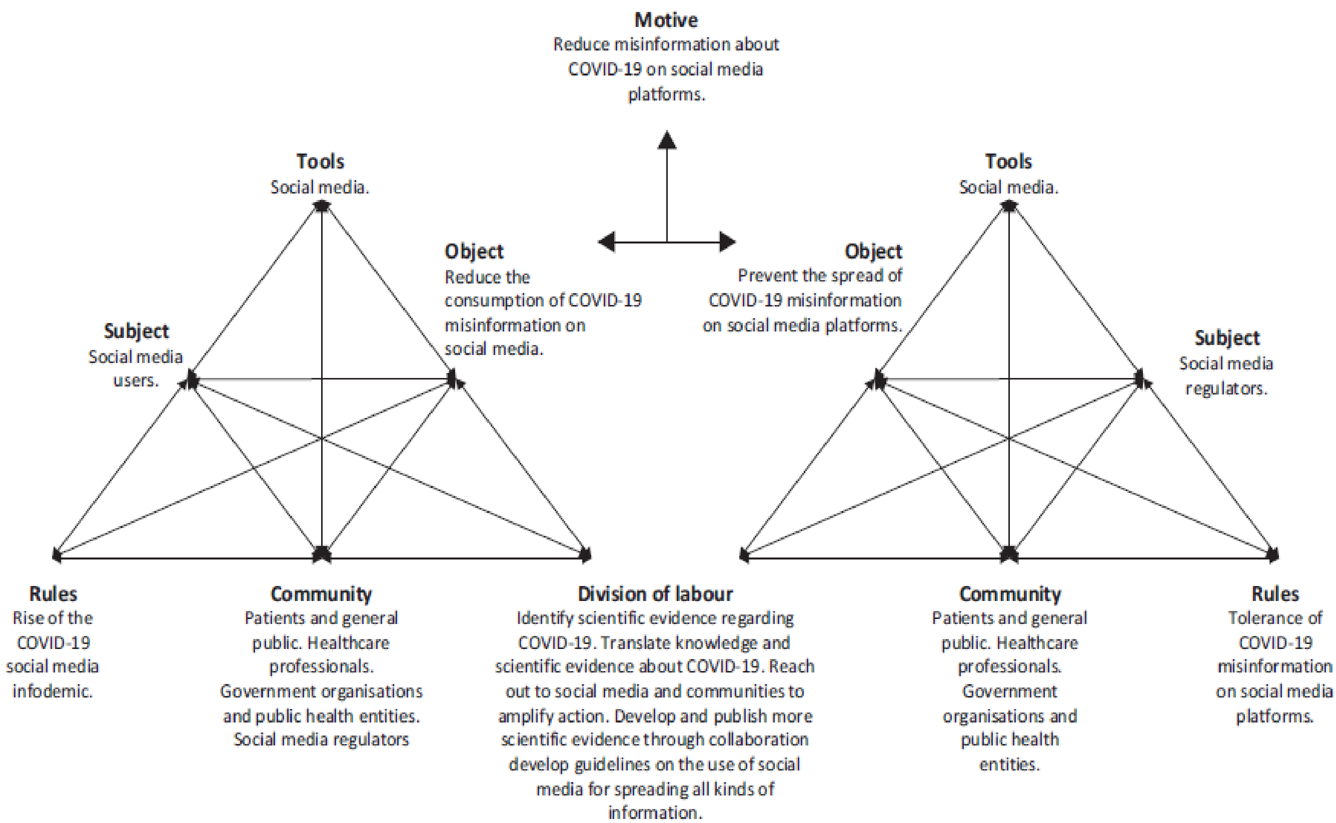


Figure 2.3: Proposed framework for managing the coronavirus disease's 2019 infodemic (Manene et al.,2023).

Based on the literature review on COVID-19 misinformation and infodemics on social media, there is much lack of algorithmic accountability on how to combat such misinformation being populated on social media. The researcher is proposing a conceptual framework for a sense-making perspective of algorithmic accountability during infodemics. The framework will help in combatting and channeling algorithmic accountability on social media and how the spread of misinformation can be avoided by individuals, governments, and the private sector.

2.17. Chapter Summary

This chapter reviewed the related literature on sense-making perspective and algorithmic accountability during infodemics. The concepts of BDA and models that were used for their deployment in organizations were presented; secondly, the literature relating to COVID-19 infodemics was reviewed. Based on this literature, there were many methodological challenges related to producing, gathering, analyzing, reporting, and publishing data in condensed timelines required during a pandemic. The discussions, however, did not mention all of them, but this study contributed to research methodology related to COVID-19 and helped to address those other issues as well. It was customarily said that each crisis was also an opportunity.

Chapter 3: RESEARCH METHODOLOGY

3.1. Overview

Chapter Three provided a summary of the research methodology that was used in conducting the study. This section covered research methodology, area of study, research design, study population, data collection techniques, and ethical considerations. The limitations that the researcher had encountered during the research process were highlighted, for any field of study encouraged a researcher to adapt his or her study to the current situation (Sarantakos, 2007)

3.2 **RESEARCH PARADIGM**

A research paradigm is the conceptual lens through which researchers examine the methodological aspects of their research project to determine the research methods that should be used and how the data will be analyzed. Research paradigm comprises four elements, namely, ontology, epistemology, axiology, and methodology (Kivunj et al., 2017).

3.2.1. Ontology

On the 31st of December 2019, the WHO in China was informed by the Chinese authorities of a series of pneumonia cases with unknown causes in Wuhan, Hubei, China, with clinical presentations that greatly resembled viral pneumonia. The Chinese authorities had isolated a causal agent on 7th January 2020, which was identified as a new type of coronavirus. Initially, the virus was named ‘2019 novel coronaviruses, but the official names of the virus and the disease it caused were announced soon after; the official name of the virus was ‘severe acute respiratory syndrome coronavirus 2’ (SARS-CoV-2), and the disease it caused was the coronavirus disease. The coronavirus disease was the first pandemic in history in which technology and social media were being used on a massive scale to keep people safe, informed, productive, and connected. At the same time, the technology relied on to keep people connected and informed was enabling and amplifying infodemics that continued to undermine the global response and jeopardize measures to control the pandemic.

An infodemic, as explained earlier, was an overabundance of information, both online and offline. It included deliberate attempts to disseminate wrong information to undermine the public health responses and advance alternative agendas of groups or individuals (Culp, 2020). Mis- and disinformation can be harmful to people's physical and mental health, increase stigmatization, threaten precious health gains, and lead to poor observance of public health measures, thus, reducing their effectiveness and endangering countries' ability to stop the pandemic.

3.2.2. Epistemology

Misinformation costs lives. Without the appropriate trust and correct information, diagnostic tests go unused, immunization campaigns will not meet their targets, and the virus will continue to thrive. Disinformation polarized public debate on topics related to COVID-19, amplified hate speech, heightened the risk of conflict, violence, and human rights violations, and threatened long-term prospects for advancing democracy, human rights, and social cohesion. This study aimed to give a sense-making perspective on the algorithmic accountability on infodemics during this pandemic, for which a framework was developed.

3.2.3. Methodology

Methodology articulates the logic and flow of the systematic processes followed in conducting a research project, to gain knowledge about a research problem. The methodology includes assumptions made, limitations encountered, and how they were mitigated or minimized. Mixed methods were used in this research as the topic required a purposeful mixing of methods in data collection, data analysis, and interpretation of the evidence. The key word is 'mixed'; this is an essential step in data linkage or integration at an appropriate stage in the research process.

3.2.4. Axiology

Axiology involves defining, evaluating, and understanding concepts of right and wrong behavior relating to research. Axiology considers what value shall be attributed to the different aspects of research - the participants, the data, and the audience to which we shall report the results of the research. This study used the COVID-19 misinformation where the emergence of social media as a key source of news content, had created a new ecosystem for the spreading of misinformation. This situation was illustrated by the recent rise of an old form of misinformation - blatant false news stories that are presented as if they are legitimate. Many people consume news via social media; therefore, it is desirable to reduce social media users' exposure to low-quality news content. One possible intervention is for social media ranking algorithms to show relatively less content from sources that users deem to be untrustworthy. The question, however, is - *Are lay people's judgments reliable indicators of quality, or are they corrupted by either partisan bias or lack of information?* Perhaps surprisingly, I find that lay people on average are quite good at distinguishing between low- and high-quality sources. These results indicate that incorporating the trust ratings of lay people into social media ranking algorithms may prove an effective intervention against misinformation and news content with heavy political bias.

3.3. Research Design

This study was conducted in two phases, beginning with the collection of COVID-19 misinformation posts using a combination of web scraping on websites and filtering the public streaming Twitter application programming interface for keywords associated with suspects of COVID-19 misinformation in which was NLP. The second phase involved data analysis using (NLP) which is a theoretically motivated range of computational techniques for analyzing and representing naturally occurring texts at one or more levels of linguistic analysis to achieve human-like language processing for a range of tasks or applications (Zhao et al., 2020). The NLP, therefore, will be used to identify potential misinformation in the African News Updates, black opinion.co.za, mzansitimes.com, Nationalnewsbulletin.com, News24-TV, Pretoria Live websites.

The Design Science method was utilized to understand the link between algorithmic accountability and infodemics at the national level and qualitative techniques were used to achieve the study objectives. A qualitative case studies using NLP was used. NLP methods for textual content analysis is a field that studies the computational processing of all natural languages, such as English or Swahili. NLP methods can be used to analyses content on social media platforms, such as Instagram captions or Twitter posts. Applications of NLP methods in conservation science have been rare but hold much potential for extracting useful information from textual content (Becken et al., 2017). NLP methods allow, for instance, automatic language identification, sentiment analysis, and named entity recognition for extracting the names of locations, organizations, individuals, and species mentioned in social media posts (Toivonen et al., 2019). The researcher developed a scraper in Python to fetch the data and news from the sources in which the sentiment analysis and topic modelling were used as the data analysis method to obtain the findings. Topic modelling helps in exploring large amounts of text data, finding clusters of words, the similarities between documents, and discovering abstract topics. Topic modelling is also used in search engines wherein the search string is matched with the results (Mining, 2018) and again the topic is defined as a group of similar articles which roughly talk about the same subject (Albanese et al., 2020).

3.4. Experimental research design

Experimental research is a scientific approach to research, where one or more independent variables are manipulated and applied to one or more dependent variables to measure their effect on the latter. This was an experimental study as it used the Natural Language Processes whose aim was to gather COVID-19 misinformation using Python as a language; the intention here was to discover how human beings understand and use languages, so that appropriate tools and techniques can be developed to make computer systems understand and manipulate natural language to perform desired tasks (Gobinda et al., 2005).

3.5. Topic modeling in studying algorithmic accountability.

In a recent study, the use of technological methods in communication had attracted increasing interest. However, applying these methods required both expertise and material resources, such as servers for executing programs (Smith.,2005). In this study, the researcher investigated algorithmic accountability during infodemics, in which AI played a major role, and analyzed the empirical Twitter data from users.

The researcher investigated what Python language could provide for multi-method media event research, where the media was the starting point, by running topic modeling for the data consisting of 21 of 508 tweets. Topic modeling provided a way to compress the big data qualitative phase of the research.

3.6. Qualitative, Quantitative, and Mixed Approaches

Qualitative and quantitative research approaches are utilized rather frequently in different disciplines, such as sociology, psychology, and history. Regarding the research approaches, there persists the so-called ‘paradigm wars’ in which researchers belong to two distinct camps—interpretivism and positivism. The positivistic researchers believe that the social world consists of concrete and unchangeable reality which can be quantified objectively; whereas the interpretive researchers oppose the positivistic belief of reality and argue that, instead, reality is socially constructed by humans, which can be changed and understood subjectively (Rahman, 2016).

Quantitative analysis between algorithmic accountability and the context of infodemics involved examining numerical data to identify patterns, trends, correlations, and associations. For instance, researchers utilized statistical techniques to explore the impact of algorithmic decision-making on the spread of misinformation during infodemics (Smith et al., 2020). Machine learning algorithms were employed to analyze large datasets and predict misinformation dissemination patterns (Jones et al., 2019). Network analysis methods were also utilized to examine the structure of information networks and identify influential actors in the spread of misinformation (Gao et al., 2019). These quantitative approaches provided valuable insights into the complex dynamics between algorithmic accountability and infodemics, enabling researchers to inform evidence-based interventions and policy responses. Data visualization tools and techniques were employed to visually represent quantitative findings and insights. Visualizations such as charts, graphs, and heatmaps helped communicate complex relationships and trends between algorithmic accountability and infodemics in a clear and understandable manner.

Mixed-method research requires a purposeful mixing of methods in data collection, data analysis, and interpretation of the evidence. The key word is ‘mixed’; this is an essential step in data linkage or integration at an appropriate stage in the research process. Purposeful data integration enables researchers to seek a more comprehensive view of their research landscape, viewing phenomena from different viewpoints and through diverse research lenses (Shorten et al, (2017). Selecting the right research method starts with identifying the research question and study aims. A mixed methods design is appropriate for answering research questions that neither quantitative nor qualitative methods could answer alone.

In the study, a mixed-research approach was employed to comprehensively explore the relationship between algorithmic accountability and the context of infodemics. This methodological choice was motivated by the acknowledgment that leveraging both qualitative and quantitative methodologies provided distinct advantages in unraveling the complex dynamics inherent in this intersection.

Through the combination of qualitative and quantitative data collection methods, researchers were able to gather a diverse range of insights and perspectives. Qualitative methods, such as interviews, focus groups, and case studies, facilitated an in-depth exploration of the intricate nuances and underlying mechanisms shaping algorithmic accountability practices within the infodemics context. Conversely, quantitative techniques, including data mining, statistical analysis, and network mapping, provided a broader perspective by quantifying patterns, trends, and correlations in the vast datasets characterizing digital information dissemination during infodemics. The integration of qualitative and quantitative data enabled triangulation, where findings from one method could be corroborated, enriched, or challenged by the other. This triangulation enhanced the validity and reliability of the study's findings by providing multiple lines of evidence and perspectives. For instance, qualitative insights into stakeholders' perceptions and experiences of algorithmic accountability were complemented by quantitative analyses of algorithmic impacts on information dynamics and public discourse.

The mixed-research approach facilitates a more holistic understanding of the connections or contradictions between qualitative and quantitative data. It enables researchers to identify convergent patterns, discrepancies, or divergent findings across different data sources, thereby uncovering deeper insights into the underlying mechanisms driving algorithmic decision-making and its implications within the infodemics landscape.

By leveraging the complementary strengths of qualitative and quantitative methodologies, the mixed-research approach employed in this study aimed to provide a nuanced and comprehensive understanding of the intricate linkages between algorithmic accountability and the context of infodemics. Through rigorous analysis and synthesis of diverse data sources, this approach sought to illuminate key insights that could inform evidence-based interventions, policy recommendations, and future research directions in this critical domain. These two approaches provided opportunities for participants to have a strong voice and share their experiences across the research process, and they facilitated different avenues of exploration that enriched the evidence and enabled questions to be answered in-depth (Smith et al., 2017).

3.7. Research Strategy

This study allowed the adoption of a research approach that was qualitative. The research map below summarizes the overall research strategy.

	Research Element	***Research Focus***	***Objectives***	***Methods***
HISTORY	CONTEXT	*Macro Social Organization,* for example, infodemics networks and algorithmic accountability.	Determine the influence of algorithmic systems on infodemics	➢ Qualitative Case study using (NLP)
	SETTING	Uses a series of COVID-19 case studies to illustrate how a lack of transparency can cause problems.	Algorithmic systems problematize transparency.	
	SITUATED ACTIVITY	Understanding algorithms in social settings requires paying attention to individual sense-making practices and interpretation.	Algorithmic accountability can be realized in practice.	
	SELF	Necessary precaution to obtain - valid information, comforting information, and perplexing information.	Coronavirus misinformation can be confronted.	

Table 2: Research Strategy

To make sense of the link between algorithmic accountability and the Infodemics context, the mixed-research approach was guided by the following research objectives:

1. **Main objective: To determine the influence of algorithmic systems during infodemics**: Qualitative case studies using NLP.
2. **Algorithmic systems problematize transparency.** This objective uses a series of COVID-19 case studies to illustrate how a lack of transparency can cause problems, if undertaken before exploring the consequences that such a lack of transparency can have, as well as the complexities inherent in trying to achieve transparency in any given societal context.
3. **Algorithmic accountability can be realized in practice.** Understanding algorithms in social settings requires paying attention to individual sense-making practices and interpretation.
4. **Coronavirus misinformation can be confronted.** This can be done if the necessary precaution, information-valid, comforting information, and perplexing information are taken into consideration.

3.8. DATA COLLECTION PROCEDURES

This study first applied a systematic approach for conducting data mining on Twitter by filtering the public streaming API (Application Programming Interface) for keywords associated with COVID-19 misinformation, to collect a large corpus of general COVID-19–related conversations for the 21 days of South Africa’s lockdown. The selection of the 21-day period during South Africa's lockdown likely corresponded to a specific phase or timeframe of the COVID-19 pandemic in the region. South Africa had implemented various levels of lockdown measures during the pandemic to control the spread of the virus. The 21-day period may have been chosen based on its significance as a duration for observing changes in public discourse, misinformation trends, and policy responses during a specific phase of the pandemic. Researchers might have selected this period to capture a snapshot of Twitter conversations and misinformation dynamics during a critical juncture of the pandemic, such as the initial stages of the lockdown, a period of heightened public concern, or the emergence of significant events or developments related to COVID-19.

Additionally, the 21-day timeframe likely had practical considerations, such as ensuring a manageable scope for data collection and analysis while still capturing sufficient data to draw meaningful conclusions. The criteria for selecting specific keywords likely involved a combination of relevance to the research topic, prevalence in public discourse, and potential to capture relevant conversations on Twitter related to COVID-19 misinformation. Researchers might have considered keywords such as COVID-19, coronavirus, pandemic, lockdown, vaccine, misinformation, fake news, hoax, conspiracy, and variations thereof. The rationale behind selecting these keywords was based on their significance in discussions about COVID-19, their potential to capture different aspects of misinformation, and their relevance to the research objectives. The researcher identified general COVID-19–related keywords based on manual searches on each of the platforms, which included different iterations of COVID-19 misinformation (for example, COVID-19, COVID-19 misinformation); these keywords were converted into hashtags to conduct searches on Twitter. Texts of tweets posts were captured, as well as retweets and other metadata including - likes, favorites, comments, replies, use of similar hashtags, and associated media, hyperlinks, and metadata of posts. This metadata was used to identify any potential temporal trends associated with misinformation posts, and account characteristics, and to characterize hyperlinks to external websites that were embedded in misleading/misinformation posts. The data collected was populated in the Table in Appendix C.

3.9. DATA ANALYSIS

Quantitative research techniques generated a mass of numbers that needed to be summarized, described, and analyzed. Characteristics of the data were described and explored by drawing graphs and charts, doing cross-tabulations, and calculating means and standard deviations. Further analysis built on these initial findings, seeking patterns and relationships in the data by performing multiple regression or an analysis of variance. Advanced modeling techniques might eventually have been used to build sophisticated explanations of how the data addressed the original question; however, many quantitative research projects would never have needed to go that far; the question would have been answered by simple descriptive statistics. (Kraska et al., 2020).

This study discussed sentiment analysis, a method used to determine the sentiment expressed in text, which could be positive, negative, or neutral. It also highlighted how sentiment analysis was sometimes referred to as 'material polarity' or 'mining of opinions.'

The study then emphasized the growth and advancement of social media platforms, which had attracted a large number of users. Specifically, it mentioned Twitter, where users could post tweets limited to 280 characters. The concise nature of tweets made them suitable for sentiment analysis. Additionally, the statement provided context by mentioning that 550 million tweets were posted daily on Twitter.

Overall, the study underscored the significance of sentiment analysis in analyzing user sentiments expressed on social media platforms like Twitter, given the vast amount of data generated daily. It also highlighted how the characteristics of Twitter, such as the character limit, contributed to the feasibility of sentiment analysis. Twitter also represented all age groups of people and provided a fair representation of gender; therefore, the sentiment analysis of Twitter data became somewhat general sentiments of society (Mandloi et al., 2020).

This study began by asserting the purpose of qualitative research, which was to delve into the intricacies and significance of social phenomena (Pope et al.,2000). It emphasized the qualitative approach to understanding complex social dynamics.

Moving forward, the study outlined a methodological approach: data was collected from websites and Twitter posts, then filtered for misinformation. It described the subsequent data processing steps, including the removal of hashtags and stop-words before engaging in textual analysis. These steps were crucial for preparing the data for further analysis, ensuring that irrelevant or potentially biasing elements were eliminated.

The study concluded by highlighting a gap in the researcher's knowledge: the absence of a specific training set for detecting COVID-19 misinformation suspects through sentiment and topic modeling analysis using Natural Language Processing (NLP). This implied a challenge or limitation in the research process, indicating a need for innovative approaches or adaptations to address this gap. (Hassan et al., 2020).

3.10. Research Ethics

Ethical considerations and concerns in these fields were many and manifold. For one, the culture of Big Data contributed to the emergence of new health issues, tied to the usage and proliferation of digital and algorithmic cultures. Online platforms, such as social media, online web browsing, internet searches, forums interactions, as well as data input through smartphone applications, also provided access to new types of data that were rendered through digital and computerized processes, hence, quantified. New types of data renditions, thus, could be understood and rendered biologically (Ashofteh et al., 2020). Ethical approval and consent to participate were not relevant to this study. All information collected for this study was from the public domain, and the study did not involve any interaction with users. Definable user information was removed from the study results.

CHAPTER 4: Study Findings and Discussion

4.1 Study Findings and Discussion

This chapter presented the results of an analysis of data which had been obtained from Twitter Application Programming Interface and Websites using the NLP analysis. The purpose of this study was to assess the extent of misinformation during infodemics to address the lack of algorithmic accountability.

The information and knowledge gained, thus, helped to clarify questions and problems about algorithmic accountability.

Python language was used to analyze qualitative data. The findings presented were in line with the research aim, objectives, and research questions outlined in Chapter One of the study. The chapter commenced with extracted Tweepy from the CSV files which contained the conversations, dates, full texts, hashtags, and URLs. The researcher used Spyder Python for extracting data from the Twitter API and then used Feedly, which was a news aggregator application for various web browsers and mobile devices running iOS and Android; it was also available as a cloud-based service. The application was used to compile news feeds from a variety of online sources for the researcher to customize and share with others. After filtering for the COVID-19 tweets, there were 21 of 508 records. Appendix C represents a Table that consists of extracted Tweepy from the CSV files which contain the conversations, dates, full texts, hashtags, and URLs.

Extracted Covid-19 misinformation data tweeted on tweeter:

Retweet Counts: Retweets indicated the number of times a tweet had been shared by users on their own timelines. Political tweets that gained significant attention, contained thought-provoking or controversial statements, or were shared by influential political figures or media outlets, tended to have higher retweet counts. Major political events, such as elections, policy announcements, or international conflicts, often generated a considerable number of retweets.

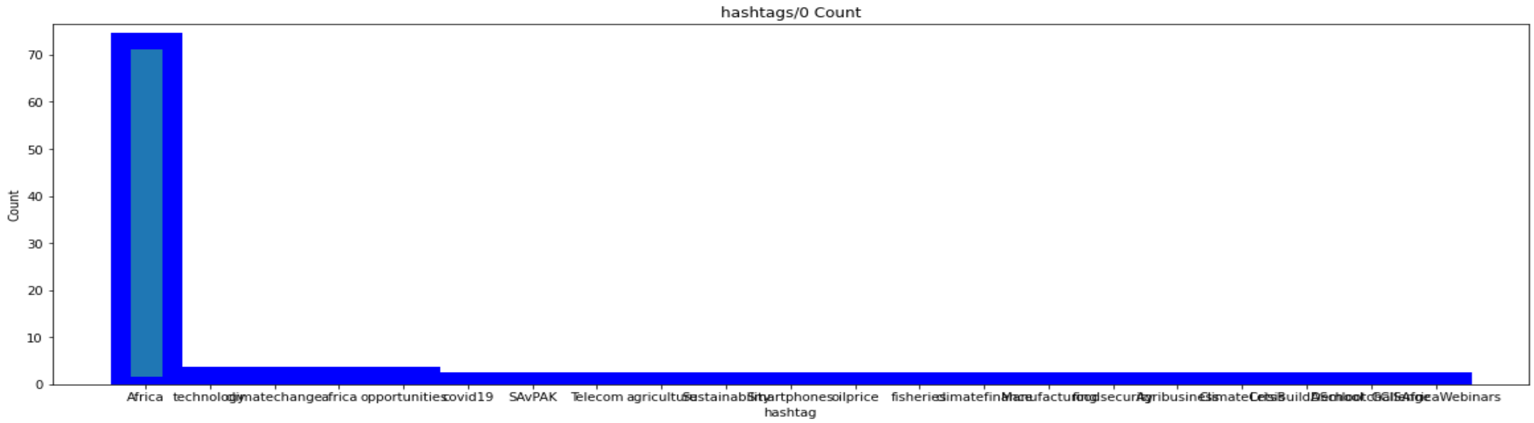
Reply Counts: Replies represented the number of responses or comments a tweet received from other users. Political tweets that prompted discussions, expressed strong opinions, or asked for public input were more likely to receive higher replies. Tweets that were perceived as controversial or that addressed sensitive issues tended to generate more replies as they sparked debates among Twitter users.

Code 4.1: Filtered columns from excel spreadsheet.

Code 4.1 is the extraction of filtered columns and rows using hashtags of COVID-19 tweets that contains misinformation.

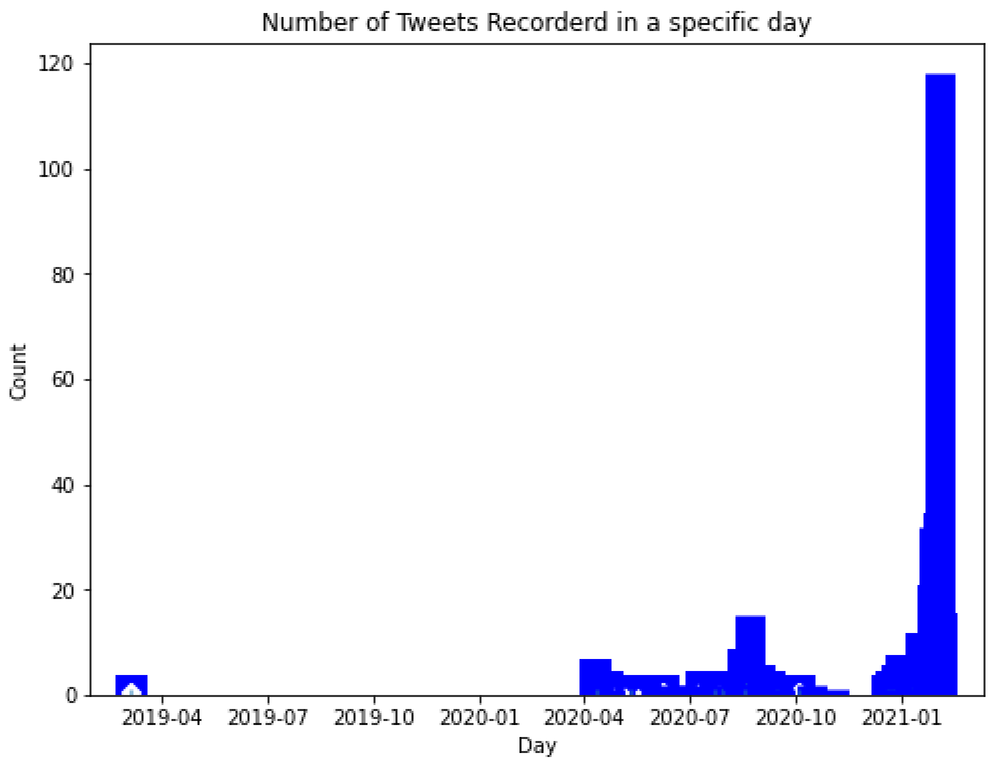
```
import pandas as pd
import seaborn as sb
import matplotlib.pyplot as plt

# filtered columns from the excel sheet
columns = ['id','reply_count','retweet_count','created_at','full_text','hashtags/0','hashtags/1','hashtags/2']

#read data from excel sheet into pandas dataframe
tweets_df = pd.read_excel("/content/drive/MyDrive/Colab Notebooks/Ramaru Tweets/tweets
extracted.xlsx" , usecols = columns)
tweets_df['created_at'] = pd.to_datetime(tweets_df['created_at']).dt.date
#drop rows with empty columns
tweets_df = tweets_df.dropna()

#peak at the excel sheet
tweets_df.head()

#bar-graph for hashtags/0 count
tweet_count = tweets_df['hashtags/0'].value_counts()
fig,ax = plt.subplots(figsize = (20,6))
ax.set_title('hashtags/0 Count')
ax.set_xlabel("hashtag")
ax.set_ylabel("Count")
ax.bar(tweet_count.index,tweet_count.values , edgecolor = '#0000ff' , linewidth = 15 )
```

Graph 4.1: Filtered columns from excel spreadsheet.



Graph 4.1. Extraction of filtered columns and rows using hashtags of COVID-19 tweets that contained misinformation.

The researcher investigated the differential diffusion of all the verified misinformation stories distributed on Twitter from 2019 to 2021. The data comprised 20 conversation tweets, 16 links, and 8 websites in South Africa which were extracted using Python language. The researcher classified news as either ‘true’ or ‘false’ using information from the South African Government and WHO. According to Vosoughi, Roy and Aral (2018), there is a 70% likelihood that, on social media, false, or misinformation will be shared, before factual information (Manene et al., 2023). Misinformation diffused significantly faster and more broadly than the truth in all forms of information; and the effects were more pronounced for false political news than for false news about health organization, science, and financial information. Across two studies with more than 1,700 U.S, In the study with adults recruited online, evidence was presented that people shared false claims about COVID-19, partly because they simply failed to think sufficiently about whether or not the content was accurate when deciding what to share (Pennycook, et al., 2020). From the results, the researcher found that fake news was more novel than true information, which suggested that people were more likely to share novel information. Misinformation inspired fear, confusion, and conflict, whereas true information inspired anticipation, sadness, joy, and trust. Graph 4.2 represented the number of COVID-19 tweets recorded in a specific day, dated from 2019 to 2021.



Graph 4.2: Number of COVID-19 tweets recorded in a specific day from 2019 to 2021.

During the COVID-19 pandemic, politics played a significant and crucial role as governments and politicians addressed the crisis. Measures taken by authorities, political responses, and debates surrounding the pandemic generated a substantial volume of political tweets.

In 2021, numerous political events occurred worldwide, including elections, policy changes, and social movements. For instance, the United States held its presidential inauguration, and there were also significant political developments in other countries. These events likely led to increased discussions on Twitter about politics. As COVID-19 turned into a full-fledged public health crisis, multiple theories regarding the virus' origin took hold on the internet, all with a common theme - the virus was artificially created in a laboratory by a rogue government with an agenda. This misinformation originated from social media accounts and websites with no credible evidence to support their claims. These posts amassed over 20 million engagements, rising each day, and the theories continued to gain traction and following on the internet, despite scientists from multiple nations analyzing the genome of COVID-19 and coming to the decisive conclusion that the virus originated in nature from an animal source. (Mian et al., 2020).

Pałka (2018) argued that machine-learning-powered tools can be used to automatically detect unclear, open-ended, and conditional statements. This was helpful for consumer organizations (often with limited resources) wishing to raise social awareness, rebuke companies, and/or potentially notify the regulators. The researcher argued that exposing the scale of the problem was necessary to increase social discontent, and that this discontent, properly channeled, could be a tool in the quest of having corporations become much more transparent about their data practices.

CHAPTER 5: Conclusions and Recommendations

Conclusions

During coronavirus disease in 2019, the term 'infodemic' was used to depict the abundance of information about COVID-19 on social media that may have overwhelmed users, as well as misinformation about the virus because of the lack of authentication of information posted on social media. Both the WHO and United Nations have warned that infodemics can become a severe threat to healthcare if misinformation on social media is not addressed in a timely manner (Manene et al,.2023).

The analysis suggested that misinformation about COVID-19 comes in many different forms, from many different sources, and makes many different claims. It often rearranges existing or accurate content relatively than formulating it on a large scale, and where it is manipulated, it is edited with simple tools. Given the breadth of the pandemic, independent media and actions by platforms and others played a vital role in addressing virus-related misinformation. Government websites and the WHO can help sort false from true material, and accurate from misleading assertions.

The findings show that much misinformation, directly and indirectly questions the actions, suitability, and legitimacy of public authorities; this includes governments, health authorities, and international organizations that suggest it will be difficult for those institutions to address or correct it directly without running into multiple problems. For instance - *How many people will accept as credible a government trying to mock counter-misinformation that casts that very same government in a negative light?* In contrast, independent fact-checkers can provide authoritative analysis of misinformation while helping platforms identify misleading and problematic content, just as independent news media can report, credibly, on how governments and others are responding to any kind of pandemic.

Social media does not regulate health information posted, so there is no way of knowing if the sources are reliable, whether the information is accurate, and the originator's intent when posting the information (Manene et al., 2023).

This research analysis also found that prominent public figures continue to play a major role in spreading misinformation about COVID-19. While only a small percentage of the individual pieces of misinformation in the sample came from prominent politicians, celebrities, and other public figures, however, these claims often have very high levels of engagement on various social media platforms. The growing willingness of some news media to call out misinformation and lies from prominent politicians can, perhaps, help counter this, similarly to the decision taken by Twitter, Facebook, and YouTube.

The data used does not capture misinformation from outstanding public figures, still, it can also spread extensively through other channels such as TV and newspapers. While fact-checks rarely spread, either as widely or in the same networks (Bounegru et al. 2017) as the misinformation it corrects, it is imperative that trusted fact-checking and media organizations continue to hold outstanding figures to account for assertations they make across all channels so that they can find new ways to distribute and publicize their work.

Describing the view and the nature of COVID-19 misinformation as an infodemics captures the scale, however, the analysis recommends that there is a risk of mis-characterizing the nature of the pandemic. As the results demonstrate, there is extensive diversity in the types of misinformation circulating, the claims made concerning the virus, and the motivations behind their production. Unlike the pandemic itself, there is no single root cause behind the spread of misinformation about the coronavirus, instead, COVID-19 appears to be a contributing factor for very different actors with a scope of diverse motives and objectives to produce a diversity of types of misinformation about many diverse topics. In this sense, misinformation about COVID-19 is as diverse as information about it.

The risk is in not observing the diversity in the view of coronavirus misinformation but is assuming there could be a sole solution to such problems; instead, the findings show that there will be no cure for misinformation about the new coronavirus, as a virus is not constant. Addressing the spread of misinformation about COVID-19 will take constant integration attempts by independent news, media, platform companies, fact-checkers, and public authorities, to help the public understand and maneuver this pandemic and others.

Individuals fully understand and accept what data about them is being collected and how it will be used; the collection and usage occur in a way ensuring that data is not biased (or de-biased) and trained algorithms are deployed in an unfair, discriminatory manner ensuring algorithmic fairness.

REFLECTIONS ON THE RESEARCH OBJECTIVES

The first research objective was - to determine the Influence of Algorithmic Accountability on Infodemics.

It was found that from the literature review, algorithmic accountability influences infodemics because from the social networks there are no rules set on how to manage algorithms.

The second objective was to examine if algorithmic accountability can problematize transparency.

In the literature review, it was briefly shown the complex nature of the models under careful examination, and this links smoothly into the third glitch-informed dimension that problematizes the concept of transparency, as well as the complexity and irreducibility of the algorithmic. Glitches have always been associated with the chaotic and unpredictable aspects of computational analyses. The reality of algorithms today is that even for experts the precise operations and potential ramifications of an assemblage of algorithms remain obscure. Transparency supposes a holistic model of planned codes and instructions that captures all possibilities, but the nature of algorithmic mediation complicates such a vision. Even if an algorithmic model is made entirely transparent, not all of its potential effects and faculties can be inferred from this gesture. The complexity of the algorithmic, along with its autonomous capacities necessarily means that part of its potential remains closed off (Kemper et al., 2019). The rapid turnout of new versions, the exponential complexity of algorithmic planning, and the exclusive autonomy of the algorithmic are all situations that hamper an effective assessment of algorithms and that are not sufficiently remedied by transparency solus. There is, therefore, a need for empirical studies of algorithmic models used in practice and the need to assess the conditions in which measures of transparency actually yield positive effects by fostering a productive relationship with an audience, while also acknowledging the necessary limits of such a relationship. Such research can help us to test guidelines for algorithmic accountability. Only then can we begin to develop guidelines that are sound and fit within existing practices (Kemper et al., 2019).

The third objective was to examine if algorithmic accountability can be realized in practice.

Technological capabilities have advanced, governmental agencies have increasingly delegated important administrative decisions to algorithms. Decisions concerning everything, from hiring to allocation of public benefits, and even law enforcement, often involve the use of algorithms today. These algorithms, however, frequently operate as a black box, making decisions without transparency and making it difficult or impossible to assess the accuracy and fairness of their performance. In contrast to the robust procedures that exist in many administrative contexts to promote the accountability of human decisions in government - from freedom of information laws to adjudicatory due process protections and appeals procedures - algorithms largely operate today in an accountability-free zone (Need et al., 2022).

The last objective was to propose strategies on how COVID-19 and other pandemics' misinformation can be confronted.

Based on the literature review, this study aimed to establish the extent of Misinformation during infodemics to address the lack of algorithmic accountability on confronting the COVID-19 pandemic. To better understand COVID-19 misinformation trends, the study analyzed Twitter Trends, identifying the infodemic labels circulating in South Africa by observing the titles of the most-read articles and websites between January and March 2020. It defined any term, query, hashtag, or phrase feeding misinformation on the internet as an infodemic label. The main terms the researcher found includes - South Africa coronavirus, face masks, symptoms of the coronavirus, and vaccine for coronavirus. The researcher also found searches with - coronavirus conspiracy and coronavirus laboratory - identified as the most dangerous and showing a tendency of racism and xenophobia.

CONTRIBUTIONS OF THE STUDY

This study contributed significantly to the literature on sense-making perspectives on algorithmic accountability. The process of decision-making failed many organizations due to a lack of guidelines. The identified factors in this study will assist management in organizations to address issues related to algorithmic accountability in infodemics and aligned the benefits and capabilities of technologies.

Algorithmic Transparency Requirements:

Strategy: Enact regulations mandating transparency in algorithms used to disseminate information during an infodemic.

Implementation Plan: Require social media platforms and other information-sharing platforms to disclose information about their algorithms, including how content is prioritized, recommended, and moderated. Establish standards for transparent reporting on algorithmic outcomes and impacts during infodemics.

Implications: Increased transparency can help identify and address algorithmic biases, misinformation amplification, and other risks associated with algorithmic decision-making during infodemics.

Real-time Monitoring and Reporting:

Strategy: Implement real-time monitoring and reporting mechanisms to track the spread and impact of misinformation during an infodemic.

Implementation Plan: Develop AI-powered tools and algorithms capable of identifying and flagging misinformation in real-time. Establish partnerships with fact-checking organizations and research institutions to verify and validate information. Create public dashboards or reports to provide real-time updates on misinformation trends and patterns.

Implications: Real-time monitoring and reporting can enable swift responses to misinformation outbreaks, mitigate the spread of false information, and promote public awareness of misinformation risks during infodemics.

Algorithmic Bias Mitigation:

Strategy: Develop strategies to mitigate algorithmic biases in content recommendation and moderation algorithms during infodemics.

Implementation Plan: Implement bias detection and correction algorithms to identify and mitigate biases in algorithmic decision-making. Conduct regular audits and evaluations of algorithmic systems to assess their fairness, accuracy, and effectiveness. Provide training and resources to algorithm developers and content moderators on identifying and addressing biases.

Implications: Mitigating algorithmic biases can help prevent the amplification of misinformation, reduce the spread of false narratives, and promote the dissemination of accurate and trustworthy information during infodemics.

User Empowerment and Digital Literacy:

Strategy: Empower users with tools and resources to identify and critically evaluate information during an infodemic.

Implementation Plan: Develop educational campaigns and initiatives to promote digital literacy, critical thinking, and media literacy skills among users. Provide users with access to fact-checking tools, resources, and guidelines for verifying information. Encourage users to report misinformation and share accurate information with their networks.

Implications: User empowerment and digital literacy initiatives can help individuals make informed decisions, discern credible information from misinformation, and contribute to combating the spread of false narratives during infodemics.

Cross-sector Collaboration and Governance:

Strategy: Foster collaboration among governments, tech companies, civil society organizations, and academia to address algorithmic accountability during infodemics.

Implementation Plan: Establish multi-stakeholder task forces, working groups, or advisory boards to develop and implement policies, guidelines, and best practices for algorithmic accountability. Facilitate information-sharing, data collaboration, and joint initiatives to address misinformation challenges across sectors.

Implications: Cross-sector collaboration and governance can promote coordination, alignment, and collective action in addressing algorithmic accountability during infodemics, enhancing the effectiveness and impact of interventions.

This will be a significant contribution to practice and management as the empirical evidence of this study will be used to guide decision-making.

LIMITATIONS OF THE STUDY

The study focused on misinformation on the coronavirus pandemic on algorithmic accountability during Infodemics. The literature available on the COVID-19 pandemic is mostly available in the White Papers, company websites and government websites. Very little literature has been academically recorded on algorithmic accountability during infodemics in terms of the COVID-19 pandemic. This study tested the Technology-Organization-Environment (TOE) to assess algorithmic accountability during infodemics, hence, this study will contribute significantly to the literature of the sense-making perspective on algorithmic accountability. The process of decision-making fails many organizations due to lack of guidelines, therefore, the identified factors in this study will assist management in organizations to address issues related to algorithmic accountability in infodemics and align the benefits and capabilities of technologies. This will be a significant contribution to practice and management as the empirical evidence of this study will be used to guide decision-making.

Recommendations

Governments and figures in the media should utilize the knowledge of experts, particularly from the Centre for Disease Control and Prevention and WHO, to accurately deliver information in a comprehensible and precise manner, not to incite panic amongst the public. The appearance of the COVID-19 virus has offered an opportunity for the public and medical-health professionals to fight in unity against this common threat. If health bodies appropriately manage, educate, and address people’s concerns, there is an opportunity to bridge the level of distrust that has arisen by anti-scientific movements in recent times (Mian et al,.2020). Social media platforms have taken steps to mitigate the spread of COVID-19 misinformation by implementing policies such as Twitter’s recent ‘COVID-19 Misleading Information Policy’, which prohibits users from using the platform’s services to share false or misleading information about COVID-19. Olan et al., (2022) reported that societal acceptance of information depends on the verification features available on social media platforms. The lack of verification further highlights the need for social media platforms to introduce mechanisms that fact-check and verify the information to reduce the spread of misinformation.

In light of the identified negative outcomes of the discussed COVID-19 social media infodemic and the spread of misinformation on social media, it can be concluded that there is a need for health information management on social media. This study has developed a conceptual framework that can be used to mitigate the sharing of misinformation about COVID-19 and other pandemics on social media platforms during an infodemic. By implementing the set of developed strategies and activities, this framework aims to distribute these activities among stakeholders, to facilitate the management of misinformation on social media; despite the validity and effectiveness of the findings discovered throughout this study, one major limitation of this research is that the data collected is secondary. For these results to be more effective, gathering primary data should be facilitated in future studies and research in this field. Gathering qualitative data could further facilitate innovative applications to the developed framework and possibly allow for the framework’s application in other environments, outside of social media.

The study’s researcher suggests that officials and individuals should consider taking action to combat COVID-19 and other pandemics’ misinformation. These are:

- The dissemination of compelling evidence for the practice should continue.
- Developing standards for patrolling social media, catching and convincingly refuting misinformation before it has the chance to take root.
- Strategically supplying fact-checking personnel with information that debunks the most prevalent misconceptions first, thereby, maximize the use of resources.

- Supplying conservative outlets with more accurate information through the placing of public service announcements, supplying persuasive interview subjects, and promoting websites where consumers can find accurate information.
- Encouraging online subscription news outlets to take down their paywalls around COVID-19 content to make accurate information accessible to a larger audience.

References

Aggarwal, C. C. (2011). Social Network Data Analytics. Social Network Data Analytics, 1–2. https://doi.org/10.1007/978-1-4419-8462-3

Ahinkorah BO, Ameyaw EK, Hagan JE Jr, Seidu A-A and Schack T (2020). Rising Above Misinformation or Fake News in Africa: Another Strategy to Control COVID-19 Spread. Front. Commun. 5:45. doi: 10.3389/fcomm.2020.00045.

Ahinkorah, B. O., Ameyaw, E. K., Hagan, J. E., Seidu, A.-A., & Schack, T. (2020). Rising Above Misinformation or Fake News in Africa: Another Strategy to Control COVID-19 Spread. Frontiers in Communication, 5 (June), 2018–2021. https://doi.org/10.3389/fcomm.2020.00045

Al Dawood, H. M. A. (2021). Bridging the gap between AI systems and society: algorithmic accountability and ethics within small AI organisations in Sweden: Is Sweden ready for AI Ethics.

Algorithms, R. (n.d) Combating fake news during the COVID-19 pandemic presentation outline (CSIR, 2020).

Alhassan, I., Sammon, D., & Daly, M. (2016). Data governance activities: an analysis of literature. Journal of Decision Systems, 25, 64–75. https://doi.org/10.1080/12460125.2016.1187397.

Amini, S., Gerostathopoulos, I., & Prehofer, C. (2017). Big data analytics architecture for real-time traffic control. In: 2017 5th IEEE international conference on Models and Technologies for intelligent transportation systems (MT-ITS), 26 Jun 2017 (pp. 710–715).

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society, 20(3), 973–989. https://doi.org/10.1177/1461444816676645

Ancona, D. (2012). Sense-making: Framing and Acting in the Unknown. The Handbook for Teaching Leadership: Knowing, Doing, and Being, 3–19.

Anshari M, Almunawar MN, Lim SA, Al-Mudimigh A. (2019). Customer relationship management and big data enabled: personalization & customization of services. Applied Computing and Informatics. 2019; 15(2):94–101. doi: 10.1016/j.aci.2018.05.004. [CrossRef] [Google Scholar]

Ashofteh, A., & Bravo, J. M. (2020). A study on the quality of novel coronavirus (COVID-19) official datasets. Statistical Journal of the IAOS, 36(2), 291–301. https://doi.org/10.3233/SJI-200674

Ashofteh, A., & Bravo, J. M. (2020). A study on the quality of novel coronavirus (COVID-19) official datasets. Statistical Journal of the IAOS, 36(2), 291–301. https://doi.org/10.3233/SJI-200674

Bai Y, Yao L, Wei T, Tian F, Jin DY, Chen L, Wang M. 2020 Presumed asymptomatic carrier transmission of COVID-19. JAMA. 2020 doi: 10.1001/jama.2020.2565. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

Bailey, J. (2008). The first step in qualitative data analysis: Transcribing. Family Practice, 25(2), 127–131. https://doi.org/10.1093/fampra/cmn003.

Bangalee, A., & Bangalee, V. (2021, October 26). Fake news and fallacies: Exploring vaccine hesitancy in South Africa. South African Family Practice, 63(1). https://doi.org/10.4102/safp.v63i1.5345

Batrinca, B., Treleaven, P.C., (2015). Social media analytics: a survey of techniques, tools, and platforms. AI & Soc. 30, 89 https://doi.org/10.1007/s00146-014-0549-4.

Beninger, K., Fry, A., Jago, N., Lepps, H., Nass, L., & Silvester, H. (2014). Research using social media; users' views. NatCen Social Research, 20.

Brasoveanu, A., Moodie, M., & Agrawal, R. (2020). Textual evidence for the perfunctoriness of independent medical reviews. CEUR Workshop Proceedings, 2657(1), 1–9. https://doi.org/10.1145/

Brennen, J. S., Simon, F. M., Howard, P. N., & Nielsen, R. K. (2020). Types, Sources, and Claims of COVID-19 Misinformation. Oxford University Press, (April), 1–13.

Britten, N., Jones, R., Murphy, E., & Stacy, R. (1995). Qualitative research methods in general practice and primary care. Family practice, 12(1), 104-114.

Brown, A. D., Colville, I., & Pye, A. (2015). Making Sense of Sensemaking in Organization Studies. Organization Studies, 36(2), 265–277. https://doi.org/10.1177/0170840614559259

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. Public Administration Review, 81(5), 825-836.

Cilliers, Charles. "Zweli Mkhize says '5G causing Covid-19 deaths' is just fake news". The Citizen. Archived from the original on 29 July 2020. Retrieved 1 April 2020

Cilliers, Charles. 2020 "Zweli Mkhize says '5G causing Covid-19 deaths' is just fake news". The Citizen. Archived from the original on 29 July 2020. Retrieved 1 April 2020

Citation: Ahinkorah BO, Ameyaw EK, Hagan JE Jr, Seidu A-A and Schack T (2020) Rising Above Misinformation or Fake News in Africa: Another Strategy to Control COVID-19 Spread. Front. Commun. 5:45. doi 10.3389/comm.2020.00045

COVID-19 Test Kits. (2022, April 15). Biomedical Safety & Standards, 52(7), 53–53. https://doi.org/10.1097/01.bmsas.0000825204.71280.c3

Crampton, J.W., Graham, M., Poorthuis, A., Shelton, T., Stephens, M., Wilson, M.W., Zook, M., 2013. Beyond the geotag situating 'big data' and leveraging the potential of the geoweb. Cartogr. Geogr. Inf. Sci. 40, 130–139. https://doi.org/10.1080/15230406.2013.777137.

Culp, W. C. (2020). Coronavirus Disease 2019. A & A Practice, 14(6), e01218. https://doi.org/10.1213/xaa.0000000000001218

Data governance: A conceptual framework, structured review, and research agenda. (2019, August 10). Data Governance: A Conceptual Framework, Structured Review, and Research Agenda - ScienceDirect. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Data governance: A conceptual framework, structured review, and research agenda. (2019, August 10). Data Governance: A Conceptual Framework, Structured Review, and Research Agenda - ScienceDirect. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Davis, Rebecca. 2020 "Coronavirus Fact Check#1: Viral outbreak: Fake news spreads in SA in tandem with Covid-19". Daily Maverick. Archived from the original on 16 August 2020. Retrieved 6 April 2020

Davis, Rebecca. 2020"Coronavirus Fact Check#1: Viral outbreak: Fake news spreads in SA in tandem with Covid-19". Daily Maverick. Archived from the original on 16 August 2020. Retrieved 6 April 2020

Dennison Himmelfarb, C. R., & Baptiste, D. (2020). Coronavirus Disease (COVID-19). Journal of Cardiovascular Nursing, (Publish Ahead of Print). (August). https://doi.org/10.1097/jcn.0000000000000710

Diakopoulos, N. (2014). Algorithmic accountability reporting: On the investigation of black boxes.

Diakopoulos, N. (2016). Algorithmic accountability reporting: On the investigation of black boxes. Tow Center for Digital Journalism.

Dirk Lindebaum, Mikko Vesa, and Frank den Hond, 2020: Insights From "The Machine Stops" to Better Understand Rational Assumptions in Algorithmic Decision Making and Its Implications for Organizations. AMR, 45, 247–263, https://doi.org/10.5465/amr.2018.0181

Edelman Trust Barometer. (2021). Trust and the Covid-19 Pandemic. https://www.edelman.com/research/trust-2021-spring-update.

Eprs, T., & Parliamentary, E. (2019). Scientific Foresight Unit (STOA), a governance framework for algorithmic accountability and transparency, Brussels: European Union 2019. Retrieved from http://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf.

Fatemeh Torabi Asr and Maite Taboada. 2018. The Data Challenge in Misinformation Detection: Source Reputation vs. Content Veracity. In Proceedings of the First Workshop on Fact Extraction and Verification (FEVER), pages 10–15, Brussels, Belgium. Association for Computational Linguistics

Fleming, N. (2020). Coronavirus misinformation, and how scientists can help to fight it. Nature, 583(7814), 155–156. https://doi.org/10.1038/d41586-020-01834-3

Frenkel, S., Alba, D., and Zhong, R. (2020). Surge of virus misinformation stumps Facebook and Twitter. The New York Times. Retrieved from https://www.nytimes.com/2020/03/08/technology/coronavirus-misinformation-social-media.html.

Frenkel, S., Alba, D., and Zhong, R. (2020). Surge of virus misinformation stumps Facebook and Twitter. The New York Times. Retrieved from https://www.nytimes.com/2020/03/08/technology/coronavirus-misinformation-social-media.html

Fuller, M. (Ed.). (2008). Software studies: A lexicon. Mit Press.

Garritty C, Stevens A, Hamel C, Golfam M, Hutton B, Wolfe D. 2019 Knowledge synthesis in evidence-based medicine. Semin Nucl Med. 2019; 49(2):136 44.https://doi.org/10.1053/j.semnuclmed.2018.11.006 Google Scholar

Gorgoni, G. (2020). Stay Human. The quest for responsibility in the algorithmic society. Journal of Ethics and Legal Technologies, 2(JELT-Volume 2 Issue 1), 31-47.

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science Advances, 5(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586.

Haleem A, Javaid M, Khan IH, Vaishya R. Significant Applications of Big Data in COVID-19 Pandemic. Indian J Orthop. 2020 May 7;54(4):526-528. doi: 10.1007/s43465-020-00129-z. PMID: 32382166; PMCID: PMC7204193.

Haleem Abid. 2020 "Significant Applications of Big Data in COVID-19 Pandemic." Indian Journal of Orthopaedics, vol. 54, 4 1-3. 7 May. 2020, doi: 10.1007/s43465-020-00129-z

Haleem, A., Javaid, M., Khan, I. H., & Vaishya, R. (2020). Significant Applications of Big Data in the COVID-19 Pandemic. Indian journal of orthopedics, 54(4), 1–3. Advanced online publication. https://doi.org/10.1007/s43465-020-00129-z

Haleem, A., Javaid, M., Khan, I. H., & Vaishya, R. (n.d.). Significant Applications of Big Data in the COVID-19 Pandemic. PubMed Central (PMC). https://doi.org/10.1007/s43465-020-00129-z.

Hassan S-U, Saleem A, Soroya SH, 2020. Sentiment analysis of tweets through Altmetrics: A machine learning approach. Journal of Information Science. June 2020.

Hobbs, R., & Mihailidis, P. (2019). News literacy: Understanding digital age journalism. SAGE Publications.

Hollowood, E., Mostrous, A. (2020). 'Fake news in the time of C-19'. Tortoise. (Accessed Mar. 2020). https:// members.tortoisemedia.com/2020/03/23/the-Infodemics-fake-news-coronavirus/content.html

Hossain, T. (2021). COVID Lies: Detecting COVID-19 misinformation on social media. University of California, Irvine.

https://datasociety.net/library/algorithmic-accountability-a-primer/ RESOURCE · APRIL 18, 2018, ALGORITHMIC ACCOUNTABILITY: A PRIMER.

https://www.jonesday.com/en/insights/2019/06/proposed-algorithmic-accountability-act

https://www.medrxiv.org/content/early/recent

IIOM. (2020). Misinformation, fake news and the instrumentalization of the COVID-19 pandemic. IOM Social Media Campaign in Mexico 20 April 2020.

J. Han, M. Kamber, and J. Pei. 2011 Data Mining: Concepts and Techniques. Morgan Kaufmann, San Francisco, 2011.

Jin, Y., Li, G., & Zhang, C. (2015). Study of data mining in social network analysis. Journal of Information and Computational Science, 12(16), 5947–5955. https://doi.org/10.12733/jics20106851

Johnson, N. F., Velasquez, N., Jha, O. K., Niyazi, H., Leahy, R., Restrepo, N. J., ... & Wuchty, S. (2020). Covid-19 infodemic reveals new tipping point epidemiology and a revised $ R $ formula. arXiv preprint arXiv:2008.08513.

K. Priyanka and N. Kulennavar, 2014 "A survey on big data analytics in health care," International Journal of Computer Science and Information Technologies, vol. 5, no. 4, pp. 5865–5868, 2014).

Kaplan, A.M., Haenlein, M., (2010). Users of the world, unite! The challenges and opportunities of social media. Bus. Horiz. 53, 59–68. https://doi.org/10.1016/j.bushor. 2009.09.003.

Kelly SE, Moher D, Clifford TJ. 2016 Quality of conduct and reporting in rapid reviews: an exploration of compliance with PRISMA and AMSTAR guidelines. Syst Rev. 2016; 5:79. https://doi.org/10.1186/s13643-016-0258-9.

Kemper, J., & Kolkman, D. (2019). Transparent to whom? No algorithmic accountability without a critical audience a critical audience. 4462. https://doi.org/10.1080/1369118X.2018.1477967

Kivunja, C., & Kuyini, A. B. (2017). Understanding and Applying Research Paradigms in Educational Contexts. International Journal of Higher Education, 6(5), 26. https://doi.org/10.5430/ijhe.v6n5p26

Kraska, P. B., Brent, J. J., & Neuman, W. L. (2020). Qualitative Research and Analysis. Criminal Justice and Criminology Research Methods, 333–362. https://doi.org/10.4324/9780429026256-11

L. Mandloi and R. Patel. 2020, "Twitter Sentiments Analysis Using Machine Learning Methods,". 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-5, doi: 10.1109/INCET49848.2020.9154183).

Lange, P.G., (2007). Publicly private and privately public: social networking on YouTube. J. Comput. Commun. 13, 361–380.

Liu, C., & Graham, R. (2020). Making Sense of Algorithms: Relational Perception of Contact Tracing and Risk Assessment during COVID-19. SSRN Electronic Journal, 1–26. https://doi.org/10.2139/ssrn.3687843

Liu, Q., Liu, W., Sha, D., Kumar, S., Chang, E., Arora, V., Yang, C. (2020). An Environmental Data Collection for COVID-19 Pandemic Research. Data, 5(3), 68. https://doi.org/10.3390/data5030068n.

Lomborg, S., Bechmann, A., (2014). Using APIs for data collection on social media. Inf. Soc. 30, 256–265. https://doi.org/10.1080/01972243.2014.915276.

M. Lippi, G. Contissa, A. Jablonowska, F. Lagioia, H.-W. Micklitz, P. Palka, G. Sartor, and P. Torroni. 2020, The force awakens Artificial intelligence for consumer law. Journal of Artificial Intelligence Research, 67:169–190, 2020.

Machmud, M., Irawan, B., Karinda, K., Susilo, J., & Salahudin, S. (2021). Analysis of the intensity of communication and coordination of government officials on twitter social media during the Covid-19 handling in Indonesia. Academic Journal of Interdisciplinary Studies, 10(3), 319-334.

MacLaren, G., Fisher, D., & Brodie, D. (2020). Preparing for the most critically ill patients with COVID-19: the potential role of extracorporeal membrane oxygenation. Jama, 323(13), 1245-1246.

Mahdi Jalili, Matjaž Perc, 2017**.** Information cascades in complex networks., Journal of Complex Networks, Volume 5, Issue 5, October 2017, Pages 665–693, https://doi.org/10.1093/comnet/cnx019.

Manene, S., Hove, C., & Cilliers, L. (2023). Mitigating misinformation about the COVID-19 infodemic on social media: A conceptual framework. Jàmbá Journal of Disaster Risk Studies, 15(1), 1–8. https://doi.org/10.4102/jamba.v15i1.1416.

Mian, A., & Khan, S. (2020). Coronavirus: The spread of misinformation. BMC Medicine, 18(1), 18–19. https://doi.org/10.1186/s12916-020-01556-3

Mian, A., and Khan, S. (2020). Coronavirus: the spread of misinformation. BMC Med. 18, 1–2. doi: 10.1186/s12916-020-01556-3

Mining, T. (2018). Text Mining 101: A Stepwise Introduction to Topic Modeling using Latent Semantic Analysis (using Python). 1–12.

Modjarrad K, Moorthy VS, Millett P, Gsell P-S, Roth C, Kieny M-P. 2016 Developing global norms for sharing data and results during public health emergencies. PLoS Med. 2016; 13(1): e1001935. https://doi.org/10.1371/journal.pmed.1001935.

Nandi, G., & Das, a. (2013). A Survey on Using Data Mining Techniques for Online Social Network Analysis. International Journal of Computer Science Issues, 10(6), 162–167.

http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=16940784&AN=93404019&h=MWYzsNVeP2Q8klAFiWFHW3PUpgJLZxRIpB1jfSK4qJfBbaMUEp4nY/oJdYPRHc4xHL0dBYfuGxhZsmiP7ToLBg==&crl=c.

Need, T., Approach, N., When, A., Functions, G., Performed, A., By, A., Freedom, M., & Clinic, I. A. (2022). A LGORITHMIC ACCOUNTABILITY: Media Freedom & Information Access Clinic, Yale Law School.

Nick Seaver. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. Big Data & Society 4, 2 (2017), 205395171773810. https: //doi.org/10.1177/2053951717738104.

Oxford Analytica (2020a). Misinformation Will Undermine Coronavirus Responses. Emerald Expert Briefings (Emerald Insights).

Oxford Analytica (2020a). Misinformation Will Undermine Coronavirus Responses. Emerald Expert Briefings (Emerald Insights).

Pagoto S, Waring ME, Xu R. 2019. A call for a public health agenda for social media research. J Med Internet Res. 2019; 21(12): e16661)

Pałka, P., & Yousefzadeh, R. (2018). Social Fairness, Accountability and Transparency of the Data Economy: Using Machine Learning to Combat the Emptiness of Privacy Policies. http://laurent.risser.free.fr/TMP_SHARE/ICML2020 LML/Using_Machine_Learning_to_Combat_the_Emptiness_of_Privacy_Policies_Acc essed 19_June_2020.pdf.

Paris, B., Donovan, J. 2019. Deepfakes and Cheap Fakes: The Manipulation of Audio and Visual Evidence. Data & Society. (Accessed Mar. 2020). https://datasociety. net/wp-content/uploads/2019/09/DS Deepfakes_ Cheap_FakesFinal-1-1.pdf

Pasquale, F. (2011). Restoring transparency to automated authority. J. on Telecomm. & High Tech. L., 9, 235.

Pasquale, Frank. The black box society: The secret algorithms that control money and information. Harvard University Press, 2015. https://doi.org/10.4159/harvard.9780674736061

Peng, L., Yang, W., Zhang, D., Zhuge, C., & Hong, L. (2020). Epidemic analysis of COVID-19 in China by dynamical modeling. arXiv preprint arXiv:2002.06563.

Pennycook, G., Mcphetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 Misinformation on social media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. https://doi.org/10.1177/0956797620939054.

Pham, Q.-V., Nguyen, D. C., Huynh-The, T., Hwang, W.-J., & Pathirana, P. N. (2020). Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts. IEEE Access, 8(April), 130820–130839. https://doi.org/10.1109/access.2020.3009328

Philips, W., Milner, R. 2017. The Ambivalent Internet: Mischief, Oddity, and Antagonism Online. Cambridge, UK: Polity Press.

Pope C, Mays N. 2000. Qualitative Research in Health Care., (2nd edn.) London: BMJ Books.

Powles, J., & Nissenbaum, H. (2018). The seductive diversion of 'solving 'bias in artificial intelligence. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2133), 20180080.

Public Administration Review, 81(5), 825–836. https://doi.org/10.1111/puar.13293.

PubMed Abstract | CrossRef Full Text | Google Scholar.

Pushpam, C. A., & Jayanthi, J. G. (2017). Overview of Data Mining in social media. International Journal of Computer Sciences and Engineering, 5(11), 147–157. https://doi.org/10.26438/ijcse/v5i11.147157

Rahman, M. S. (2016). The Advantages and Disadvantages of Using Qualitative and Quantitative Approaches and Methods in Language "Testing and Assessment" Research: A Literature Review. Journal of Education and Learning, 6(1), 102. https://doi.org/10.5539/jel.v6n1p102

Raymond, A. H., Young, E. A. S., & Shackelford, S. J. (2017). Building a better HAL 9000: algorithms, the market, and the need to prevent the engraining of bias. Nw. J. Tech. & Intell. Prop., 15, 215.

Runjic E, Behmen D, Pieper D,2019 Following Cochrane review protocols to completion 10 years later: a retrospective cohort study and author survey. J Clin Epidemiol. 2019; 111:41–8. https://doi.org/10.1016/j.jclinepi.2019.03.006.

Runjic E, Rombey T, Pieper D, Puljak L. Half of systematic reviews about pain registered in PROSPERO were not published and the majority had inaccurate status. J Clin Epidemiol. 2019; 116:114–21. https://doi.org/10.1016/ j. jclinepi.2019.08.010

Russonello, G. (2020). Afraid of coronavirus? That might say something about your politics. The New York Times. Retrieved from https://www.nytimes.com/2020/03/13/us/politics/coronavirus-trump-polling.html.

Russonello, G. (2020). Afraid of coronavirus? That might say something about your politics. The New York Times. Retrieved from https://www.nytimes.com/2020/03/13/us/politics/coronavirus-trump-polling.html.

Saúde, O. P.-A. Da. (2020). Understanding the Infodemics and Misinformation in the fight against COVID-19 What is the Infodemics**.** World Health Organization, Pan American Health Organization. (2020). Understanding the infodemic and misinformation in the fight against COVID-19 | DIGITAL TRANSFORMATION TOOLKIT. Pan American Health Organization, 2. https://iris.paho.org/bitstream/handle/10665.2/52052/Factsheet-infodemic_eng.pdf.

Schildt, H., Mantere, S., & Cornelissen, J. (2020). Power in Sensemaking Processes. Organization Studies, 41(2), 241–265. https://doi.org/10.1177/0170840619847718

Schillinger, D., Chittamuru, D., & Ramírez, A. S. (2020). From “Infodemics” to Health Promotion: A Novel Framework for the Role of social media in Public Health. American Journal of Public Health, 110(9), 1393–1396. https://doi.org/10.2105/AJPH.2020.305746.

Scott, M. 2020. ‘Facebook’s Private Groups are Abuzz with Coronavirus Fake News.’ Politico. (Accessed Mar. 2020). https://www.politico.eu/article/ Facebook-misinformation-fake-news-coronavirus- covid19/

Sell, T. K., Hosangadi, D., & Trotochaud, M. (2020). Call for Special Feature Papers: Infodemiology and Infodemics. Health Security, 18(5), 345–346. https://doi.org/10.1089/hs.2020.29001.cfp).

Shah, Saloni & Mulahuwaish, Aos & Ghafoor, Kayhan & Maghdid, Halgurd. (2022). Prediction of global spread of COVID-19 pandemic: a review and research challenges. Artificial Intelligence Review. , 55. 1-22. 10.1007/s10462-021-09988-w.

Shorten, A., & Smith, J. (2017). Mixed methods research: expanding the evidence base. Journal of Practical Studies in Education ,20(3), 74–75.

Sigfrid L, Moore C, Salam AP. 2019. Rapid research needs appraisal methodology to identify evidence gaps to inform clinical research priorities in response to outbreak results from the Lassa fever pilot. BMC Med. 2019; 17(1):107. https://doi.org/10.1186/s12916-019-1338-1.

Sloan, S. M. (2020). Behind the 'Curve': COVID-19, Infodemics, and Oral History. The Oral History Review, 47(2), 193–202. https://doi.org/10.1080/00940798.2020.1798256

Smith, G., & Rustagi, I. (2020). The Problem With COVID-19 Artificial Intelligence Solutions and How to Fix Them. Stanford Social Innovation Review. https://doi.org/10.48558/NEME-4M86

Smith, J. and Nair, R., 2005. Virtual machines: versatile platforms for systems and processes. Elsevier.

Southwell BG, Niederdeppe J, Cappella JN,2019. Misinformation as a misunderstood challenge to public health. Am J Prev Med. 2019; 57 (2): 282–285).

Stoica, A. A. (2020, May). Algorithmic Fairness for Networked Algorithms. In Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (pp. 2214-2216).

Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. Science, 361(6404), 751-752.

Template, 2023: COVID-19 pandemic data - Wikipedia. (2023, May 21). Template: COVID-19 Pandemic Data - Wikipedia. http://en.wikipedia.org/wiki/Template:COVID-19_pandemic_data.

Toivanen, P., Huhtamäki, J., Valaskivi, K., & Tikka, M. (2020). English abstracts Topic modeling in studying hybrid media events and circulation of meanings A transversal algorithmic literacy education model. 43.

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., Tenkanen, H., & Di Minin, E. (2019). Social media data for conservation science: A methodological overview. Biological Conservation, 233; (January), 298 315. https://doi.org/10.1016/j.biocon.2019.01.023.

Unceta, I., Nin, J., & Pujol, O. (2020). Risk mitigation in algorithmic accountability: The role of machine learning copies. PLoS ONE, 15; (11 November), 18–20. https://doi.org/10.1371/journal.pone.0241286

Vedder, A., & Naudts, L. (2017). Accountability for the use of algorithms in a big data environment. International Review of Law, Computers, and Technology, 31(2), 206–224. https://doi.org/10.1080/13600869.2017.1298547

Vijay Khatri and Carol V. Brown. (2010). Designing data governance. Commun. ACM 53, 1 (January 2010), 148–152. https://doi.org/10.1145/1629175.1629210.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. Science, 359(6380), 1146-1151. https://doi.org/10.1126/science.aap9559

Vraga, E., Bode, V. (2020). ‘Defining Misinformation and Understanding its Bounded Nature: Using Expertise and Evidence for Describing Misinformation.’ Political Communication 37(1): 136-144. https://doi. org/10.1080/10584609.2020.1716500.

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explain automated decision-making does not exist in the general data protection regulation. International Data Privacy Law, 7(2), 76-99.

Wang CJ, Ng CY, Brook RH. 2020. Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. JAMA. 2020 doi: 10.1001/jama.2020.3151. [PubMed] [CrossRef] [Google Scholar]

Wardle, C. 2017. ‘Fake news.: Iit’s complicated.’ UK: First Draft. (Accessed Mar. 2020). https:// firstdraftnews.org/latest/fake-news-complicated/

Wardle, C. 2019. ‘First Draft’s Essential Guide to Understanding Information Disorder’. UK: First Draft. (Accessed Mar. 2020). https:// firstdraftnews.org/wp-content/uploads/2019/10/ Information_Disorder_Digital_AW.pdf? x76701

Weick, K. E. (1995). Sensemaking in organizations. Thousand Oaks, CA: Sage.

WHO. (2020). Novel Coronavirus (2019-nCoV) Situation Report - 1. WHO Bulletin, (JANUARY), 17.

Wieringa, M. (2020). What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability. FAT* 2020 Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 1–18. https://doi.org/10.1145/3351095.3372833.

Wieringa, M. (2020, January). What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability. In Proceedings of the 2020 Conference on Fairness, accountability, and Transparency (pp. 1-18).

Wolkewitz M, Puljak L. 2020. Methodological challenges of analysing COVID-19 data during the pandemic. BMC Med Res Methodol. 2020 Apr 14;20(1):81. doi: 10.1186/s12874-020-00972-6. PMID: 32290816; PMCID: PMC7154065.

Wolkewitz, M., & Puljak, L. (2020). Methodological challenges of analyzing COVID-19 data during the pandemic. BMC medical research methodology. 6, 4–7.

Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, Wang Q, Xu Y, Li M, Li X, Zheng M, Chen L, Li H. 2020. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharmaceutica Sinica B. 2020 doi: 10.1016/j.apsb.2020.02.008. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

Xia W, Sanyi T, Yong C, Xiaomei F, Yanni X, Zongben X. 2020**.** When will the resumption of work in Wuhan and its surrounding areas during the COVID-19 epidemic? A data-driven network modeling analysis. SCIENTIA SINICA Mathematica. 2020 doi: 10.1360/SSM-2020-0037. [CrossRef] [Google Scholar]

Zhao, L., Alhoshan, W., Ferrari, A., Letsholo, K. J., Ajagbe, M. A., Chioasca, E. V., & Batista-Navarro, R. T. (2020). Natural Language Processing (NLP) for requirements engineering: A systematic mapping study. ArXiv, v. ACM Computing Surveys (CSUR).

Ziewitz, M.2016**.** “Governing algorithms: Myth, mess, and methods.' Science, Technology, & Human Values 41, no. 1 (2016): 3-16.

APPENDICES

Appendix A

ETHICS APPROVAL CERTIFICATE

RESEARCH AND INNOVATION
OFFICE OF THE DIRECTOR

NAME OF RESEARCHER/INVESTIGATOR:
Mr RR Ramaru

STUDENT NO:
11573950

PROJECT TITLE: A Sense making Perspective of Algorithmic Accountability during Infodemics.

ETHICAL CLEARENCE NO: FMCL/23/BIS/05/0601

SUPERVISORS/ CO-RESEARCHERS/ CO-INVESTIGATORS

NAME	INSTITUTION & DEPARTMENT	ROLE
Dr. W Munyoka	UNIVEN, Business Information System	Supervisor
Prof. A Kadyamatimba	UNIVEN, Business Information System	Co-Supervisor
Mr RR Ramaru	UNIVEN, Business Information System	Investigator – Student

Type: **Master's Research**
Risk: **Straightforward research without ethical problems (Category 2)**
Approval Period: **May 2023 – May 2024**

The Research Ethics Social Sciences Committee (RESSC) hereby approves your project as indicated above.

General Conditions

While this ethics approval is subject to all declarations, undertakings and agreements incorporated and signed in the application form, please note the following.

- The project leader (principal investigator) must report in the prescribed format to the REC:
  - Annually (or as otherwise requested) on the progress of the project, and upon completion of the project.
  - Within 48hrs in case of any adverse event (or any matter that interrupts sound ethical principles) during the course of the project.
  - Annually a number of projects may be randomly selected for an external audit.
- The approval applies strictly to the protocol as stipulated in the application form. Would any changes to the protocol be deemed necessary during the course of the project, the project leader must apply for approval of these changes at the REC. Would there be deviated from the project protocol without the necessary approval of such changes, the ethics approval is immediately and automatically forfeited.
- The date of approval indicates the first date that the project may be started. Would the project have to continue after the expiry date; a new application must be made to the REC and new approval received before or on the expiry date.
- In the interest of ethical responsibility, the REC retains the right to:
  - Request access to any information or data at any time during the course or after completion of the project,
  - To ask further questions; Seek additional information; Require further modification or monitor the conduct of your research or the informed consent process.
  - withdraw or postpone approval if:
  - Any unethical principles or practices of the project are revealed or suspected.
  - It becomes apparent that any relevant information was withheld from the REC or that information has been false or misrepresented.
  - The required annual report and reporting of adverse events was not done timely and accurately,
  - New institutional rules, national legislation or international conventions A it necessary

ISSUED BY:
UNIVERSITY OF VENDA, RESEARCH ETHICS COMMITTEE
Date Considered: May 2023

Name of the RESSC Chairperson of the Committee: Prof TS Mashau

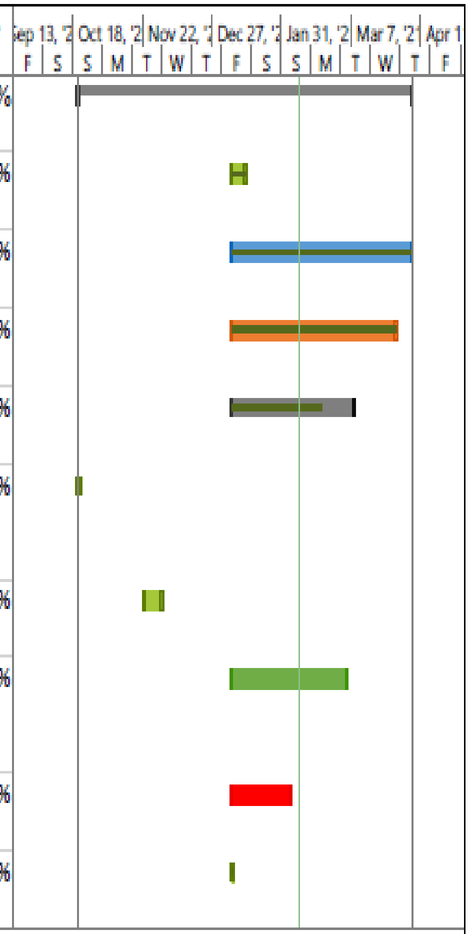
Signature PTMashau

UNIVERSITY OF VENDA
OFFICE OF THE DIRECTOR
RESEARCH AND INNOVATION
2023 -06- 01
Private Bag X5050
Thohoyandou 0950

Appendix B

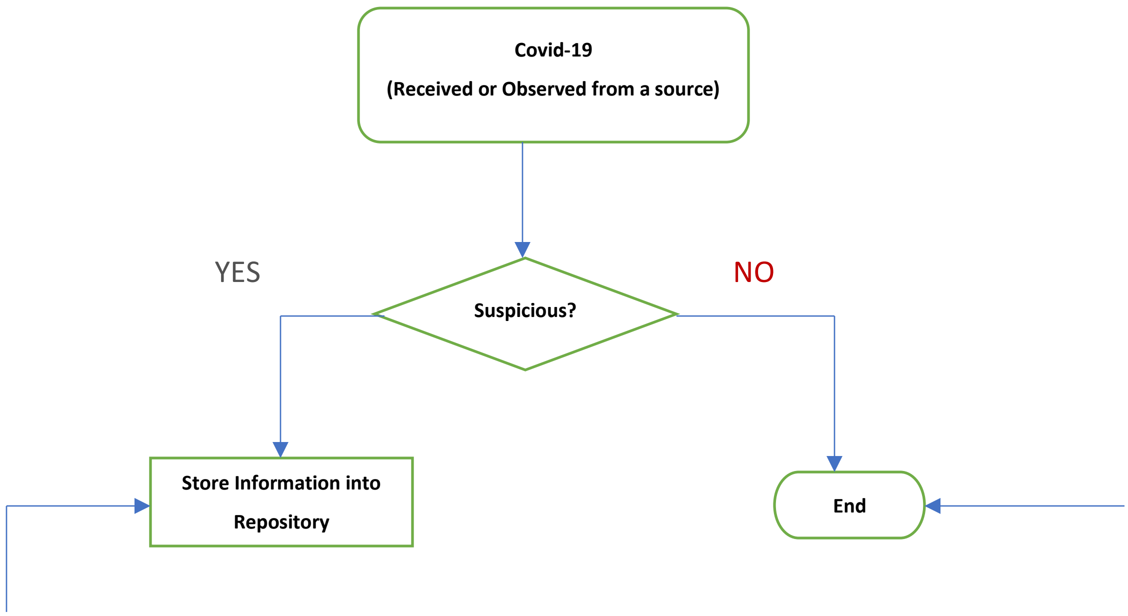
GANTT CHART

ID	Task Name	Start	Finish	Late Start	Late Finish	Free Slack	% Complete
0	A Sense Making Persp	ed 10/21/20	Tue 4/6/21	d 10/21/20	Tue 4/6/21	0 days	66%
1	1 Research Topic Description	Wed 1/6/21	Tue 1/12/21	Wed 1/6/21	Tue 1/12/21	0 days	100%
2	2 Research Proposal Chapter 1	Wed 1/6/21	Tue 4/6/21	Wed 1/6/21	Tue 4/6/21	0 days	100%
3	3 Chapter 2 Literature Review	Wed 1/6/21	Mon 3/29/21	Wed 1/6/21	Mon 3/29/21	0 days	100%
4	4 Chapter 3 Research	Wed 1/6/21	Mon 3/8/21	Thu 2/4/21	Tue 4/6/21	21 days	75%
5	5 Research Proposal departmental Preseantation	Wed 10/21/20	Wed 10/21/20	Tue 4/6/21	Tue 4/6/21	119 days	0%
6	6 Research English Edit	Mon 11/23/20	Tue 12/1/20	Mon 3/29/21	Tue 4/6/21	90 days	0%
7	7 Chapter 4 Study Findings and Discussion	Wed 1/6/21	Thu 3/4/21	Mon 2/8/21	Tue 4/6/21	23 days	0%
8	8 Chapter 5 Conclutions and	Wed 1/6/21	Thu 2/4/21	Mon 3/8/21	Tue 4/6/21	43 days	0%
9	9 References	Wed 1/6/21	Wed 1/6/21	Tue 4/6/21	Tue 4/6/21	64 days	75%



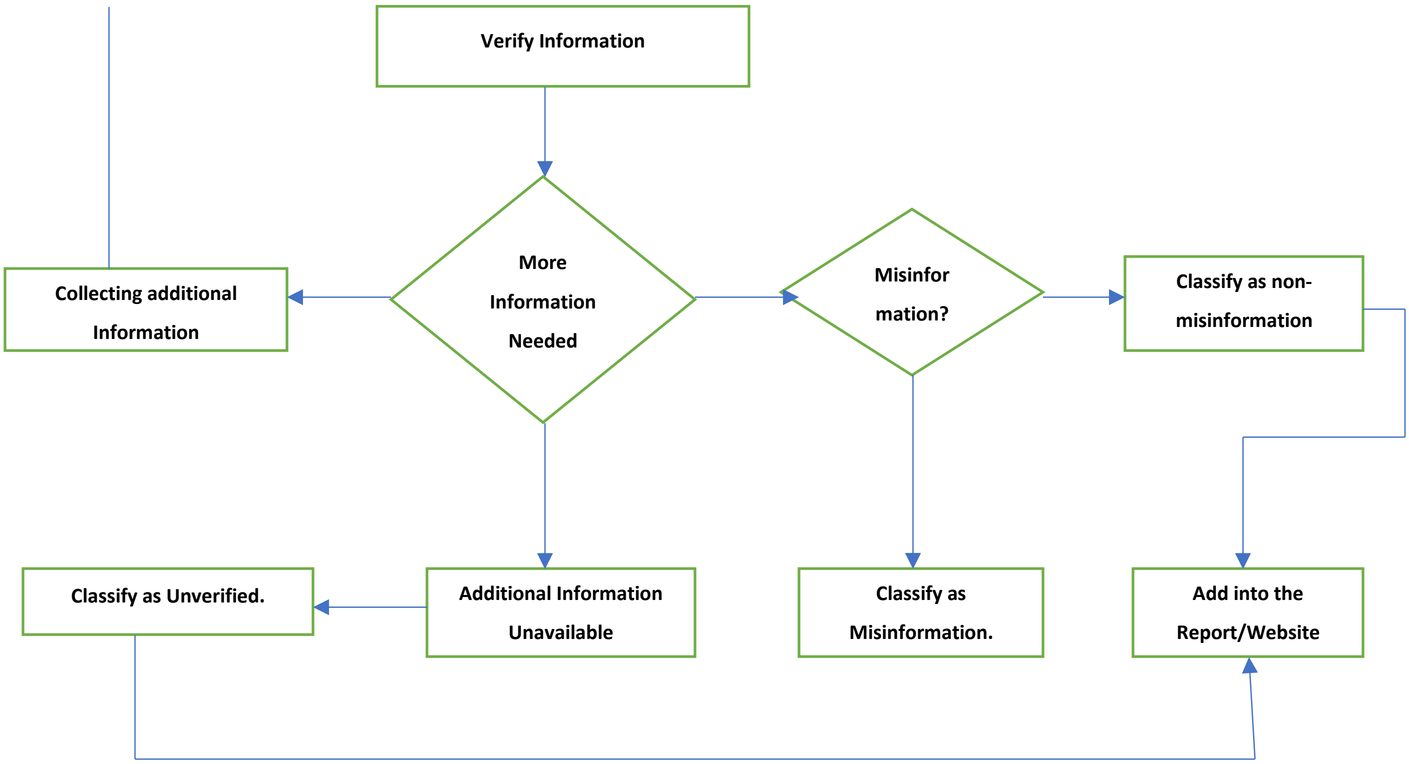
Appendix C

Collecting misinformation



- The process of identifying and collecting misinformation
- Categorizing the misinformation accordingly

Verifying misinformation



- The process of verifying misinformation
- Use of tools and existing fact-checking sources
- E.G. Twitter, WHO, Government website and other fact-checking website

Appendix D

The Table consist of extracted Tweepy from the CSV files which contain the conversations, dates, full texts, hashtags, and URLs. The researcher used Spyder python for extracting data from the Twitter API and then used Feedly, which is a news aggregator application, for various web browsers and mobile devices running iOS and Android. It is also available as a cloud-based service. It was used to compile news feeds from a variety of online sources for the researcher to customize and share with others. After filtering for the Covid-19 Tweets there were 21 of 508 records.

conversation	created at	full text	hashtags/0	hashtags/1	URL	URLs/0/display URL	URLs/0/expanded URL	URLs/0/URL
1.35E+18	1/22/2021 4:50	Global Covid-19 Daily Data Update #COVID19 #Data https://t.co/1NQ9w9hyz1	COVID19	Data	https://twitter.com/SusaAfrica1/status/1352448456443486212	susafrica.com/2020/11/14/glo...	https://susafrica.com/2020/11/14/global-covid-19-daily-data-cases-deaths-recovered/	https://t.co/1NQ9w9hyz1

1.35E+18	1/6/2021 10:40	Global Covid-19 Daily Data Update #COVID19 #data https://t.co/1NQ9w9hyz1	COVID19	data	https://twitter.com/SusaAfrica1/status/1346738368906993664	susafrica.com/2020/11/14/glo...	https://susafrica.com/2020/11/14/global-covid-19-daily-data-cases-deaths-recovered/	https://t.co/1NQ9w9hyz1
1.34E+18	12/27/2020 5:57	Global Covid-19 Daily Data Update #COVID19 #data https://t.co/	COVID19	data	https://twitter.com/SusaAfrica1/status/1343043359179419648	susafrica.com/2020/11/14/glo...	https://susafrica.com/2020/11/14/global-covid-19-daily-data-cases-deaths-recovered/	https://t.co/1NQ9w9hyz1

		1NQ9w9hy z1						

1.34E +18	12/23/ 2020 11:59	Daily Briefing: UN Food Assistance Programme Hit As COVID-19 Dries Funding in Uganda #foodsecurity #COVID19 #Refugees #Uganda #SouthSudan #DRC #Burundi #SDGs https://t.co/VLJnk3K6Iu	foodsecurity	COVID19	https://twitter.com/SusaAfrica1/status/1341684864370794496	susafrica.com/2020/12/23/dai…	https://susafrica.com/2020/12/23/daily-briefing-un-food-assistance-programme-hit-as-covid-19-dries-funding-in-uganda/	https://t.co/VLJnk3K6Iu

1.34E +18	12/25/ 2020 10:51	Global Covid-19 Daily Data Update #COVID19 #data https://t.co/ 1NQ9w9hy z1	COVID1 9	data	https://twitter.com/SusaAfri ca1/status/13423924329950 16707	susafrica.com /2020/11/14/ glo...	https://susafrica.com/2020/11 /14/global-covid-19-daily- data-cases-deaths-recovered/	https://t.co /1NQ9w9h yz1
1.36E +18	2/3/20 21 15:53	Namibia: Covid-19 Accelerates Commercial Farming Dreams https://t.co/ KOHbKGXw n4 #newspape rs #feedly	newspa pers	Feedly	https://twitter.com/anamafa lda1992/status/1356964036 945002496	allafrica.com/ stories/2021 02...	https://allafrica.com/stories/2 02102030687.html	https://t.co /KOHbKGX wn4

1.36E +18	2/3/2021 15:13	COVID-19: Take vaccine or die, FG warns Nigerians https://t.co/mzmZy1ihK8 #newspapers #feedly	newspapers	Feedly	https://twitter.com/anamafalda1992/status/1356953987854065665	legit.ng/1401074-covid-...	https://www.legit.ng/1401074-covid-19-take-vaccine-die-fg-warns-nigerians.html	https://t.co/mzmZy1ihK8
1.36E +18	2/3/2021 15:14	BREAKING: Governor Ortom recovers from COVID-19, returns to office https://t.co/w6i0IgdaEa #newspapers #feedly	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356954262996213770	legit.ng/1401100-govern...	https://www.legit.ng/1401100-governor-ortom-recovers-covid-19-returns-office.html	https://t.co/w6i0IgdaEa

1.36E +18	2/3/2021 15:04	South Africa to Begin Testing COVID-19 Vaccines Before Launching Program https://t.co/zMoUkF4sB3 #newspapers #feedly	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356951724980908032	voanews.com/covid-19-pande...	https://www.voanews.com/covid-19-pandemic/south-africa-begin-testing-covid-19-vaccines-launching-program	https://t.co/zMoUkF4sB3
1.36E +18	2/3/2021 15:01	How a garage filled with dusty old machines is helping South Africa combat	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356951005037006851	rss.cnn.com/~r/rss/cnn_top...	http://rss.cnn.com/~r/rss/cnn_topstories/~3/TwJGVRG8wyo/index.html	https://t.co/jXpyRjweh8

		Covid-19 https://t.co/jXpyRjweh8 #newspapers #feedly						
1.36E+18	2/3/2021 15:02	Africa: Continent Nears 3.6 Million Confirmed Cases of Covid-19 https://t.co/C7M74bpHBd #newspapers #feedly	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356951189183676417	allafrica.com/stories/202102...	https://allafrica.com/stories/202102030278.html	https://t.co/C7M74bpHBd
1.36E+18	2/3/2021 15:02	Fear in NYSC camps as senior official dies of COVID-	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356951157214752770	legit.ng/1401036-covid-...	https://www.legit.ng/1401036-covid-19-fear-nysc-camps-senior-official-dies-infection.html	https://t.co/VIGdDYwncP

		19 https://t.co/VIGdDYwncP #newspapers #feedly						
1.36E+18	2/3/2021 10:53	Nigeria: 22 LGAs Account for 95% of Recent Covid-19 Cases (Full List) https://t.co/xx433B85xA #newspapers #feedly	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356888651364466693	allafrica.com/stories/202102...	https://allafrica.com/stories/202102030182.html	https://t.co/xx433B85xA
1.36E+18	2/3/2021 15:01	Africa: Covid-19 Death Toll Exceeds 92,000	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356950956970291200	allafrica.com/stories/202102...	https://allafrica.com/stories/202102030235.html	https://t.co/3dTfWlEW0w

		Across Continent https://t.co/3dTfWlEW0w #newspapers #feedly						
1.36E+18	2/3/2021 10:55	How a garage filled with dusty old machines is helping South Africa combat Covid-19 https://t.co/YFYdesQKIu #newspapers #feedly	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356888949831122944	rss.cnn.com/~r/rss/cnn_lat…	http://rss.cnn.com/~r/rss/cnn_latest/~3/TwJGVRG8wyo/index.html	https://t.co/YFYdesQKIu

1.36E +18	2/3/2021 15:04	Britain Identifies 105 Cases of South African COVID-19 Variant https://t.co/HnoVeAa2r7 #newspapers #feedly	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356951703350820865	voanews.com/covid-19-pande...	https://www.voanews.com/covid-19-pandemic/britain-identifies-105-cases-south-african-covid-19-variant	https://t.co/HnoVeAa2r7
1.36E +18	2/3/2021 10:55	Nigeria: Kano State Shuts Hospital for Attending to Severe Covid-19 Cases https://t.co/mPkatPR8lq	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356889052134465536	allafrica.com/stories/202102...	https://allafrica.com/stories/202102030203.html	https://t.co/mPkatPR8lq

		#newspapers #feedly						
1.36E+18	2/3/2021 10:55	Zimbabwe: to Receive Covid-19 Vaccine https://t.co/okiNP3PCyc #newspapers #feedly	newspapers	feedly	https://twitter.com/anamafalda1992/status/1356889125056569345	allafrica.com/stories/202102...	https://allafrica.com/stories/202102030219.html	https://t.co/okiNP3PCyc
1.36E+18	2/3/2021 15:01	Ensuring Equitable Covid-19 Vaccine Distribution in South Africa https://t.co/ahPH3sU33K #newspapers #feedly	newspapers	Feedly	https://twitter.com/anamafalda1992/status/1356950886266896387	hrw.org/news/2021/02/0...	https://www.hrw.org/news/2021/02/03/ensuring-equitable-covid-19-vaccine-distribution-south-africa	https://t.co/ahPH3sU33K

1.36E +18	2/3/2021 10:54	Zimbabwe: Illegal Mining Activities Threaten Covid-19 Fight https://t.co/9YgVfObq9P #newspapers #feedly	newspapers	Feedly	https://twitter.com/anamafalda1992/status/1356888823263879169	allafrica.com/stories/202102...	https://allafrica.com/stories/202102030190.html	https://t.co/9YgVfObq9P

1.35E+18	1/27/2021 22:12	South African Broadcasting Corporation (SABC) News Anchor @NzingaQ will be the Moderator for Africa's Webinar on COVID-19: Vaccine Roll out Plan in Africa Date: 28 January 2021 Time:11h30	GCISAfricaWebinars	VaccineStrategySA	https://twitter.com/GCISMedia/status/1354522666477170688	gcis.zoom.us/j/98512742435	https://gcis.zoom.us/j/98512742435	https://t.co/zv2MEc0cCj

		CAT Zoom Link:https://t.co/zv2MEc0cCj #GCISAfricaWebinars #VaccineStrategySA #COVID19						

Appendix D

How to find stories using Twitter's API

- Download and install Spyder for free.
- Apply for Twitter developer keys and tokens.
- Add the keys and tokens to your Python script.
- Draft a set of questions you want to be answered from Twitter's data.
- Add your query – a hashtag or keyword to the script and run it.
- Analyze the data for patterns or outliers.
- Check to see if your questions are answered.
- Consider expanding your research with new keywords and hashtags from the analysis.

Researcher's Signature: ~~RamaruRR~~ Date: 3/8/2023