INFORMATION TECHNOLOGY AND
GLOBAL GOVERNANCE

Data Governance and Policy in Africa

Edited by
Bitange Ndemo · Njuguna Ndung’u
Scholastica Odhiambo · Abebe Shimeles

OPEN ACCESS

palgrave
macmillan

Information Technology and Global Governance

Series Editor
Derrick L. Cogburn
American University
Bethesda, MD, USA

Information Technology and Global Governance focuses on the complex interrelationships between the social, political, and economic processes of global governance that occur at national, regional, and international levels. These processes are influenced by the rapid and ongoing developments in information and communication technologies. At the same time, they affect numerous areas, create new opportunities and mechanisms for participation in global governance processes, and influence how governance is studied. Books in this series examine these relationships and influences.

Bitange Ndemo
Njuguna Ndung'u
Scholastica Odhiambo • Abebe Shimeles
Editors

Data Governance and Policy in Africa

palgrave
macmillan

Editors
Bitange Ndemo
Kenya's Ambassador
Belgium and the EU Mission
Brussels, Belgium

Scholastica Odhiambo
African Economic Research
Consortium
Nairobi, Kenya

Njuguna Ndung'u
African Economic Research
Consortium
Nairobi, Kenya

Abebe Shimeles
African Economic Research
Consortium
Nairobi, Kenya

Information Technology and Global Governance
ISBN 978-3-031-24497-1 ISBN 978-3-031-24498-8 (eBook)
https://doi.org/10.1007/978-3-031-24498-8

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover illustration: piranka \ getty images

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The utilization of data offers vast potential to transform African economies. Advances in data processing and storage capacity, as well as new insights possible through frontier data analytics in fields such as machine learning and artificial intelligence, are realizing tangible gains today. As a result, it is increasing efficiency and lowering costs, thereby boosting productivity growth with the potential to improve living standards across the continent. Furthermore, applications are leveraging data to close information gaps and increase transparency, thereby increasing the accessibility of finance amongst individuals, particularly within lower-income segments, and increasing their available opportunities. In addition, the vast potential of data to improve public sector governance and accountability through innovative and evidence-based policy-making, and therefore empower citizens, is increasingly being tapped.

However, some critical challenges exist. Large amounts of private data collected about individuals have the potential to be misused. Information can be collected and shared with third parties without the knowledge or consent of the data subject, thereby violating individual data privacy. Across Africa, although a large share of households has mobile phones, access to data is much lower and inequitable. Additionally, the potential of data to improve living standards is hindered by the small number of digitized datasets and accuracy concerns over collected data. In addition, the accessibility of digitized datasets, particularly those collected by the government using public resources, is frequently limited.

Africa must devise a continental data governance framework with the ultimate goal of improving living standards by maximizing the use of data and ensuring productive cross-border data flows while protecting individual rights. However, the continent lacks integrated regional structures for governing data, and individual African nations have wide disparities in existing infrastructure and legislation to safeguard data use. Therefore, developing a practical data governance framework requires scrutiny of the economic, legal, technological, and institutional issues attendant to such regulation, as well as the establishment of proper standards for the exchange and protection of data.

This book draws on research by experts across the continent in order to establish a foundation for a continental data governance framework. The issues covered include data collection and accuracy, the regulatory regimes for data across counties in Africa, challenges, and opportunities arising from the digitization of data and finance, the opportunities arising from new technologies such as blockchain, and the effects of data governance on the data scientist. These fundamental contributions lay a strong foundation for further policy work in making data work for Africa.

Nairobi, Kenya Bitange Ndemo

Acknowledgements

The authors thank the generous financial support provided by the Hewlett Foundation. Particularly, we are grateful to Mrs Sarah Lucas, former Program Officer, Hewlett Foundation, for her valuable guidance, insightful comments, and all rounded support since the inception of the project. We are also thankful to two anonymous reviewers for their comments and suggestions that improved a lot the quality of the book. The following discussants helped refine and sharpen individual chapters: Dr Miriam Omolo, Dr Ben Mkalama, Prof. Alex B. Makulilo, Dr Andrew Rens, and Dr Margaret N. Ndung'u. We are also thankful to the following distinguished policy panellists at the Regional Policy Forum on Data Governance and Policy held virtually on 22 November 2021, who provided their unique perspectives on the challenges, opportunities, and options for robust data governance and policy in Africa: Hon Joseph Mucheru, Cabinet Secretary, Ministry of ICT, Innovation and Youth Affairs, Kenya; Mrs Lesly Goh, Senior Technology Advisor, World Bank; Dr Michael Pisa, Policy Fellow, Centre for Global Development; Mrs Sonia Jorge, Executive Director, Alliance for Affordable Internet, Head of Digital Inclusion Program, Web Foundation; Maximilian Ghantz, Senior Policy Researcher, Mozilla, Berlin Office; and Dr Oliver Chinganya, Director, Africa Centre for Statistics, UNECA. The authors express their gratitude to Dianah Muchai, Obadia Miroro, Cecilia Mutanu, and Sheila Lyaga for their valuable support during the duration of this project.

Contents

Notes on Contributors

Olumide Babalola is a digital rights, privacy, and data protection lawyer in Nigeria. He holds a Master's degree in International Commercial Law with ICT and Commerce from the University of Reading, United Kingdom, and is a PhD candidate at the University of Portsmouth, United Kingdom. Olumide co-founded Digital Rights Lawyers Initiative (DRLI) as a civil society and network of lawyers with the principal objective of promotion of digital rights. He is a widely published author on the topic of digital rights, privacy, and data protection.

Zachary Mwangi Chege has over 30 years' experience in the public service with vast expertise in official statistics, public policy analysis, planning, budgeting, project management, and strategic management. He formerly served as the Director General of the Kenya National Bureau of Statistics (KNBS) for eight years, Chairperson of the United Nations Statistics Commission (UNSC) and African Centre for Statistics, Economic Commission for Africa (ACS–ECA), coordinating the work on modernization and transformation of national statistical systems and official statistics in Africa. He holds a Master of Arts (MA) degree in Economic Policy Management from Makerere University, Uganda, and a Bachelor of Arts (BA) degree in Economics, First Class Honours, from the University of Nairobi, Kenya.

Hanani Hlomani has worked with the African Legal Information Institute, the Democratic Governance Rights Unit, the Intellectual Property Unit at the University of Cape Town, and Research ICT Africa. He has experience with the African Union's Continental Data Policy

Framework and the Southern Africa Development Cooperation (SADC) Digital Economy Model Law Framework projects. He is an associate at AfriConsult Firm. He holds a Bachelor of Laws (LLB *cum laude*) from the University of Fort Hare and a Master of Laws (LLM) in Intellectual Property Law and is pursuing a PhD in Commercial Law at the University of Cape Town, within the DSI-NRF SARChI Research Chair in Intellectual Property, Innovation research group. He is a Mandela Rhodes Foundation alumni scholar (Class of 2019), and his key intellectual interests include innovation, research and development, data regulation, the arts and music, and financial technology.

Vukosi Marivate is the ABSA UP Chair of Data Science at the University of Pretoria. Marivate works on developing Machine Learning/Artificial Intelligence methods to extract insights from data. A large part of his work over the last few years has been in the intersection of Machine Learning and Natural Language Processing. Marivate's work in this area focuses on techniques to improve tools for and availability of data for local languages or low resource languages. Marivate is a co-founder of Deep Learning Indaba (https://deeplearningindaba.com/). He currently serves as a co-founder and chief investigator on the Masakhane NLP project (https://www.masakhane.io/).

Ben Mkalama is a research fellow and lecturer at the Faculty of Business and Management and Management Science at the University of Nairobi. A well-grounded professional with a unique blend of experience that ranges from academia, research and industry, in addition to his university experience, he has over 28 years of working and consultancy experience in senior leadership and management positions with international financial and local organisations in Africa. His current research interests are on entrepreneurship, inequalities in venture capital financing and different aspects of digital innovation.

Caroline B. Ncube is the Department of Science and Technology-National Research Foundation -South African Research Chair Initiative (DST-NRF SARChI) Research Chair in Intellectual Property, Innovation, and Development in the Department of Commercial Law at the University of Cape Town (UCT). She holds a PhD in IP Law from the University of Cape Town, an LLM from the University of Cambridge, and an LLB from the University of Zimbabwe. She is a co-leader of the Open African Innovation Research Partnership and a member of Academy of Science of South Africa (ASSAf); African Policy, Research & Advisory Group on STI and the African Continental Free Trade Area (AfCFTA) Trade and Industrial Development Advisory Council; African Union Scientific Technical Research Commission (ASRIC) Taskforce on IP Protection in Joint Research and Collaboration During Outbreaks. She is a widely published author: a co-editor of the *South African Intellectual Property Law Journal* and *Journal of World Intellectual Property.*

Bitange Ndemo is Kenya's Ambassador to Belgium and the EU Mission. He is an advisor and board member to several organizations, including Safaricom. His other assignments include Senior Advisor to the UN's Global Pulse (Big Data initiatives), the UNCDF's Better than Cash Alliance, and UNESCO's Innovation Council. He is a member of the MIT Artificial Intelligence Policy Forum (MIT AIPF) Panel member, OECD Expert Panel on Artificial Intelligence and Blockchain, and World Economic Forum's Global Blockchain Council. He formerly served as Professor of Entrepreneurship at the University of Nairobi, Chairman of the Kenya Distributed Ledgers and Artificial Intelligence Taskforce, and Permanent Secretary of Kenya's Ministry of Information and Communication in Kenya. His research centres on the link between ICTs and small and medium enterprises, emphasizing how ICTs influence economic development in Africa.

Njuguna Ndung'u is the Cabinet Secretary, the National Treasury and Planning, Government of Kenya, and the Outgoing Executive Director of the African Economic Research Consortium (AERC). He is Associate Professor of Economics at the University of Nairobi, Kenya. He formerly served as: Governor of the Central Bank of Kenya; member of Global Advisory Council (GAC) of the World Economic Forum; Visiting Fellow of Practice at the Blavatnik School of Government, Oxford University; Director of Training at the African Economic Research Consortium; International Development Research Centre of Canada (IDRC) and the Kenya Institute for Public Policy Research and Analysis (KIPPRA). He holds a PhD in Economics from the University of Gothenburg, Sweden. He has published widely in international journals as well as chapters in various books on economic policy issues. He is a member of the Brookings Africa Growth Initiative (AGI) Distinguished Advisory Group, a member of the Advisory Committee of the Alliance for Financial Inclusion (AFI) that coordinates financial inclusion policies in Africa, Asia, and Latin America, and a senior advisor for the UNCDF-based Better Than Cash Alliance.

Scholastica Odhiambo is a Manager of Research at the African Economic Research Consortium (AERC) and Senior Lecturer of Economics at Maseno University. She is managing both thematic and collaborative research programmes at the AERC. She has previously worked on research projects funded by INCLUDE, the European Union, and Japan International Cooperation Agency (JICA). Her expertise encompasses policy analysis teaching and research in microeconomic policy, macroeconomic policy, econometrics, health, and environmental economics. She is an AERC Scholar and holds a PhD in Economics and a Master of Arts degree in Economics from University of Nairobi, Kenya, and Chancellor College, University of Malawi, respectively. She is a widely published author.

Miriam Omolo is the Executive Director at the African Policy Research Institute (APRI). She has over 20 years of experience in public policy research and analysis. She holds a doctorate degree in economics (D.Litt. et Phil. Economics) from the University of South Africa. Her areas of research work are on trade, development and poverty, public finance management. She is also working on emerging areas of extractive sector and digital economy. She has experience in both macro and micro econometric modeling using large survey data sets. Miriam is a network member of the African Economic Research Consortium (AERC) and African Growth Policy Modeling (AGRODEP) Network hosted by the International Food Policy Research Institute (IFPRI).

Abebe Shimeles is an honorary professor, Department of Economics, University of Cape Town. He is the immediate former Director of Research where he provided intellectual and strategic leadership for the thematic and collaborative research programmes at the African Economic Research Consortium. He has over 30 years of professional experience in policy research working for academia, NGOs (Action Aid), and other international organizations (United Nations Economic Commission for Africa, the World Bank, and the African Development Bank) with a proven track record of publications in peer-reviewed journals, policy dialogue, and impacting policy through research. His recent position at the African Development Bank was Division Manager of the Macroeconomic Policy, Forecasting, and Research department. He holds a PhD in Economics from Goteborg University, an MA from Delhi School of Economics, and a BA from Addis Ababa University. He started his research career at the AERC in the mid1990s and continued to serve as a resource person since 2009.

Aaron Thegeya is an economist and computer scientist. He has worked as a Senior Economic Advisor to the Deputy Chief of Staff in the Executive Office of the President of Kenya, and also as an Economic Advisor the Chair, Commission on Revenue Allocation. He holds a PhD in Economics from the University of Oxford, where he studied as a Rhodes Scholar, and where he also completed a Post-Doctoral Fellowship at the Said Business School. He also holds a Bachelor's degree in Economics and Computer Science from Johns Hopkins University. He has worked as an Economist at the International Monetary Fund and the World Bank, and is currently director at Aliquot Limited, an economics and technology consultancy.

Peter Maina Wanjohi is an economist/statistician by profession and currently an assistant manager, Labour and Prices Statistics at Kenya National Bureau of Statistics. He holds Masters of Arts in Economics from Waseda University, Japan and Bachelor of Arts in Economics (*First Class Honors*) from University of Nairobi. He has nine years' experience in the field of statistics and has actively participated in collection, processing, analysis, reporting and dissemination of official statistics mainly in areas of Education, Labour, Prices and Population Statistics. He has a strong passion in research in the field of economics and statistics.

List of Figures

List of Tables

CHAPTER 1

Introduction

Bitange Ndemo, Njuguna Ndung'u,
Scholastica Odhiambo, and Abebe Shimeles

1.1 Introduction

Sound economic policy presupposes availability of timely, comprehensive, credible, and multi-purpose data that many African countries have lacked for a long period. It is not long ago when major policy reforms were implemented based on findings drawn from faulty data. A recent project

B. Ndemo
Kenya's Ambassador, Belgium and the EU Mission, Brussels, Belgium
e-mail: bndemo@bitangendemo.me

N. Ndung'u
National Treasury and Economic Planning, Nairobi, Kenya
e-mail: njuguna.ndungu@aercafrica.org

S. Odhiambo (✉)
African Economic Research Consortium, Nairobi, Kenya
e-mail: scholastica.odhiambo@aercafrica.org

A. Shimeles
Department of Economics, University of Capetown, Capetown, South Africa
e-mail: abebe.shimeles@uct.ac.za

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8_1

by the World Bank on "Agriculture in Africa: telling facts from myth[1]" is a clear illustration of how, for many years, policymakers in Africa formulated their policies toward smallholder farmers based on stylized facts that were inaccurate or untrue. The project identified over 16 well-established "myths" that generally had been taken as facts and informed decision-making in most African countries. We can cite many other examples as well that triggered significant policy actions driven by faulty data.[2] Some researchers also argued that Africa's national statistics are significantly affected by measurement errors, poor data management, and weak capacity, making national development strategies incoherent.[3] On the other hand, relatively reliable data generated from large household or labor force surveys or census in most cases remain unused for policymaking for various reasons. It is here that AERC has played a major role in facilitating the use of such data sets for the analysis of Africa's labor markets, poverty and inequality, small holder land and labor productivity, and so forth and made significant contributions to the understanding of African economies. Still Africa has remained the most under-researched continent to date.

The last two decades, however, has seen significant improvements in the availability of high-quality data that can be used for effective policy evaluation. The advent of behavioral economics, the availability of big data through satellite imagery by NASA and others, such as changes in temperature, rainfall, soil quality, vegetation, night light data, movement of people and traffic across borders through Google Maps, and most of all data generated by mobile telephony present unique opportunities for Africa. Recent years have seen a rise in cutting-edge research on African economies utilizing such data as well as well-designed field experiments. A combination of these two has given researchers unique advantages to explore frontier issues that were unthinkable a few decades ago. The digital footprint that flourished in the continent in the financial sector and the

[1] See details of the project in https://www.worldbank.org/en/programs/africa-myths-and-facts

[2] One good example is the 'finding' that private returns to education is the highest in Africa at primary level. This then induced development partners, particularly the World Bank and other donors, to push African governments to prioritize spending on primary education, while neglecting secondary and higher education. Today, Africa is the lowest in terms of number of scientists per thousands of population, research patents, centers of excellence, etc.

[3] Please see the book by Morten Jerven (2015): *Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It.*

use of social media and other platforms also present advantages that enable economic research in Africa. The African Economic Research Consortium, with the support of the Hewlett Foundation, undertook a scoping study to bridge the gap between evidence and data, as well as evidence and policy, but more importantly to draw attention to appropriate data management policy, data use, data protection, and data governance in the context of Africa. The introduction section presents the key findings from the seven chapters contained in this volume and outlines the way forward.

The contributors to this book consist of economists, lawyers, statisticians, and data technology experts to assess the opportunities, challenges, and risks existent in the current state of data generations, sharing protocols, and consistency of legislations. Desirous of encouraging the widespread use of large data from different sources, ensure its reliability, and facilitate availability, the book seeks to identify opportunities, constraints, and impediments to the use of data and evidence to inform economic policy decision-making in sub-Saharan Africa and, in so doing, to engage policymakers on the implications for data governance that could provide safe access and foster its widespread use. The main objective of this book is twofold. First, to create a platform on data governance that raises awareness on basic principles/tenets of international norms and shares experiences and practices from across the world and Africa. Second, to initiate, promote, and advocate for data governance protocols in the era of digital revolution and assess the potentials for improving the digital market to enhance benefits to African consumers, governments, and businesses. To achieve these objectives, the chapters covered broad themes that provide better understanding of the challenges, opportunities, and risks in the process of data production, consumption, and utilization. Chapter 2, “A Prototype Data Governance Framework for Africa” by Bitange Ndemo and Aaron Thegeya, outlines a framework for data governance that ensures sovereignty, while at the same time enhances productivity within each African country and bolsters cross-country collaborations. The chapter proposes a continent-wide data governance strategy that “abide by core principles such as preserving accountability, ensuring data accuracy and quality, and facilitating interoperability and standardization of data.” The chapter discusses global practices that provide the principles and framework for data governance that promotes accountability in improving data quality, integrity, privacy, and security of data subjects, as well as ethical use of data.

Implementation of a robust data governance strategy requires a thorough understanding of the process involved in translating data into information and policy action. Chapter 3, "A Value Chain Approach to Data Production, Use, and Governance for Sound Policymaking in Africa" by Zachary Mwani Chege and Peter Maina Wanjohi, provides a schematic approach to data production, consumption, and application starting downstream (data generation) all the way up to data utilization and its application for decision-making. Using National Statistical Offices (NSOs) as the prime examples entrusted with the responsibility of overseeing publicly available data, the chapter presents the guiding principles that govern data generation, sharing, dissemination, and protocols that are enshrined in United Nations Fundamental Principle of Official Statistics (UNFPOS).[4] These principles touch upon a wide range of issues such as relevance, impartiality, and equal access; professionalism; accountability; prevention of misuse; cost-effectiveness; confidentiality; legislation; national co-coordination; international coordination and cooperation; as well as comprehensive documentation. The chapter reviews where Africa stands with respect to these principles and identifies gaps and areas for further improvement considering the rapid progress in digital technology, data revolution, and big data. Chapter 4, "Data Protection Legal Regime and Data Governance in Africa: An Overview" by Olumide Babalola, and Chap. 5, "Data Regulation in Africa, Free Flow of Data, Open Data Regimes and Cybersecurity" by Hanani Hlomani and Caroline B. Ncube, explore the regulatory and legal challenges with respect to data governance and policy in Africa. Chapter 4 presents the "legal framework around data protection in Africa in the light of their salient provisions, adequacy, efficiency, and enforceability in relation to data governance on the continent," with emphasis on ratifications and declarations made under the auspices of Africa Union, such as the Malabo Convention. The paper's key message revolves around the following: first, "free movement of both personal and non-personal data is best supported by the adoption and use of open standards for data and appropriate cybersecurity regulation." Second, policymakers and legislators ought to consider the above threefold key message in formulating policies and drafting laws. Third, the chapter also suggests that "further research may be carried out on data localization laws and follow the developments pertaining to the African Union Commission's Draft Data Policy Framework."

[4] See, for example, https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx

Advances in digital technology, geospatial data, satellite imageries, and big data have transformed the landscape in which data governance has to be applied. Chapter 6, "Digitalization and Financial Data Governance in Africa: Challenges and Opportunities" by Bitange Ndemo and Ben Mkalama, elaborates and discusses how emerging technologies in the areas of artificial intelligence, robotics, Internet of things, and big data analytics have shaped business models across the globe. Africa needs to be ready to take advantage of these technologies, while mitigating the risks they pose in many areas including national security, privacy protection of data subjects, unethical use of data for purposes of disinformation and misinformation, and so forth. The paper notes that despite such significant opportunities afforded and heightened risk posed by the digital technology, many African countries are behind the curve in devising legislations and installing infrastructure necessary for data protection and management. A sharper example is found in Chap. 7, "The Economics of Blockchains Within Africa" by Aaron Thegeya, which discusses in great detail the upcoming blockchain technology that has transformed economic data in unprecedented manner with a potential "to boost levels of productivity and unlock capital flows to underserved sectors, in addition to leveraging the increasing returns of information as an input to production to spur economic growth." The challenge, however, is that only very few countries in Africa (Mauritius and Kenya) have devised the framework, including legislation and regulatory provisions, to take advantage of blockchain technology in equitable and sustainable manner. The chapters identified structural and institutional challenges African countries face to exploit fully the advantages blockchain technology offers and the risk it poses for disadvantaged groups and regions. The final chapter, "More Than Just a Policy—Day-to-Day Effects of Data Governance on the Data" by Vukosi Marivate, presents the need and imperatives for data governance from the perspective of data science, which is an evolving discipline that applies algorithms and systems to extract patterns, and knowledge from both well-structured and noisy data. The chapter highlights the need for robust data governance systems to manage the proliferation of analysis, decision-making, and other impactful measures adopted by businesses, government entities, and others using data science where the margin for error and corruption is paramount.

1.2 The Way Forward

The chapters in this volume underscored that data intrinsically are generated in unlimited quantity, and, in most cases, it is a non-rival good that could be shared or used across users without creating any disincentives or diseconomies. These unique characteristics of data, coupled with the increasingly data-driven global economy, could create the risk of enormous power imbalances, inequalities, and diverging development clubs. Examples include the unequal access to broadband Internet across the globe, which has widened over time, leading to unequal generation and utilization of data. Divergent cybersecurity capabilities led to different degrees of protection of privacy and exposure to the misappropriation of national- and individual-level data. When it comes to Africa, the chapters note that national statistical offices generally are behind the curve in the understanding and utilization of digital technology in data creation as well as making it ready for use by researchers, decision-makers, and private citizens. The Covid-19 provided an opportunity for many African governments to recognize and experience the huge potential presented by the digital technology in generating, storing, and distributing data for purposes of real-time decision-making. Noting these insights, the AERC anticipates further engagement in the realm of data governance and policy in Africa. Some of the ideas are presented below.

1.3 Potential for Future Research

As noted above, African governments are at different stages of data governance and policy practices. But common to all seems the lack of appreciation of data policy to enhance decision-making based on solid empirical evidence, of which many African governments generally have subscribed to and aspire to inform their strategies and policies. This disconnect between the absence of data policy and the desire to influence policy through evidence is evident in almost all African countries. The following issues encapsulate future research to close such a gap:

- *Data policy and inclusive development in Africa*

Many experts, including those in the policy panel, suggested that critical socioeconomic data are collected in unreasonable time intervals, are rarely available to inform policy on time, and are not accessible to researchers to conduct studies to build the stock of knowledge that could be transformative. It is not known how much such data management practice costs Africa in terms of lost development. For example, few African governments have reliable and up-to-date data on profile of their labor force, stock of skills, and employment opportunities. Yet, they formulate various labor market policies and legislations and develop plans on employment generation and even social protection programs based on outdated and incomplete data sets. Assessing the extent to which such data policies have prevented African countries from formulating reliable and achievable development policies is an important research undertaking.

- *Benchmarking Africa in the global data policy and governance*

The Fourth Industrial Revolution is riding on the back of data and information. The full potentials of emerging new technologies are realized in countries where the data infrastructure and system are compatible with the global norm. Data policy is not just about protection. It is also about the right of full access, principle of usability, and integrity. Such a policy requires data-generating bodies to fully comply with the rules of data sharing, availability, and integrity protocols. Benchmarking African countries with the global norm could be an eye opener to various government agencies, private operators, and other data-generating agencies in terms of the opportunities Africa is losing to catch up with the rest of the world.

- *Mapping data interoperability in Africa and the future of policymaking*

One recurring theme raised during the RPF was the significant gap observed in Africa with regard to various socioeconomic and other individual-level data being fragmented, scattered, and poorly organized, costing governments enormous resources and time to manage the economy. Examples include important socioeconomic surveys designed and collected independently from each other, costing millions in resources and at the same time limiting their significance for policymaking. Examples include census surveys, labor force surveys, living standard measurement surveys, demographic and health surveys, which are designed and collected independently of labor force surveys and many other regularly conducted national surveys. These are also not linked with other vital statistics such as income taxes, financial transactions, health, and education services. Most developing countries have now started to use unique identification codes for citizens that, with the right privacy policies, could be integrated with other data platforms. That offers policymakers important information to formulate public policies that are innovative and inclusive.

CHAPTER 2

A Prototype Data Governance Framework for Africa

Bitange Ndemo and Aaron Thegeya

2.1 Introduction

In their simplest form, data are frequently defined as a collection of symbols that are the properties of observables or the representation of facts. Data within a given context translate into information—and information in perspective, integrated into a viewpoint based on experience—is what we think of as knowledge (Ackoff, 1989). Despite the distinction between data and information, the terms are often interchangeable in practice. Data are an important component of total factor productivity and contribute in important ways to growth, in addition to labor and capital. There are massive economies of scale to be gained from combining different data sets to yield insights that would be otherwise unavailable or difficult to capture. In addition, improvements in data processing, data storage, and data analytics through machine learning and artificial intelligence can support

B. Ndemo (✉)
Kenya's Ambassador, Belgium and the EU Mission, Brussels, Belgium
e-mail: bndemo@bitangendemo.me

A. Thegeya
Aliquot Limited, Nairobi, Kenya

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8_2

productivity gains, boost efficiency, and decrease costs—advances that can drive economic growth, increase prosperity, and improve the standard of living on the continent.

Data governance involves establishing principles to enable an environment for the sharing of data, with the ultimate goal of improving living standards, while at the same time recognizing and protecting the rights of data originators and users. Given the central role of data in today's global economy, a system of effective data governance is essential. That said, the development of any such framework requires careful scrutiny of the economic, legal, and institutional issues attendant to such regulation, as well as the establishment of proper standards for the exchange and protection of data.

At the micro or firm level, data governance has historically referred to managing the availability, usability, integrity, and security of data. From a global perspective, the World Bank (2021a) has deemed that data governance "entails creating an environment of implementing norms, infrastructure policies and technical mechanisms, laws and regulations for data, related economic policies, and institutions that can effectively enable the safe, trustworthy use of data to achieve development outcomes" (p. 38). To leverage the vast opportunities of data utilization, Africa must develop a data strategy underpinned by a governance framework. Such a strategy would establish data sovereignty and render the continent more competitive and better positioned to engage in cross-country collaboration during the digital age. Data could be reused by promoting practices protective of privacy, including personal and other sensitive data, through techniques including anonymization, pseudonymization, differential privacy, generalization, suppression, and randomization.

Africa must define a continental strategy and devise a governance framework that maximizes the use of data while ensuring productive cross-border data flows and protecting individual rights. The continent lacks sturdy and expansive national or regional structures for governing data, and individual African nations have yet to develop legislation to safeguard data use and digital transactions—an absence likely to cause market fragmentation due to insufficient harmonization (United Nations Congress on Trade and Development, 2021). Meanwhile, the data governance frameworks that do exist, albeit in their limited form, lack coherence in terms of principles, scope, and enforceability across jurisdictions. The rapidly changing landscape of data generation, storage, and mining capacity—as well as the dearth of human and financial resources, reliable institutions, and enforcement capacity to support an efficient data governance environment—will, absent immediate action by key stakeholders, cause the continent to regress at the moment when it is arguably positioned to show its greatest progress ever.

Without a coherent data governance framework, data generated in Africa risk being improperly utilized within each country and in other parts of the world, leading to an unbalanced platform of data exchange with countries where data are closely regulated. African countries, therefore, must take the lead in establishing the appropriate frameworks that will serve not only their own national interests but also those of the continent as a whole. Governments, development institutions, and nonstate actors should collaborate to implement and enforce data governance laws and policies that can make the continent's digital economy more competitive while, at the same time, enhancing transparency, trust, and digital inclusiveness for all users.

2.2 Background and Literature Review

The history of data is closely intertwined with the evolution of mankind. The earliest examples of data being stored and analyzed by humans date back to about 18,000 BCE, in what is now Uganda, when humans were recorded using the Ishango bone for the purposes of tallying (Marr, 2015). The bones were marked with notches to keep track of trading activity, and notches were compared between bones to carry out rudimentary calculations on supplies. Subsequently, the abacus, the first device constructed specifically for performing calculations, was invented around 2400 BCE. The first data libraries appeared during roughly the same period, marking mankind's initial endeavor toward mass data storage. The year 1663 saw the emergence of statistics as a distinct mode of analysis, when John Graunt recorded mortality information in London and used his figures and framework to design an early warning system to alert the population about the spread of the bubonic plague that had been ravaging Europe. The central concept of the modern computer emerged thereafter, based on the ideas of Alan Turing, who, in 1936, presented the notion of a universal machine (Zimmermann, 2017), paving the way for the first digital computers in the following decade. Finally, data became ubiquitous with the advent of the Internet, announced by Tim Berners-Lee, in 1991, thus setting the stage for the modern age of big data.

Historically, people struggled to collect data because they lacked the necessary tools and infrastructure; the digital revolution, however, led to dramatic changes in the scope and types of data collected, and the volume of data sets collected has increased compared to only a few decades ago. What's more, when governments fail to do the collecting, private firms and individuals can now use new digital platforms to gather data for private use, for commercial purposes, or to promote accountability and governance—such as platforms used to report violence or discrimination. The cataloguing of information from Africa's past, through the digitization of archived records and the utilization of disparate data sources for analysis of economic activity, climate, and terrain, for example, has increased the set of data available as well as analytical findings from both research and commercial perspectives. This transformation has encouraged insightful publications on Africa's past (see, e.g., Fourie, 2016) that would not have been possible without today's data infrastructure.

Data consumption needs have increased significantly over time, and consumption of data varies by region. Daily usage statistics are staggering: From the advent of civilization to 2003, for example, five exabytes of data were created[1]; but, by only seven years later, that amount of data was being generated every two days (World Bank, 2021b). By 2025, it is estimated that 463 exabytes of data will be created around the world each day (Desjardins, 2019). Presently, the entire universe of data is estimated at 44 zettabytes, a total that accounts, for example, for the 294 billion emails sent, the 5 billion Internet searches that occur, and the 65 billion messages transmitted each day through messaging services such as WhatsApp (World Bank, 2021b).

A World Bank (2021b) study looking at minimum data consumption using data from six developing and emerging countries found that the most frequent online activities, which included visits to public service websites, learning, shopping, health information and news, consumed 660 megabytes of data per user, per month. When looking beyond data requirements for solely welfare-improving activities like those just mentioned, individuals in these countries needed an additional 5.2 gigabytes for recreational activities on social media per month, putting total monthly data demand in these economies at approximately 6 gigabytes per person (2021b). Never has a data governance framework been more necessary than it is now.

[1] An exabyte corresponds to 10^{21} bytes, and a zettabyte corresponds to 10^{24} bytes.

A review of the literature shows that the body of research on data governance has been carried out mostly from an organizational perspective. Given data's role as a strategic and monetizable asset, organizations have researched holistic data governance frameworks to facilitate effective utilization of data with a profit motive, while respecting privacy rights (see, e.g., Khatri & Brown, 2010; Otto, 2011; Weber et al., 2009). From a regulatory perspective, countries are in the process of defining data governance frameworks. For example, in November 2020, the European Commission proposed rules on data governance to boost data sharing and support European data spaces, in line with principles such as personal data protection (General Data Protection Regulation), consumer protection, and competition. The World Bank has even focused its 2021 World Development Report on data issues pertinent to developing economies.

Micheli et al. (2020) investigated the emerging models of data governance in the age of datafication and, in addressing the politics of data, considered actors' competitive struggles. This conceptualization brought to the forefront the multifaceted economic and social interactions, as well as power relations, within data governance models—particularly those at work in corporate environments. Public bodies and civil society are, within these models, key players for both redistributing any value produced via data and democratizing its governance. Further, Micheli et al. found that data trust and intermediaries were included in nearly every investigated model, leading the researchers to underscore the importance of data infrastructure as fundamental to improving trust in data.

Research has also revealed a wide variety of views and minimal agreement across stakeholders on the issue of data governance frameworks. Within the context of academia, Kouper et al. (2020) carried out an exploratory study on data governance in the United States, involving individuals who worked in research and academic institutions, aiming to understand the entities central to decision-making and governance on data and research-related issues. This group's findings showed considerable complexity and diversity across stakeholders in terms of both identity and ideas on the governance of data. To account for such diversity, Kouper et al. proposed to frame data governance in research around common governance bodies, arguing as well that, to ensure effective data governance in research, voices of people from different literacy and income levels should always be incorporated in shaping policy and making decisions.

Several approaches have been used to determine data governance activities. For instance, Alhassan et al. (2016) used key words to identify papers on data governance activities using open-coding approaches and identified 31 articles that mentioned such activities. Their analysis identified 110 data governance activities across five decision domains of their framework (data principles, metadata, data quality, data life cycle, and data access), with each domain implicating a different critical aspect of data governance.

The rapid growth in digital financial services presents concerns over data protection and privacy for low-income individuals, especially those in developing countries. Vidal and Medine (2019), for example, analyzed whether data privacy is desirable in a corporate world. Their analysis included experiments in India and Kenya, where several products with varying degrees of data protection and a range of privacy options were offered to low-income individuals, thereby allowing the researchers to evaluate the demand for individual safeguards within markets with limited or no frameworks in place to protect individual privacy, and they found that low-income individuals were willing to pay for their data privacy. For instance, in Kenya, 64% of low-income individuals surveyed chose options with a greater degree of data privacy, despite the imposition of a non-trivial 10% fee attached to this option. Even more, results in Bangalore were similar to those of Kenya, with 66% of survey participants choosing this option.

Freely available public data could generate economies of scale through reuse, and the benefits of these types of data in terms of the public good present a case for protecting the availability of some classes of data from public sources relative to private firms. Beraja et al. (2020) analyzed the state of artificial intelligence within China by gathering comprehensive data from government and firm-procurement contracts within the artificial intelligence industry and found that sharing data improved productivity in both private and public institutions. Their results also indicated that the ability to access government data outweighed the feasibility of providing these same data through commercial means; accessible government data should not, they concluded, be substituted by private markets.

2.3 An Organizational Framework for Data Governance

A sound data governance framework requires that institutions and stakeholders have the right incentives to produce, protect, and share data; a comprehensive understanding of data governance also demands consideration of key dimensions including (a) the relevant stakeholders who use data and those who are impacted by the use of data; (b) the life cycle of data from creation to destruction; (c) the typology of data, reflecting relevant characteristics that impact processing, storage, and accuracy; and (d) enabling pillars such as economic, legal, and institutional aspects that create the necessary infrastructure for using data and maximize its productivity. These key dimensions are illustrated in Fig. 2.1.

The key stakeholders who generate and use data include households, the private sector, governments, and civil society, with households and the private sector being major data producers and/or consumers and

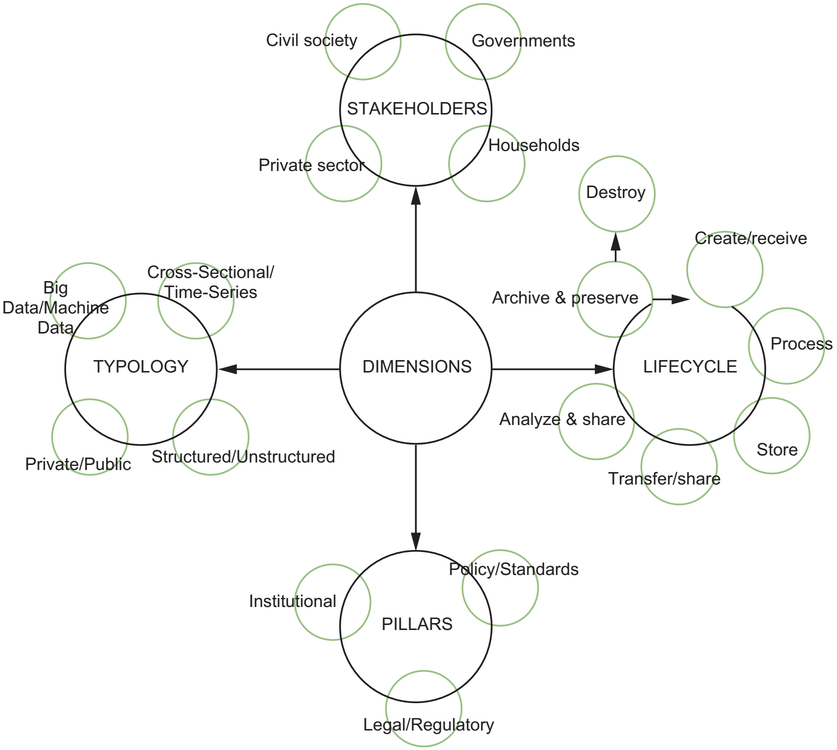


Fig. 2.1 Organizational framework for data governance. Note: Figure conceived and designed by the authors

governments and civil society offering essential safeguards concerning its use. The government is central in formulating policies. and regulations, while civil society helps hold other stakeholders accountable. The needs of each stakeholder bear consideration within a data governance framework. Data privacy concerns, for example, vary across stakeholders and are relevant particularly for households and the private sector; conversely, certain data collected by governments merit classification under public data, particularly where the utilization of these data improves productivity and creates economies of scale—and also given that the data are collected using public resources.

The data life cycle details the key steps that occur between the creation and destruction or reuse of data, specifically the collection, processing, and storage of data; transferring or sharing of data among users; analysis and value addition; archiving and preservation for future use; and destruction of data at the end of the cycle. Stored or archived data are usually available for reuse, and an enabling infrastructure is essential during each step, including security of storage and transmission through encryption—protocols that enable data transfer across systems, allow its destruction at the end of the cycle, and maintain integrity and accuracy of data by preventing unauthorized manipulation.

Data collection and processing methods help determine accuracy, in turn promoting greater trust in data sets. Established data collection techniques for public data include the collection of population statistics by an official authority, or the collection of sample statistics using rigorous sample-design techniques. These methods yield accurate and trusted data sets structured in nature, but they also tend to command significant resources during the collection, depending on the level of disaggregation required, either in terms of subpopulations or regions of interest. Due to the financial cost, as well as planning and logistical requirements associated with collection, these techniques tend to be implemented infrequently. Therefore, analyses based on these data usually have gaps, either in their level of disaggregation or across time.

A large number of distinct typologies of data exist, determined by the multidimensional aspects inherent in data, as well as the lens or perspective through which data are viewed. Data can be classified according to whether they are for private or public use, a distinction that, in turn, determines how widely available they might be as well as their cost to access. Data collected for commercial use are treated as a private good, and those who own them enjoy a competitive advantage as well as the ability to collect fees when selling them. Public data, by contrast, which were collected used public resources, are intended to be widely available. Publicly available data usually provide social value and are useful inputs for other economic activities. Open data, for example, are a type of public data shared to fortify public governance and increase transparency, while also generating commercial opportunities.

For purposes of classification, structured data are organized according to some predefined model and stored electronically, typically in a relational database within a tabular format. Databases allow for efficient searching, editing, and error detection; they can also be more easily manipulated by programming languages. Conversely, unstructured data are less organized, are typically text-heavy, and require more flexible data structures. These data are more difficult for programs to process. Additionally, data can be classified according to cross-sectional and temporal dimensions, with cross-sectional data including many observations on subjects recorded at a fixed point and time-stamped data accounting for observations on one or many subjects recorded over time. Spatio-temporal data describe both the time and location of a particular event.

The utilization of big data is a subset of new collection and analysis techniques involving unstructured data, techniques made possible by accessible, less expensive, and more expansive storage capacity, as well as by advances in machine learning related to processing capacity and the increasing production of large amounts of digital data. These techniques reveal patterns from high frequency data, in real time, while low statistical errors within the data are supported by the large number of observations. Machine-learning algorithms depend on the availability of large data sets, with the predictive power of the algorithms increasing as the data become more available, even as the effectiveness of the algorithms continues to depend on the accuracy of the training data being used.

Newer techniques for collecting data rely on the availability of digital data and depend on both advances in machine learning and estimation theory at the small-area level. These methods offer advantages relative to traditional methods in terms of cost, frequency, and coverage; the use of small-area estimation techniques, for example, allows interpolation of statistics at a disaggregated level based on the combination of population, sample, and even satellite data. These methods may also include data from unstructured sources, and the accuracy of these methods remains an area of active research.

Underpinning the rights of stakeholders, the flow of data within its life cycle and the various data typologies are the enabling pillars of an effective data governance framework. The pillars include the economic, legal, and institutional framework that facilitates policies enabling the appropriate use of data while protecting data privacy, standards that embed data accuracy and make possible the secure storage and transfer of data, and the implementation and enforcement of appropriate regulation for the use of data. Establishing appropriate legally empowered institutions to create and regulate the data space is a critical dimension of the enabling framework.

A number of data governance frameworks exist, varying in terms of membership and degree of implementation. In 2014, the African Union adopted the Malabo Convention, which sought to encourage cybersecurity and personal data protection among partner countries, although the Convention was not fully implemented and thus not enforceable. According to the Convention, data need not be stored once the purpose for which it had been collected was met. This would have called for personal data to be protected by deletion when its purposes were achieved, meaning that data controllers would need to follow up with other data users to ascertain the destruction of personal data.

The Economic Community of West African States (ECOWAS) was established to promote the integration and economic growth of its member states. The member states adopted a Personal Data Protection Act in 2010, an agreement covering personal data and consent by the subject, recipient, and third parties, as well as the role of data processors and a data protection authority. However, the Act does not cover other important dimensions, such as profiling, anonymization, personal data breaches, and pseudonymization—matters of particular relevance with respect to cross-border data flows within the ECOWAS community. The Act requires member states to develop independent data-processing agreements for their citizens, guaranteeing their professional secrecy, impartiality, and power to punish errant parties. According to the Act, the processing of personal data is legitimate when carried out with the owner's consent and approval.

The Asia-Pacific Economic Cooperation (APEC) has developed a privacy framework for Asian countries, specifically in the Pacific region. APEC aims to promote flexible and effective information flows within the APEC community, while ensuring well-managed data protection. In 2020, to protect their government institutions, firms, and individuals against harm or the risk of private data being exposed, as well as to promote trade and ensure trust among member states, New Zealand, Chile, and Singapore signed a Digital Economy Partnership Agreement (DEPA) governing and protecting the sharing and processing of electronic data.

Other well-established data governance frameworks can be found in the European Union and the United States. In 2018, the European Union put in place its General Data Protection Regulation (GDPR), enshrining it as the legally recognized framework for data privacy and protection among member states. This framework governs the European Union's member states and their trading partners in all matters of data governance. In a rather different and definitively disaggregated manner, the United States offers both state and federal laws to protect personal online data and privacy.

2.4 A Prototype Data Governance Framework for Africa

Establishing an effective data governance framework for Africa requires a clear delineation of its objectives and careful attention to the unique characteristics of the continent. Africa has a large informal sector, an agricultural sector that dominates in production, and most of its commercial entities are small businesses. Much of the population connects through mobile phones, even as data access levels are much lower, averaging about 20% of the population, and while access to high speed data connections is even lower, still. Additionally, data access across households is highly uneven and depends on geographic location and economic status. For most Africans, the costs of enjoying Internet access are prohibitive, meaning that uneven access to data at a national level is mirrored by a large disparity in access at the continental level.

Digital-format data are highly limited in Africa; many public data sets are not digitized, and wide access to those that are digitized is low, fragmented, and inconsistent. Low levels of Internet connectivity also deter households and the private sector from generating new digital data, which, in turn, presents a major barrier to producing the high levels of data concentration that can spur innovative activity, enable the utilization of big data and machine learning techniques for data mining, and increase productivity. Data strategies and governance frameworks do not exist in many African countries, and, where frameworks do exist, they are typically incomplete, disjointed, or not fully aligned with other existing and relevant legislation already in place, such as laws protecting individual rights. What's more, due to weak institutions, governance issues, or limited capacity, levels of enforceability within existing frameworks are low across the continent. The countries that already have data governance frameworks in place require close levels of coordination to avoid suffering fragmentation and, as a consequence, diminished effectiveness.

A pan-African data governance infrastructure can help the continent realize a single market for data, thereby enabling the creation, use, and reuse of data by individuals across Africa and spurring economic growth and development while protecting the rights of data subjects. A necessary prerequisite for a single data market is the generation of sufficient data to allow economies of scale through utilization. This means that an appropriate framework to rapidly increase data digitization and widespread access must be developed in parallel with a data governance framework, in addition to the establishment of other key legal and regulatory frameworks that comprehensively govern the data life cycle. Realizing an effective data governance framework is contingent on the establishment of country-level guidelines that provide a template instructing nations on the precise components necessary for a comprehensive framework, while also establishing principles to ensure coherence across the components within a country. Further, a complementary overall framework linked to and interoperable with national frameworks should be established at the continental level.

An effective framework requires a clear set of definitions and categories for different types of data as well as rules pertaining to the use and reuse of data within each category. In this regard, a framework should clearly define private versus public data and should offer clear guidelines on the use of each type. But effective implementation should also go a step further, designating key public data sets to be shared both nationally and across borders—data sets that should be identified according to the strategic interests of countries, thereby calling for a concurrent effort to determine and prioritize interests that will maximally promote the sharing of data. These may include, for example, expanding regional trade, boosting agricultural productivity and promoting food security, or dealing with climate-related threats. Cross-border data sharing can leverage principles employed in existing systems that effectively utilize information transcending borders, such as monitoring systems for infectious diseases.

A comprehensive data governance framework must rest on the widespread engagement of all stakeholders in a social contract that defines the protection of individual data, thereby building trust, creating an enabling environment that adds value to data, and promoting an equitable system (World Bank, 2021a). Such a social contract could overcome negative externalities resulting in the underutilization of data for productive activity, and, if properly implemented, it could define the role of and cultivate trust in data intermediaries—those figures or institutions central to the eventual success of a data governance framework.

Civil society has a central role to play in shaping the social contract across all other stakeholders by influencing policies on appropriate and optimal levels of data openness, transparency, and accountability. Through civil technology innovations such as open government platforms, civil society can leverage big data to improve tracking of the performance of public institutions in fulfilling their mandates, ensure wider dissemination of performance data to citizens to improve transparency, and integrate community feedback to allow broader active citizen participation in local government. Additionally, as the level of digitalization increases, civil society has a central role to play in ensuring that all individual rights are protected, in particular the rights of vulnerable individuals or those without an adequate level of awareness about their individual rights. Additionally, civil society is key to ensuring that big data is leveraged to increase the base of opportunities available to all individuals, and equally that digitalization does not increase income and data inequality.

Key elements of data property rights include guidance on the establishment of data ownership, as well as the appropriate level of control on data sharing. Property rights management is a critical part of any data management process; data owners have an interest in understanding how other users will utilize their data, and they also seek to ensure that ethical, legal, and professional obligations are observed. Data property rights are also important from the perspective of equity, as poor legal and governance structures can encourage misuse of information and render vulnerable those who enjoy neither authority nor influence.

A data governance framework can take a number of perspectives on data ownership, either creating a centralized authority responsible for monitoring and enforcing data-sharing regulations or following a more decentralized framework where data sharing resides at the individual level. Within this context, individual preferences can be brought to bear in terms of the utility each individual obtains from maintaining data privacy, relative to the advantages that may accrue from data sharing, such as better matching and personalization of services. The appropriate framework can be implemented by passing on the control of data access protocols to individual users—for example, by enabling individuals to choose their level of access to different types of information generated by their devices. Indeed, privacy can be fragile and fleeting when third parties have access to sensitive data. When users share their data on different online platforms, they reveal signals about other users' preferences, based on shared exogenous characteristics. For example, the preferences of a teenager of a given age in a given school may signal the preferences of that person's circle of friends, perhaps creating negative externalities against those holding information about their preferences shared without their approval and limiting the scope of their control over personal information (Acemoglu et al., 2019).

Market failure may arise due to a lack of data rights. Data are non-rival and excludable,[2] creating incentives to hoard data and allowing the collection of rents as well as the maintenance of a dominant market position. In such situations, significant positive externalities to data sharing that could have a major impact on economic growth may fail to occur. In addition, organizations that collect data lack sufficient incentives to protect the privacy of users who have shared data, given that they do not internalize users' utility from privacy. In such cases, oversharing of data may occur.

With the increasing digitalization of information and improvements in algorithmic analysis, large amounts of data have been collected by companies in order to keep track of individual behavioral patterns, thereby enabling profiling and prediction of future actions. These data are then sold to third-parties for the purposes of monetization, by allowing these entities to market products more effectively to individuals. Enabled by a weak regulatory environment, individual data have been collected from devices and transmitted to companies without the awareness or consent of individuals, and thereafter sold to third-parties. This has prompted responses from various regulatory authorities, with some emphasizing the protection of individual privacy, and others leveraging the power of analytics and data to implement innovations such as social credit systems, which effectively increase surveillance over individuals. However, effective

[2] Non-rivalry of data means that data can be consumed or processed by multiple users without depleting its quality and/or supply. Excludability of data occurs when some groups or individuals are excluded from accessing or using the data.

regulation is complicated by the likelihood that regulatory authorities are behind the curve of innovative activity within major technology firms. Additionally, major technology corporations exert significant influence in shaping the policy environment for the collection and use of data.

An effective data governance structure must promote access that offers benefits to small businesses in particular, and the costs of adhering to the framework must not be prohibitive. Additionally, the realization of a single market for African data must be balanced with incentives for data localization, which is defined as a mandatory administrative or legal requirement indirectly or directly stipulating that data be stored and processed, non-exclusively or exclusively, within a specified jurisdiction. There are aspects of data localization that both advance and detract from effective data governance; although data localization can enhance data privacy and security, it can also inhibit trans-border data flows and lead to various negative consequences attendant to such a slowdown.

2.4.1 Principles

To fulfill its objectives, the data governance framework should adhere to certain central principles, including (a) promoting an agile framework to allow for innovation and experimentation; (b) ensuring accountability of all stakeholders within the data life cycle; (c) establishing standards for data accuracy and quality; (d) developing protocols for the standardization of data, thereby underpinning data quality and enabling interoperability; (e) preserving transparency in the utilization of data; (f) enabling equitable access of public data to all data users; (g) securing non-prohibitive costs of compliance to regulations relating to data; (h) promoting competition in the use and reuse of data; and (i) seeing to it that data sharing at an international level, outside Africa, occurs in full compliance with the rules of Africa's data governance framework.

The data governance framework should be governed by the principle of light-touch regulation, allowing innovation and experimentation, while still possessing the agility to respond quickly to information and implement lessons learned. The large number of use cases for data are unknown, and a restrictive regulatory stance discourages the realization of their full potential. A conducive environment for innovation can be established through regulatory sandboxes, as well as by leveraging the global experiences of other countries as they implement their own frameworks. A conducive framework should also be promoted by regulating data applications appropriately as they are introduced, while still maintaining a principle of widespread availability of public data.

Accountability is a critical component of data governance that should be emphasized when developing both national and regional data governance frameworks. A comprehensive data governance framework covers dimensions of accountability both within and across organizations. Thus, at the macro level, an appropriate data governance framework should contemplate organizational dimensions to guide appropriate design within organizations, while also recognizing aspects of accountability from both a domestic and cross-border perspective.

Within an organization, the framework should promote a holistic view of data governance, as well as integration of data governance practices across departments, by directly and indirectly involved individuals. For example, the establishment of a data council with responsibility for data governance and with representation across departments and at all levels of seniority could formalize the creation of data policies and procedures for implementation and enable effective observation and monitoring. Organizational data governance frameworks currently in place tend to relegate data governance functions to an information technology department, often resulting in ineffective, fragmented, and partial implementation. Building integrated data governance at the organizational level will help build trust among stakeholders.

Data quality standards ensure the accuracy of data, build trust in its use, and allow for consistency in its dissemination. The sheer volume of data produced and analyzed on a daily basis, as well as its exponential growth, underlines the importance of maintaining data quality standards. Uncertainty about the quality of data discourages its use and, if relied upon, may result in erroneous decisions. In the worst case, data can be misused for malicious intent. Quality should, therefore, be ascertained to ensure data are timely, accurate, complete, and consistent—and quality standards should be compatible with other existing rules and regulations, including but not limited to those pertaining to privacy and competition.

Data standardization and the establishment of data protocols are also key enabling factors for supporting an effective cross-border data governance infrastructure and a single market for data. Data standardization contributes to ensuring data accuracy and has important implications for productivity by improving the efficiency of data processes and also encouraging usage. Grannis et al. (2019) investigated the impact of data validation and standardization on accuracy, finding that, in the case of healthcare records, standardization increased the accuracy of healthcare data. What's more, standardization improves the interoperability and portability of data. Interoperability refers to the ability to integrate data sets from different sources, while portability means the ability to transfer or share data without affecting their quality and content. Interoperability standards should be established within countries, with comprehensive coverage over sectors, geographies, and interests, and these standards ought to be underlain with a set of appropriate technical frameworks (such as interfaces for application programming) that individuals can leverage in promoting interoperability.

Transparency should be exercised in all stages of the data governance process. Data-related decisions and data processes should be communicated across all data users to ensure a clear understanding of data-handling processes and to allow users to know how their information is obtained and deployed. This will, in turn, build trust within the data governance process, encouraging users to participate within the framework and providing users with the necessary information to exercise their rights with regard to the availability and use of their data. Moreover, access to data should be provided on an equitable basis across all categories, including both private and private data, and this access should be universal and independent of data producers' and users' economic status or market power. Access costs should be low enough to allow users widespread participation, and gaps in the enabling infrastructure within countries and across the continent should be closed to promote more equitable basis. Further, the costs of compliance for participation in the data economy cannot be prohibitive, an especially relevant concern in the African context, where the vast majority of the private sector consists of small businesses with limited capacity—businesses whose active participation will require a conducive framework that promotes competition in the use and reuse of data and drives innovative activity. Finally, to ensure that the rights of African citizens are adequately protected by third parties, the data governance framework should ensure that data sharing at an international level is done with those countries and regions who comply fully with the rules established within the African data governance framework.

2.4.2 Infrastructure

Without an appropriate enabling environment, including physical and human capital infrastructure underpinning its implementation, a data governance framework cannot reach its potential. This framework also calls for institutions to secure the right cultural, legal, policy, regulatory, organizational, institutional, and technical environment to ensure that all data users can effectively and efficiently extract value from data, as well as enforcement mechanisms of national and regional data governance frameworks, where regulatory authorities enjoy administrative, agency, and financial autonomy, to ensure the security and privacy of data.

The appropriate enabling environment should invest in infrastructure that lowers data access costs while increasing its quality, with particular attention to strengthening each country's infrastructure while also addressing access gaps both within and across countries. Investment should target a rapid increase in the digitization of data, promote the sharing of currently existing high-value public data sets, improve access to data at the household level, and raise the quality of public internet connections. Additionally, the value of the enabling infrastructure should be established by quantifying the impact of increased digitization, data access, and use both within countries and across the continent. The enabling environment depends on established protocols and application programming interfaces (APIs) to promote the standardization and transfer of data both nationally and regionally. Finally, these frameworks, as well as their productive deployment, hinge on investment in a well-trained labor force.

2.5 Conclusion and Recommendations

Data sharing can vastly improve living standards through improvements in productivity. Data are also non-rival and partially excludable, meaning they can be reused infinitely without degradation. With knowledge building upon knowledge, returns to the utilization of data could increase in scale; the efficient utilization of data is thus critical to productivity growth.

Free data exchange has spawned unprecedented opportunities to people across the globe, creating jobs and industries, facilitating increased mobility, and ultimately raising standards of living. Policymakers aspire to responsible and safe data use to improve the lives of the people they serve, while minimizing the misuse or exploitation of data. To this end, a clearly defined data governance framework integrated with a data strategy is necessary to establish data sovereignty and buttress Africa's competitiveness and cross-country collaboration during the digital age. Implementing an effective data governance framework will preserve the availability, usability, integrity, and security of data across the continent—and such a framework will be both served and safeguarded by a developed data infrastructure, technical protocols, laws and regulations, and institutions suited to promoting the safe and trustworthy use of data while respecting privacy.

The central aim of a pan-African data governance infrastructure is to help bring about a single market for data, thereby allowing the creation, use, and reuse of data by individuals across the continent and spurring economic growth and development while still protecting the rights of data subjects. To fulfill its objectives, this framework should adhere to certain central underlying principles and operate within a light-touch and agile regulatory framework that encourages innovation. These principles include ensuring accountability, maintaining data accuracy and quality, and facilitating interoperability and standardization of data; but just as much, this framework must afford equitable access of data and keep the costs of compliance low, so as to promote competition.

Some key areas require additional research if the maximum utility of a data governance framework is to be enjoyed, such as (a) identifying and prioritizing key strategic interests across the continent that will benefit the most from the implementation of a data governance framework, as well as quantifying the value of the framework; (b) determining strategies to increase the pace of digitization of offline public data sources, while also increasing access to already digitized public data; and (c) mapping infrastructure gaps and cost-of-access disparities at the subnational and cross-country level and then resolving these inequities. Finally, further research must explore how to create a regulatory framework that best implements the principles of effective data governance. This task will include, for example, research to define an effective accountability framework both at the firm and national levels, as well as the development of protocols to allow data transfer and interoperability.

References

Acemoglu, D., Makhdoumi, A., Malekian, A., & Ozdaglar, A. (2019). *Too much data: Prices and inefficiencies in data markets* (NBER Working Paper Series). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w26296/w26296.pdf

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis, 126*(1), 3–9. https://softwarezen.me/wp-content/uploads/2018/01/datawisdom.pdf

Alhassan, I., Sammon, D., & Daly, M. (2016). Data governance activities: An analysis of the literature. *Journal of Decision Systems, 25*(sup1), 64–75. https://doi.org/10.1080/12460125.2016.1187397

Beraja, M., Yang, D., & Yuchtman, N. (2020). Data intensive innovation and the state: evidence from AI firms in China. NBER Working Paper No. 27723.

Desjardins, J. (2019, April 17). How much data is generated each day? *World Economic Forum*. https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-daycf4bddf29f/

Fourie, J. (2016). The long walk to economic freedom after apartheid, and the road ahead. *IDEAS.* https://ideas.repec.org/p/sza/wpaper/wpapers267.html

Grannis, S., Xu, H., Vest, J., Kasthurirathne, S., Bo, N., Moscovitch, B., Torkzadeh, R., & Rising, J. (2019). The effect of data validation and standardization on patient matching accuracy. *Journal of the American Medical Informatics Association, 26*(5), 447–456. https://doi.org/10.1093/jamia/ocy191

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM, 53*(1), 148–152. https://www.researchgate.net/publication/220426163_Designing_data_governance

Kouper, I., Raymond, A., & Giroux, S. (2020). An exploratory study of research data governance in the U.S. *Open Information Science, 2020*(4), 122–142. https://doi.org/10.1515/opis-2020-0010

Marr, B. (2015, February 25). A brief history of big data everyone should read. *World Economic Forum*. https://www.weforum.org/agenda/2015/02/a-brief-history-of-big-data-everyoneshould-read/

Micheli, M., Ponti, M., Craglia, M., & Suman, A. B. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society, 7*(2), 1–15. https://doi.org/10.1177/2053951720948087

Otto, B. (2011). *A morphology of the organization of data governance*. European Conference on Information Systems (ECIS 2011 Proceedings). https://www.researchgate.net/publication/221407900_A_morphology_of_the_organisation_of_data_governance

United Nations Congress on Trade and Development. (2021). *Trade and Development Report*. United Nations Congress on Trade and Development. https://unctad.org/webflyer/tradeand-development-report-2021

Vidal, M. F., & Medine D. (2019). *Is data privacy good for business?* Consultive Group to Assist the Poor. https://www.cgap.org/sites/default/files/publications/2019_12_Focus_Note_Is_Data_Privacy_Good_for_Business.pdf

Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all—A contingency approach to data governance. *ACM Journal of Data and Information Quality, 1*(1), 1–27. https://www.alexandria.unisg.ch/67793/1/a4-weber_external.pdf

World Bank. (2021a). *World development report 2021: Data for better lives*. World Bank Group. https://doi.org/10.1596/978-1-4648-1600-0

World Bank. (2021b). *Minimum data consumption: How much is needed to support online activities, and is it affordable*? World Bank Group. http://hdl.handle.net/10986/35149

Zimmermann, K. A. (2017, September 6). History of computers: A brief timeline. *Live Science*. https://www.livescience.com/20718-computer-history.html

CHAPTER 3

A Value Chain Approach to Data Production, Use, and Governance for Sound Policymaking in Africa

Zachary Mwangi Chege and Peter Maina Wanjohi

3.1 Introduction

Myriad of challenges have been cited as the reasons behind underdevelopment in Africa. Some of the most commonly identified challenges that have been linked to slow development in the continent include bad political leadership, poor corporate governance, insufficient capital, low infrastructure development, corruption, slow uptake of technology, and low intra-African trade. Moreover, the level of implementation of governments' strategies and programs has largely been rated poorly and also strongly linked to poor performance of many economies in Africa. Many of the African countries' official statistics have also been enjoined in the blame for being shallow in coverage and failure to correctly reflect the actual situation. Consequently, official statistics have largely been viewed as of low quality, low precision, untimely, incomplete, poor accessibility, and inconsistent. The statistics from most of the African countries are

Z. M. Chege (✉) • P. M. Wanjohi
Kenya National Bureau of Statistics, Nairobi, Kenya

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8_3

therefore regarded with mistrust and occasionally judged as insufficient in informing development needs for the continent.

In light of the necessity for accelerating development in Africa, there is need to strengthen the state of official statistics. Quality and reliable data are required to inform policy formulation at country's level as well as to monitor and evaluate progress toward internationally agreed programs such as Agenda 2063 and the Sustainable Development Goals (SDGs) among many others. It is, therefore, imperative to relook at the way that statistics are produced in Africa, with a view of addressing their quality as well as efficiency and effectiveness of the production process. Globally, many countries have enhanced the use of alternative data sources with the aim of cutting down on costs of production and improving on timeliness. Granted, many African countries have made spirited efforts along similar lines and positive results are evident, but many impediments still exist. This research recommends the need for National Statistical Offices (NSOs) to encompass value chain approach to data production and heighten the uptake of statistics for evidence-based decision-making and better governance of statistical process to address data quality for sound policymaking in Africa.

According to Porter (1985), value chain is a sequence of activities that are required to be undertaken by a firm or an organization in order to create a product or service. A firm or an organization performs a number of discrete activities in production of a good or service, starting from designing a product or service up to delivering it to customers. The main objective of value chain analysis is to improve efficiency of a firm in production of products or services by enabling it to deliver maximum value at the lowest possible cost.

In relation to statistics, the value chain entails the process of identifying the data needs, designing data collection instruments, testing data collection instruments, collecting the data, processing the data, analyzing the data, and disseminating the data for its final use. The data value chain also addresses uptake of data by potential users as well as the resultant impact of the data usage. This is deemed important in order to evaluate if the statistics are meeting their intended purposes. National Statistical System (NSS) in every country should ensure there is constant feedback between producers, users, and other stakeholders for perpetual improvements of quality of official statistics. According to European Commission, Statistics Sweden and Eurostat (1999), quality statistics are defined as those statistics that are accurate, relevant, consistence, comparable, timely, reliable, complete, impartial, and accessible to all users.

Reviewing data value chain also calls for addressing inherent data governance weaknesses that have persistently undermined development of official statistics to requisite standards in Africa in accordance with best international practice. In this respect, the research made recommendations on practical ways of managing the quality, availability, integrity, validity, interpretability, consistency, and accessibility of statistics to enhance trustworthiness and uptake of official statistics in Africa.

3.2 Methodology

This paper mainly used the review of existing literature to identify and examine gaps and challenges in regard to production and use of official statistics in Africa. Also examined using existing literature is the governance structure in production and use of official statistics in Africa where underlying gaps and weaknesses were identified. The research identified the current practice by African NSOs in production of official statistics, examined the uptake of official statistics produced in Africa, examined the governance structure of NSOs in Africa, and compared these practices with the best international practices. On the basis of these gaps, challenges, and weaknesses, value chain analysis was conducted on data production, use, and governance in Africa, and recommendations were made for future improvements.

Three case studies on Australia, Rwanda, and South Africa were conducted where their practices in relation to data production, use, and governance were reviewed. The study focused on these three countries since they have excelled in specific areas in regard to production and use of official statistics. These three case studies were used as a benchmark for enriching recommendations on how other African countries can improve value chain in production of official statistics, utilization of these statistics, and data governance. The case studies targeted areas on quality dimensions, efficiency in production process, effectiveness of NSS, use of alternative data sources (e.g., big data and citizen-generated data), and dissemination of official statistics.

3.3 Production and Dissemination of Official Statistics in Africa

According to United Nations Handbook on Management and Organization of National Statistical System (2021), official statistics are defined as statistics that are developed, produced, and disseminated by countries' national statistical systems in conformity with United Nations Fundamental Principle of Official Statistics (UNFPOS), globally agreed statistical standards, codes, and recommendations as well as relevant national laws and programs. Fundamental Principles of Official Statistics comprises of ten principles, which provide the basic rules for the production of official statistics (United Nations Economic and Social Council, 2013). The ten principles are listed below, and countries including African countries under the United Nations should follow these principles in production and dissemination of official statistics (Table 3.1). Table 3.2 presents sources of data used in the production of official statistics.

African NSOs have been relying on traditional sources of data (administrative, surveys, and censuses) to produce official statistics. According to African Data Revolution Report 2016 (UNECA et al., 2016), data revolution has led to new (non-traditional) sources of data that can be used to produce official statistics. These new sources of data include satellite data, remote sensing data, citizen-generated data, and big data, among others. United Nations (2021) defines citizen-generated data, "as data produced by non-state actors under the active consent and participation of citizens to primarily monitor, demand or drive change on issues that affect them

Table 3.1 Fundamental principles of official statistics

Fundamental principles of official statistics	
Principle 1: Impartiality, relevance, and equal access	Principle 6: Confidentiality
Principle 2: Professional standards, scientific principles, and professional ethics	Principle 7: Legislation
Principle 3: Transparency and accountability	Principle 8: National coordination
Principle 4: Prevention of misuse	Principle 9: Use of international standards
Principle 5: Sources of official statistics	Principle 10: International cooperation

United Nations; United Nations Economic and Social Council (2013)

Table 3.2 Sources of data used in the production of official statistics

Three main sources of data used in the production of official statistics		
1. Administrative Data	*2. Surveys/Censuses*	*3. Other Data Sources*
Official statistics are largely based on administrative data. Administrative data is generated by government ministries, departments, and agencies in course of their administrative operations. Therefore, this secondary source of data is availed to NSOs for statistical purposes.	Surveys and censuses are other traditional sources of data that are commonly used in the production of official statistics. Surveys and censuses collect individual data directly from the population using a statistical methodology and the collected data is used solely for statistical purposes.	Other data sources include commercial data streams from businesses, sensors data, geospatial data, and social media data, among others United Nations (2021).

United Nations (2021), "The Handbook on Management and Organization of National Statistical Systems: 4th Edition of the Handbook of Statistical Organization," United Nations, New York

directly." These new sources of data can be used as a secondary source for the production of official statistics. Statistical surveys are relatively cheap and results can be released faster compared to censuses. However, surveys and censuses are expensive, labor-intensive, and time-consuming compared to other data sources such as administrative data. One of the major challenges faced by NSOs in Africa in using administrative data is that some sources of these data provide inaccurate, incomplete, inconsistent data resulting in low-quality data (United Nations, 2021; UNECA et al., 2016) (Table 3.3).

Existing evidence indicate that, in Africa, education enrollment obtained from administrative sources is usually higher than that obtained from surveys. These discrepancies are mainly attributed to resource allocations to learning institutions based on the number of learners. On civil registration and vital statistics, existing evidence indicate that 46 countries in Africa have incomplete civil registration systems for registration of births and deaths. As a result, about 83 percent of Africans reside in a country without a well-functioning system for birth registration, while 87 percent of deaths occur in countries without a complete system for death registration. Civil registration provides a system through which a country continuously captures and keeps complete records of births and deaths.

Table 3.3 Main reasons for provision of low-quality administrative data

Main reasons for provision of low-quality administrative data
① Administrative data is primarily collected for administrative purposes and therefore, *the human capital involved may not have adequate technical knowledge of producing data for statistical purposes.*
② Most African countries *do not have appropriate administrative data systems* such as civil registration and vital statistics systems for managing administrative information, and *where these systems exist, they are weak*.
③ Since the data is generated mainly for administrative purposes such as for budget allocations, in some cases that *data is either over reported or under reported*.
④ Some government authorities or public bodies may be *reluctant to provide accurate and timely data since it may reflect negatively on their performance* in discharging their duties or delivering services to the public.

United Nations (2021), PARIS21 and The Mo Ibrahim Foundation (2021), UNECA et al. (2016)

These data are then used by policymakers for planning and monitoring provision of public goods and services on education, health, employment, social protection, housing, and electoral processes, among others. Therefore, policy decisions made based on these incomplete statistics results to inadequate provision of public goods and services thus, unsatisfactory outcomes (PARIS21 and The Mo Ibrahim Foundation, 2021; UNECA et al., 2016).

There is increasing demand for timely and disaggregated statistical indicators, and despite reduction of their budgets, NSOs, especially in Africa, are faced with challenges of producing these indicators for monitoring global, regional, and national goals (United Nations, 2021). In order to get these indicators, African NSOs are commonly funded by development partners to conduct surveys and censuses. For example, between 2010 and 2014, donors' annual commitments to support statistics across African countries increased from USD 6.5 million to USD 26.7 million, and it stood at USD 26.0 million in 2017. Even though the indicators generated using donor-funded statistical activities provide useful official statistics, this type of arrangement may result in production of data that is not aligned with national priorities, which may lead to inconsistent and unsustainable production of data (PARIS21 and The Mo Ibrahim Foundation, 2021). The above review indicates that there exist gaps and challenges in regard to basic statistical data, which are used in the production of official statistics in Africa. Figure 3.1 provides the definition of big data and their use in production of official statistics.

Big Data

Big data is one of new sources of data that can be used in production of official statistics.

Big data is defined as data generated in **large volumes, varieties and velocity**.

It is generated through **business transactions, communication devices, phone logs, social media, sensors and web scrapping among others**.

Non-traditional sources of data such as **big data has the potential to complement, supplement or partially replace traditional sources of data** due to their high population coverage and usage in daily life.

However, there is little evidence on use of big data in production of official statistics.

The technology required to process these new sources of data is available and is continuously improving; the main challenge facing NSOs is gaining access to the data and having the required technical capacity and skills to secure, process and analyze the data.

Fig. 3.1 Big data (United Nations, 2021; UNECA et al., 2016)

A review conducted by UNCTAD in 2020 on data protection and privacy legislation across the world indicates that in Africa, 28 countries (52 percent) have data protection and privacy legislations, in 9 countries (17 percent) the legislations are in the drafting stage, 13 countries (24 percent) have no legislation, while there were no data for 4 countries (7 percent) (Fig. 3.2).

African countries need to enact and implement data protection laws and regulations, which will ensure that data in the custody of NSOs and other organizations across NSS are treated with utmost confidentiality, in addition to putting in place strong data security mechanisms that protect data in their custody against the ever-evolving digital attacks (cyberattacks).

In order to guide countries on use of big data in the production of official statistics, United Nations Statistical Commission established a Global Working Group on Big Data for Official Statistics. The group has created a Global Platform through which countries can gain access to big

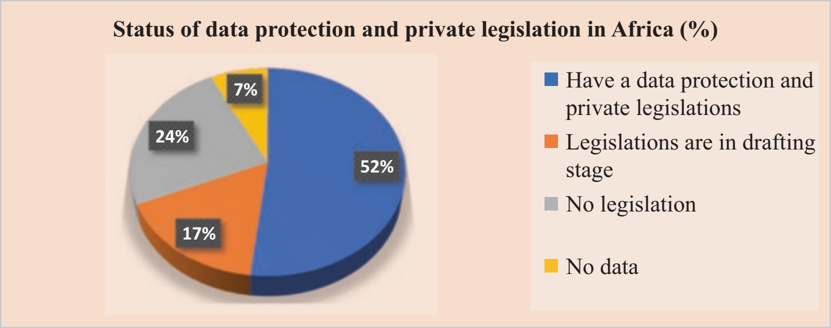


Fig. 3.2 Status of data protection and private legislation in Africa. From Data Protection and Private Legislation Worldwide, by United Nations Conference on Trade and Development (UNCTAD), ©2020 United Nations. Used with the permission of the United Nations. https://unctad.org/page/data-protection-and-privacy-legislation-worldwide (accessed 02 November, 2021).

data held by multinational corporations and modern methods used in processing and analyzing this type of data. In addition, this platform creates an environment for capacity development activities in new areas, which include privacy-preserving methods, artificial intelligence, data science, and machine learning statistics (United Nations Statistical Commission, 2020).

During Kigali Declaration of 2019, it was agreed that through the Global Platform, all countries under the United Nations should be provided with the required technological infrastructure for processing and analyzing big data, and emphasis was made on the need to support developing countries for them to access global data sets, modern technologies, and services. Rwanda is one of the African countries that is set to host the regional hub on big data, which, in collaboration with regional and international agencies, will support capacity building for African countries, to enable them utilize big data in the production of official statistics (United Nations Statistical Commission, 2020).

In Africa, the big data community is relatively small, but it is rapidly growing. African NSOs can access existing technological infrastructure (hardware and software) required to process huge data sets; however, there are two main challenges: first, getting financial resources to acquire the required technology and, second, getting the human capacities and skills required to process and analyze the data. In order to benefit from non-traditional data sources, African countries need to address issues limiting access, invest in modern technology, and recruit and retain staff with technical skills. Enacting legislation that protects personal data and guarantees data privacy will enable African countries to gain access and benefit from big data and other non-traditional data sources (PARIS21 and The Mo Ibrahim Foundation, 2021; UNECA et al., 2016).

These new sources of data, especially big data, require new techniques and methods to extract and generate statistical indicators that may be used as official statistics. Data from these new sources are not mainly generated for production of official statistics; therefore, data and statistical indicators generated from these sources need to be validated using new methodological techniques for them to be used as official statistics. This requires new skills sets. Research indicates that African NSOs are characterized by under-staffing, low-skilled workforce, and high staff turnover. This affects the quality of statistics produced (United Nations, 2021; PARIS21 and The Mo Ibrahim Foundation, 2021; UNECA et al., 2016).

African NSOs need to establish partnerships with big data sources, ICT agencies, data protection agencies, monitoring and evaluation agencies, data engineers, data scientists, data miners, and data analysts, among others, which will facilitate continuous interactions in relation to the use of new (non-traditional) data sources in the production of official statistics. For example, the International Telecommunication Union (ITU) in collaboration with Communication Authority of Kenya, Kenya National Bureau of Statistics, Mobile Network Operators, and Internet Service Providers conducted a pilot study on measuring information society using big data. The pilot study found out that big data has the capability to provide quality data that can be used for policymaking. Issues of confidentiality and data formats are some of the challenges that were experienced during the study (ITU, 2016).

3.4 Value Chain Theory in Data Production, Use, and Governance

Value chain, as explained earlier, is a stream of activities that are performed by firms or organizations to create a product or service. According to Porter (1985), value activities are classified into two broad components: primary activities and support activities. Primary activities are the activities directly undertaken by a firm in the creation of a product or service and are broadly divided into five categories: inbound logistics, operations, outbound logistics, marketing and sales, and service. Support activities are broadly divided into four categories: procurement, technology development, human resource management, and firm infrastructure.

The general value of official statistics is that, first, they provide an indispensable component in the information system of a democratic society. They serve the government, the economy, and the general public with data on the economic, demographic, social, and environmental situation. Official statistics enable decision-makers either in government or private sectors, including the general public, to make decisions based on high-quality information, which leads to better outcomes. Second, quality official statistics provide an avenue for citizens to hold government and private organizations to account, thus enhancing transparency and accountability. Third, researchers use official statistics as a complete source of evidence to do research and analysis, which leads to innovation and improved socio-economic outcomes (United Nations, 2021; UNECE, 2018).

Statistics are not solely produced by NSOs within the National Statistical Systems; there exist other organizations both private and public that are outside the official statistics system that produces statistics. However, official statistics have a comparative advantage since they are generated scientifically on the basis of international standards and methods and quality criteria provided for in the United Nations Fundamental Principles of Official Statistics, which then guarantees professional independence. A cost-benefit analysis of official statistics conducted in developed countries such as the United States, Australia, New Zealand, and the United Kingdom indicates that the benefits of official statistics outweigh the cost of their production in that they are cheap and efficient utilization of resources (United Nations, 2021; UNECE, 2018). This indicates that countries are capable of delivering maximum value of official statistics at the lowest possible cost (Table 3.4).

3.5 Value Chain Approach to Production and Dissemination of Official Statistics in Africa

This section reviews the value chain approach to production and dissemination of official statistics in Africa, where both the demand and supply sides of official statistics have been discussed. The demand side involves the identification of data needs for different users and obtaining their

Table 3.4 Value chain analysis in official statistics

Value chain analysis in official statistics	
It involves examination of a stream of primary activities and support activities by NSOs in data production.	
Primary activities	*Support activities*
Identification of data needs	Financial and human resources
Designing and testing data collection instruments	Statistical capacity
Data collection, processing, analysis, and dissemination	Procurement processes
Use of official statistics	Technologies used in the production of official statistics
	Governance structure of NSOs; comprises support activities undertaken to guide and oversee the implementation of statistical activities.

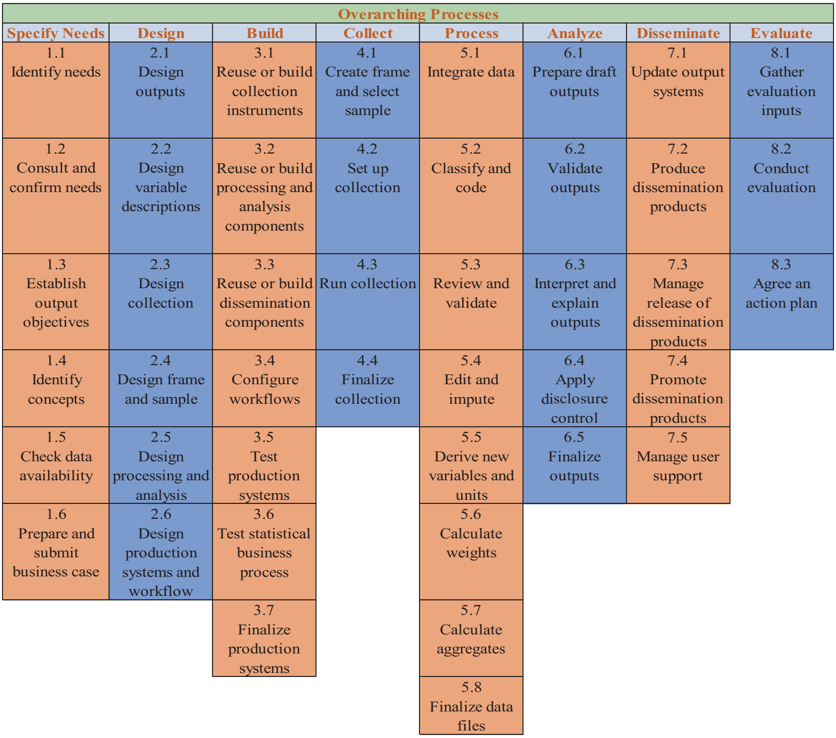
United Nations (2021), National Institute of Statistics of Rwanda (2019), Statistics Botswana (2015)

feedback in regard to use of official statistics. The supply side involves data collection, processing, analyzing, and dissemination. The use of official statistics by different users and data governance are also discussed. The value chain analysis of statistical activities in Africa has been examined using both Porter's Value Chain Model and Generic Statistical Business Process Model (GSBPM).

The GSBPM provides a sequence of all necessary activities and steps that need to be undertaken in the production of statistics (UNECE, 2019); these activities constitute primary activities under the Porter's Value Chain Model. However, an organization requires financial resources, legal framework, leadership, human capital, and suitable technology, among others, in order to carry out all the necessary activities under GSBPM. These resources provide support for carrying out all the necessary activities, and therefore, they constitute support activities under the Porter's Value Chain Model. Figure 3.3 presents level 1 and level 2 of the GSBPM.

3.5.1 *Identification of Data Needs for Different Users*

Production and dissemination of official statistics should be based on users' demand for the statistics. Producers of official statistics, more so



Overarching Processes							
Specify Needs	Design	Build	Collect	Process	Analyze	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Reuse or build collection instruments	4.1 Create frame and select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult and confirm needs	2.2 Design variable descriptions	3.2 Reuse or build processing and analysis components	4.2 Set up collection	5.2 Classify and code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Reuse or build dissemination components	4.3 Run collection	5.3 Review and validate	6.3 Interpret and explain outputs	7.3 Manage release of dissemination products	8.3 Agree an action plan
1.4 Identify concepts	2.4 Design frame and sample	3.4 Configure workflows	4.4 Finalize collection	5.4 Edit and impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing and analysis	3.5 Test production systems		5.5 Derive new variables and units	6.5 Finalize outputs	7.5 Manage user support	
1.6 Prepare and submit business case	2.6 Design production systems and workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalize production systems		5.7 Calculate aggregates			
				5.8 Finalize data files			

Fig. 3.3 The Phases (level 1) and Subprocesses (level 2) of the GSBPM. Source: (UNECE, 2019). From Generic Statistical Business Process Model (GSBPM), by United Nations Economic Commission for Europe (UNECE), ©2019 United Nations. Used with the permission of the United Nations.

National Statistical Offices, should plan and implement their statistical activities that satisfy the needs of different users. This means that NSOs should carry out demand-driven statistical activities. Users of official statistics include national and regional governments (policymakers, lawmakers, civil servants across different government offices); international and regional organizations; businesses; media; academic, research, and education community; non-government organizations (NGOs); and the general public (United Nations, 2021).

Identification of data needs for different users is the first key activity in the data value chain. At this level, NSOs should engage different users of official statistics and identify their data needs, and this should form the basis of their statistical programs and activities.

The establishment of systems for the identification of data needs for different users is still very low among African NSOs. For example, Kenya NSO noted that one of the challenges it faced in the implementation of 2013–2017 strategic plan was synchronizing the collection and compilation of data with the interest of various stakeholders (Kenya National Bureau of Statistics, 2018). In order to overcome this challenge, the NSO noted that during the implementation of 2018–2022 strategic plan, it needs to undertake an analysis of data users' needs and align production and management of statistical activities with user requirements. In addition, the NSO noted that adequate engagement of stakeholders in planning and implementation of statistical operations is important for identification of their requirements and enhancing ownership and acceptance of results.

South Africa NSO, in its strategic plan for 2020/21–2024/25, notes that there are new demands for statistics such as statistics on the digital economy, well-being, and climate change. These new demands have led to diverse data requirements and changed the structure of data users. Traditional sources could not meet these new demands. The agency, therefore, noted that to meet these new demands, it needs to conduct in-depth data analysis by integrating data from different sources, seeking collaborations with other data producers in the analysis of existing data, and exploring the use of alternative sources of data (Statistics South Africa, 2020).

Ghana NSO in its strategic plan for 2020–2024 indicates that it will conduct an assessment survey and hold focus group discussions to identify statistical needs for stakeholders in the private sector (Ghana Statistical Service, 2020). Rwanda NSO, on the other hand, conducts biennial user satisfaction survey (USS) that collects feedback on preferences of different users of official statistics. These surveys provide important feedback that help the statistical agency in planning and implementing data production activities (National Institute of Statistics of Rwanda, 2019).

These examples show that only few countries in Africa have created mechanisms to ensure continuous interactions of data users and producers of official statistics with the aim of enhancing the production of official statistics that meet user requirements and increasing the demand and uptake of official statistics. African NSOs need to appreciate the fact that the process of identifying data needs for different users should be part of their routine activities. Identification of data users' needs, referred to as specify needs phase in the GSBPM, is the first phase of the eight phases identified in the overall production of official statistics as shown in Fig. 3.3 (United Nations, 2021; United Nations Economic Commission for Europe, 2019). The choice of source(s) of data to be used by an NSO should be on the basis of its ability to meet the required statistical standards and data users' needs at the lowest cost possible.

3.5.2 Designing Data Production Processes

Once the needs of different users are identified, the next step is the design phase under GSBPM, as shown in Fig. 3.3, where NSOs are required to define the required statistical outputs that need to be produced, define variables to be collected, determine the most suitable data collection tools and methodologies, and develop and define all the data production operational processes (United Nations, 2021; United Nations Economic Commission for Europe, 2019).

National statistical agencies in Africa are usually faced with challenges while designing processes of data production to meet the needs of different users. First, data gaps in Africa are a common phenomenon, especially on civil registration and vital statistics (registration of births and deaths, migrants' records), health, poverty levels, agriculture, climate change, and digital economy, among others (PARIS21 and The Mo Ibrahim Foundation, 2021; UNECA et al., 2016). Some of these data gaps arise from weak administrative data systems, lack of these systems altogether, or delays/failure to conduct various surveys/censuses. Due to data gaps, some data users such as private sector, civil-society organizations, academia, and other non-state actors resort to independent production of statistical information for use in their operations. This is a positive outcome arising from data gaps, and therefore, statistical agencies should work together with other data producers to supplement official statistics. This can be achieved by establishing partnerships between NSOs and non-state actors' data producers, which will ensure that these actors produce data using internationally accepted standards and methods.

Surveys provide alternative source of data in cases where administrative data are missing or incomplete. This is a common phenomenon in Africa, and NSOs use surveys/censuses to obtain data, which can be generated from administrative data such as data on births, deaths, and health. Therefore, in order to fill the data gaps, these countries resort to conducting surveys to supplement and complement data from administrative sources. This increases surveys sample sizes and the number of survey/census questions, thereby increasing the cost of surveys/censuses, which negatively affects the quality of the data collected mainly due to respondents' fatigue and misreporting by fieldwork staff.

For example, first, Population and Housing Censuses conducted in Malawi and Kenya in 2018 and 2019 (Government of Malawi, 2019; Kenya National Bureau of Statistics, 2019), respectively, had several questions covering different thematic areas, namely, demographic (age, sex, relationship, marriage, fertility, mortality, disability), socioeconomic (marital status, religion, ethnicity, education, labor participation, ICT, agriculture), and household and housing characteristics (floor/roof/wall materials, tenure status, dwelling units, sources of water, energy, and waste disposal). On the other hand, United States of America 2020 census had only nine questions covering the following thematic areas: demographic (age, sex, relationship), socioeconomic (ethnicity, race, mobile phone ownership), and household and housing characteristics (tenure status) (United States Census Bureau, 2020).

Second, most African countries develop their household sampling frames based on the previous population and housing censuses. In 2010 round of population and housing censuses, only 47 African countries conducted censuses, while in 2020 round of censuses (2015–2024), as of September 2020, only 11 African countries had conducted their censuses (Bruno et al., 2020; UNECA, 2020a, b, c). Some African countries were set to conduct their censuses in 2020, but they postponed the exercise to 2021 and 2022 due to COVID-19 pandemic, where finances initially allocated to census activities were diverted to address the pandemic (UNECA, 2020a, b, c; United Nations Statistics Division, 2020). Therefore, household sampling frames currently being used in Africa are mainly based on 2010 round of population and housing censuses, implying that the samples drawn from these frames may not be representative of the entire population.

Third, African countries use different methods and standards in the production of statistics, which makes it difficult to compare data from different countries. This challenge arises from the use of different methodologies in various statistical activities, use of different concepts and definitions, and weak data technology and related infrastructure (UNECA et al., 2017). This challenge is recognized under the Strategy for the Harmonization of Statistics in Africa 2017–2026 (SHaSA 2), where

African countries aim at transforming existing statistics to make them comparable across countries, and harmonizing the methods and standards used in data production (African Union Commission et al., 2017).

As outlined in the Fundamental Principles of Official Statistics and African Charter on Statistics, national statistical agencies across Africa should coordinate the national statistical system in order to achieve an efficient and consistent statistical system and ensure quality and comparable statistical information (United Nations Economic and Social Council, 2013; African Union Commission, 2009). The coordination role of African NSOs should be anchored in law in order to enable them to play a key role in nurturing partnership, harmonization, and coordination of the national statistical system (UNECA et al., 2016). In Africa, the capacity of national statistical agencies to coordinate NSS is inadequate, which leads to low statistical capacity across NSS.

African countries need to strengthen the coordination capacity of NSOs so as to improve quality of official statistics and reduce the cost of producing data in Africa by taking advantage of data produced by other data producers to fill existing data gaps (PARIS21 and The Mo Ibrahim Foundation, 2021). This will, in turn, reduce the number of questions in surveys/censuses, thus reducing the cost of carrying out surveys/censuses and improving the quality of the data collected. In addition, the use of reliable administrative data will result in the timely production of statistical indicators for use by policymakers and other users, which will then generate better outcomes for the society.

3.5.3 Development and Testing of Data Collection Instruments and Other Key Processes

Once data production processes have been designed, the next phase in the production of official statistics is the development and testing of the data collection instruments, and development and testing of data processing, analysis, and dissemination processes. Other technical processes, logistics, and administrative arrangements related to the production of the required statistical information are also developed and tested in this phase. This phase is referred to as the build phase under the GSBPM, as shown in Fig. 3.3. This phase in the production of statistics is carried out once or when there is change in technology or methodology for regularly generated statistical outputs (United Nations, 2021; UNECE, 2019).

Most African countries have embraced the use of information and communication technology (ICT) in the production of official statistics. This is evidenced by the adoption of computer-assisted personal interview (CAPI) technology in data collection and transmission by a number of countries in Africa. The use of CAPI technology and other modern mobile technologies in data collection and processing improves the quality of data and reduces the time taken to collect, process, analyze, and disseminate the data. The main challenge for adopting CAPI technology in the production of statistics is the cost of the equipment and ICT infrastructure required (United Nations, 2021; Kenya National Bureau of Statistics, 2018, 2019). Malawi and Kenya are examples of two African countries that conducted their Population and Housing Censuses in 2018 and 2019, respectively, using CAPI technology. This enabled these countries to release their first population and housing census report in a record of three months after census enumeration for Malawi and two months after census enumeration for Kenya (Government of Malawi, 2019; Kenya National Bureau of Statistics, 2019). In 2017, about 57 percent of African countries had decided to use CAPI in their census taking during the 2020 round of population and housing censuses (Bruno et al., 2020).

Challenges facing African countries in regard to the use of modern technologies in the production of official statistics include inadequate finances, inadequate human capital, inadequate technical skills, inadequate technical capacity, and poor mobile network coverage that affects data transmission, among others (PARIS21 and The Mo Ibrahim Foundation, 2021; UNECA et al., 2016). CAPI data collection tools and related ICT infrastructure need to be tested severally and piloted before they are deployed for main data collection. Due to the challenges mentioned above, African national statistical agencies experience difficulties in testing and piloting CAPI data collection tools and related ICT infrastructure.

For example, Bruno et al. (2020) note that a number of African countries require external support in the mobilization of the required financial resources and technical assistance to strengthen their capacity to adopt modern technologies in order for them to conduct CAPI censuses in 2020 round of population and housing censuses. In addition to 2020 round of censuses, other statistical activities in Africa have also been negatively affected by COVID-19 pandemic due to disruptions of work arrangements and data collection activities (United Nations Statistics Division and World Bank, 2020).

African countries have not standardized data collection tools due to the use of silo approach in the production of official statistics across NSS. Silo approach in data ecosystems may lead to duplication of efforts and data errors, thus limiting data systems interoperability and interrupting linkages across different statistical processes (United Nations, 2021; National Institute of Statistics of Rwanda, 2019).

African NSOs need to leverage their adoption of modern technologies in the production of statistical information to develop and adopt an integrated data production system, where processes for different statistical activities are standardized and linked together through development and application of similar standards, methods, and modern technologies. This will achieve efficiency in the production of official statistics by the use of standardized data collection and processing tools, lowering the costs of conducting statistical activities such as surveys and censuses and reducing the time taken to produce statistical outputs. This system will enhance efficient and better use of available resources (United Nations, 2021).

Organizations across African Statistical Systems need to collaborate among themselves in order to improve production and use of statistical information in Africa (United Nations Economic and Social Council, 2013; African Union Commission, 2009). One of the main objectives of SHaSA 2 is to improve statistical coordination and collaboration among international partners, regional organizations, continental organizations, and national statistical institutes (African Union Commission et al., 2017). This will enable African countries to utilize available resources more efficiently. For example, Kenya is supporting African countries with mobile devices (locally assembled and used in 2019 census) and technical capacity for their respective census undertaking (Republic of Kenya, 2021).

On funding of statistical activities in Africa, a study reviewing financing of national statistical systems in Rwanda, Ethiopia, and Philippines highlighted the following factors as crucial in the establishment of well-funded statistical systems. The first factor is increase in demand for statistics required to monitor and evaluate both national and international priorities. This has helped NSOs in these countries to get sufficient and sustainable funding from both national budgets and development partners. Second, there is high level of political interest and support for statistics in these three countries, which has enabled governments to manage relations with donors and efficiently implement development plans. Third, donors in these countries have aligned their financial and technical support for statistics with government strategies. Fourth, the legal framework governing operations of NSOs in these three countries has enabled them to operate independently, thus increasing the trust of statistics produced. The fifth factor is the alignment of the National Strategies for Development of Statistics with country development plans, and the last factor is the coordination of statistical stakeholders within NSS (PARIS21, 2018).

3.5.4 Data Collection, Processing, and Analysis

Data collection involves collecting information from different sources with the aim of meeting the needs of different users. The collected data are thereafter processed and analyzed to produce required statistical indicators. These are three phases: collect phase, process phase, and analysis phase under the GSBPM, as shown in Fig. 3.3. Under the collect phase, all the necessary information is collected to meet the identified user needs. This phase also comprises of all preparatory activities aimed at ensuring that people, processes, and technology are ready to collect the required statistical information. These preparatory activities include sample creation and selection; training of data collection personnel; procurement and provision of required data collection instruments; publicity and advocacy to inform respondents about the statistical activities; close supervision and coordination of data collection personnel to minimize nonresponses and errors; and data transmission, storage, and retrieval for analysis (United Nations, 2021; UNECE, 2019).

African NSOs face a number of challenges while collecting data for the production of official statistics. Some of these challenges are similar to the challenges affecting other data production processes as mentioned earlier, such as inadequate funding, overreliance on donor funding, data gaps, low usage of alternative data sources, inadequate coordination of NSS, understaffing and inadequate human capital skills and capacity especially in modern ICT environment, insufficient update of administrative data collection instruments, and unautomated administrative data systems.

Other challenges as noted in Kenya National Bureau of Statistics 2018–2022 Strategic Plan, Statistics South Africa Strategic Plan 2020/2021–2024/2025, Statistics Botswana Strategic Plan 2015–2020, and Namibia Statistics Agency Strategic Plan 2017/18–2021/22 include delay in disbursement of allocated funds from exchequer, delay in procurement of required products and services, delay/failure to update sampling frames, limited sample sizes that do not allow disaggregation of data to the lowest levels possible, low surveys/censuses response rates, difficulties in accessing some clusters in high-end residential areas and conflict prone areas, and lack of an elaborate long-term plan indicating timelines for conducting different surveys and censuses.

The process phase involves processing of the collected data and preparing it for analysis, while under the analyze phase, statistical outputs are generated and validated before they are disseminated. The process and analyze phases are usually done concurrently. Processing and analyzing data in Africa face similar challenges as those outlined in the data collection phase.

In addition, some African countries are experiencing challenges in the computation of gross domestic product (GDP) growth rates and poverty estimates whereby the statistical foundations used to compute these statistical indicators are very weak. Computation of GDP in some African countries is done using old methods, and there is a lack of frequent poverty estimates in most African countries and those available are not comparable across time and space (Devarajan, 2013). For example, a survey conducted by United Nations Economic Commission for Africa in September 2020 showed that 30 African countries (55.6 percent) were using 2008 System of National Accounts (2008 SNA) to compute GDP, 22 countries (40.7 percent) were using 1993 SNA, 1 country (Sudan) was using 1968 SNA, while there was no information for Libya (UNECA, 2020a, b, c). In addition to the above challenges, there is fragmentation of statistical activities in Africa, especially in surveys, whereby two or more statistical activities on one topic are conducted by different organizations. These statistical activities are usually conducted using different methods, thus making it difficult to compare the outputs across time and space (Devarajan, 2013).

The adoption of Global Statistical Geospatial Framework, which facilitates the integration of geospatial and statistical information, will enable statistical agencies to analyze data up to the lowest level possible (United Nations Statistical Commission, 2019). The framework enables integration of different data sets from geospatial and statistical communities, which allows generation of standardized and harmonized geospatially enabled statistical information.

3.5.5 Dissemination and Use of Official Statistics

This is the dissemination phase under GSBPM where statistical outputs are released to users using different channels. Once the production of official statistics is complete, these statistics and related statistical information such as metadata and methodologies used need to be made available, accessible, and understandable by all types of users. Dissemination and use of statistics that provide better understanding and better decision-making in relation to the economy and society in general should be considered as the core objective of national statistical system. Statistical agencies should provide support to all data users by responding to requests for additional data and information and also responding to queries in regard to the released statistical outputs (United Nations, 2021; UNECE, 2019).

Accessibility and use of official statistics are challenges for several African countries due to the following difficulties that also affect other data production processes mentioned earlier: weak data technology, weak infrastructure, political issues, inadequate funding, and inadequate capacity. These challenges limit the capability of African NSOs to disseminate official statistics through various channels, which include national statistical agencies' websites, data portals, social media, machine to machine, hardcopy, mobile apps, Geographic Information System (GIS) portals, statistical yearbook, and dynamic visualizations. Statistical outputs from surveys and censuses conducted in Africa are usually released late. This affects the usability of the data. However, the use of technology in data collection for surveys and censuses in Africa is enabling the timely release of statistical products that helps in filling the existing data gaps (UNECA et al., 2017).

Policymakers in Africa as well as the general public are not adequately informed on the important role that statistical information can play in improving the social and economic development of an economy. This lack of awareness on the importance of statistics in the society is an impediment to the development and use of statistics in Africa, and it negatively affects the quality and availability of statistical information.

Making official statistics available and accessible to all users increases the possibility of these statistics being used in decision-making by public officials and civil servants, enhances transparency and accountability in government operations, and improves delivery of services to the public, thus leading to better socio-economic outcomes in the society. In addition, availability and accessibility of official statistics foster innovation and economic development, where businesses use these statistics to gain a better understanding of different market segments, and based on these statistics, they develop new goods and services (PARIS21 and The Mo Ibrahim Foundation, 2021; UNECA et al., 2016).

Data openness by NSOs provides information on the available data, reduces duplication across NSS, reduces the data production cost, and enables users to identify data gaps that inform data production activities. Data openness in Africa is still very low; for example, in 2018, data openness score was 12 points below the world average. In addition, African NSOs experience challenges in providing data in user-friendly formats. A review conducted by Open Data Watch in 2018 shows that out of the data published on 49 African NSOs websites, 69.5 percent of the data were in nonproprietary format, 39.7 percent of the data's metadata were available, 29.4 percent had download options, 18.5 percent of the data were machine-readable, and 12.9 percent of data had terms of use (PARIS21 and The Mo Ibrahim Foundation, 2021; UNECA et al., 2016).

Engagement of data users by NSOs in Africa should be one of their core activities, and this can be achieved by developing a user engagement strategy. This is part of the evaluate phase, the eighth phase under GSBPM. According to UNECA (2020a, b, c), a user engagement strategy outlines the methods to be used to encourage interactions among data producers and users, establish mechanisms for obtaining feedback and experiences on the use of statistical products and services, and provide a framework that guarantees that data users' feedback is considered while making data production decisions.

3.5.6 Data Governance

Data governance in the production of official statistics refers to those support activities undertaken by national statistical offices to guide and oversee the implementation of statistical activities. These support activities entail the legal frameworks governing the production of official statistics, institutional frameworks of NSOs and other institutions within the NSS, and financing frameworks and overall management of NSOs (United Nations, 2021; National Institute of Statistics of Rwanda, 2019; Statistics Botswana, 2015). The overall management of NSOs comprises the following activities: general management, human resources management, financial management, supply chain management, management of organizations' assets including ICT equipment, and quality management (Statistics Botswana, 2015; Porter, 1985).

Data governance functions are categorized into four clusters; the first cluster is strategic planning which involves the development of strategies and establishment of institutional arrangements. The second cluster is making and implementation of rules, and it entails making legislation and regulations, setting standards, and offering clarification and guidance. The third cluster is compliance that comprises enforcing, auditing, arbitrating, and remedying. Enforcing is done on a daily basis in order to ensure compliance with the established legal framework, standards, and norms. The fourth cluster is learning and evidence, and it comprises monitoring and evaluation and risk management. This function enables an organization to assess its performance and that of its staff and also evaluate achievement of the project(s) objectives (World Bank, 2021).

Laws and regulations governing statistical agencies and national statistical system significantly determine the quality of official statistics produced, their availability, accessibility, and use. Laws and regulations governing NSS in Africa should enable NSOs to produce accurate, relevant, consistent, comparable, timely, reliable, complete, impartial, and accessible data as per Fundamental Principles of Official Statistics as well as African Charter on Statistics. African NSOs should be autonomous to enable them produce official statistics without any interference, especially political interference. Some of the challenges facing African NSS, especially NSOs, is lack of political and institutional independence, which has weakened the managerial and technical ability of some NSOs to work effectively. For example, out of the 54 African countries, only 12 National Statistical Offices in Africa are considered to be autonomous (UNECA et al., 2017).

Another example is that during the 1990 round of population and housing censuses, Kenyan census results were subjected to political ridicule, while those from Nigeria were accepted after eight years. In addition, in some African countries, data on population, education, and agriculture are commonly politicized since the data are mainly used for delineation of political boundaries and resource allocations. Therefore, legal framework governing operations of NSOs in Africa should allow these offices to operate independently and should be anchored on the Fundamental Principles of Official Statistics and African Charter on Statistics. This will improve the quality, credibility, and trust in official statistics since these offices are able to ensure that there is professional independence, impartiality, and application of international standards and scientific methods in production of official statistics. The autonomy of African NSOs is not enough and should be coupled with provision of adequate human and financial resources (UNECA et al., 2016).

Evidence shows that legal autonomy enhances the trustworthiness of NSOs, thus enabling them to attract and sustain funding from both domestic sources and development partners. Rwanda, Ethiopia, and Philippines are examples of countries that have been able to attract and sustain funding for statistical activities since the existing legal framework has enabled the NSOs to operate independently (PARIS21, 2018).

Institutional frameworks entail institutions within NSS, which includes government ministries, departments, and agencies. There are a number of organizations within NSS, and these organizations operate under different mandates, and there are variations in their financing and capability (National Institute of Statistics of Rwanda, 2019). Coordination of NSS is therefore essential in ensuring that these institutions supply the required data for the production of official statistics. In Africa, the capacity of NSOs to coordinate NSS is insufficient, which negatively affects the supply of data for the production of official statistics.

Succession planning and management across NSOs in Africa is weak, and this negatively affects the production of official statistics. Experienced employees including managers exit the service consequently, affecting the management and operations of NSOs. African NSOs should therefore develop and implement a succession strategy. Another challenge is that some African NSOs do not have adequate office space, and the work environment is unfavorable (Kenya National Bureau of Statistics, 2018).

African NSOs should develop strategic plans and national statistical development strategies to guide the implementation of statistical activities and other support activities. As indicated earlier, African countries that have developed strategic plans usually face obstacles that hinder full implementation of the plans, which negatively affects the timelines of these activities. Quality management is another key element that NSOs should embrace, since it ensures production of quality data that meet the needs of different users. Development and implementation of Data Quality Assurance Framework (DQAF) is one of the ways of managing the quality of data produced by NSOs. Some African NSOs do not have these data quality assurance frameworks, which makes it difficult to assess the quality of statistics produced.

3.6 Case Studies

3.6.1 Case Study One: Australian Bureau of Statistics

Australian Bureau of Statistics (ABS) is established as an independent statutory authority under Australian laws. It is mandated to collect statistical information on a wide range of economic, demographic, social, and environmental topics and contains solid provisions for maintaining confidentiality of collected information (Australian Bureau of Statistics, 2019). ABS has been producing and providing high-quality statistics to all users including governments, general public, and researchers, and these statistics influence key decisions. ABS has been receiving adequate funding from the government, which has enabled the institution to carry out its functions effectively and efficiently (Telford et al., 2017). ABS generates independent statistics and upholds the highest standards of confidentiality, and the statistical information produced by the NSO is trusted by majority of the users. ABS is impartial, exercises professionalism in its work, and is transparent and accountable to Australian community (Telford et al., 2017).

ABS uses alternative sources of data in the production of official statistics and only conducts surveys, when necessary, in order to improve the quality of its statistical products (Australian Bureau of Statistics, 2019). In its website, ABS has a monthly release calendar that indicates the expected release dates of various statistical products in the next six months. Furthermore, the released statistical products are uploaded in ABS website, and they are easily accessible. The NSO has established an online user engagement platform, the consultation hub that is domiciled in its website. The consultation hub enables data users to provide feedback on the use of ABS statistical products and services (Australian Bureau of Statistics, 2021).

The NSO has a workforce strategy which aims at increasing its capacity and ability; identifying, monitoring, and forecasting where knowledge and skills are needed; and supporting innovative methods of working in order to achieve high performance and efficiency. ABS is improving its capability to leverage non-traditional data sources, especially big data, and also utilize artificial intelligence methodologies in production of official statistics. ABS uses both international and domestic standards, frameworks, and classifications, which ensure production of high-quality, consistent, and reliable statistics (Australian Bureau of Statistics, 2019). A study conducted by Howard in 2018 indicate that ABS experiences a number of challenges, among them being reduction of its budgetary allocations, political interference, and criticisms of its technical ability (Howard, 2019).

3.6.2 Case Study Two: National Institute of Statistics of Rwanda

The National Institute of Statistics of Rwanda (NISR) is an established independent institution under Rwandan laws, and its mandate is being the primary producer of data in Rwanda. NISR has established strong patterns of producing and disseminating data, which has improved the alignment between data demand for making policies and the supply of data. The number and frequency of official statistical products produced by NISR has increased overtime, and the NSO ensures it meets data users' needs. NISR ensures that its products comply with the policy needs and international standards and has implemented initiatives aimed at facilitating data accessibility, strengthening partnerships, and enhancing statistical capacity (National Institute of Statistics of Rwanda, 2019).

National Institute of Statistics of Rwanda usually conducts biennial user satisfaction surveys where feedback is collected from all users of official statistics. The last survey conducted in 2016/17 indicates that statistics produced by NISR meets users' priority requirements, an indication of an increase in awareness and utilization of these statistics. The survey enables NISR to map data requirements of different users who include decision-makers, analysts, technicians, investors, journalists, and citizens. Rwanda has synchronized its development plans and policies with the production of official statistics. Statistical information required for monitoring and evaluating these plans is usually mapped, and gaps are identified (National Institute of Statistics of Rwanda, 2019).

Rwanda is commonly referred to as a model for financial statistics. Official statistical activities in Rwanda are sufficiently funded domestically, especially during the first National Strategy for the Development of Statistics. In addition, funds from development partners are managed through a multi-donor basket fund. Donors' support for statistics, both financial and technical, has been aligned to national priorities (National Institute of Statistics of Rwanda, 2019; PARIS21, 2018).

Rwanda has developed a "Data Revolution Policy" whose main focus is to build capabilities of different stakeholders with the aim of equipping them with knowledge and skills needed to analyze big data. Under this policy, NISR has been carrying out a range of capacity-building initiatives, particularly on rebasing of national accounts, engagement with university students and teaching staff on availability and use of NISR statistics, engagement with secondary school students on reading data, and training media personnel on statistics, among others (National Institute of Statistics of Rwanda, 2021).

Rwanda is set to establish NISR Training Centre for building the capacity of workers within NSS, and it has established the African Centre of Excellence in Data Science domiciled in the University of Rwanda. Rwanda has established partnerships in big data, and it is exploring how to use this source of data for monitoring and evaluation of development programs (National Institute of Statistics of Rwanda, 2021).

In 2019, official statistics produced by NISR were criticized through an article in Financial Times whereby, since 2000, gross domestic product growth rate of over 7.0 percent, reduction of infant mortality rate by half, and improvement in access to education and health were doubted. Poverty statistics were also criticized, whereby the newspaper argued that between 2010 and 2014, poverty rate increased contrary to what was officially reported that poverty reduced. Political influence on official statistics and the urge to attract donors funding were singled out as some of the reasons of misreporting official data (Wilson and Blood, 2019).

In another article by Royal African Society, a comparative analysis of Rwanda's Integrated Household Living Survey conducted in 2005/06 and 2016/17 points out inconsistency of youth population aged 10–24 years. Analysis of population growth by age between these two survey periods indicates that 580,000 youth population aged 10–24 were missing and there was no explanation of the whereabouts of this population (Ansoms et al., 2021).

3.6.3 Case Study Three: Statistics South Africa

Statistics South Africa (Stats SA) is a national government department established as an independent organization under South African laws. Statistician-General of Stats SA is appointed by the president and is required by the Statistics Act to exercise professional independence and impartiality while discharging his/her duties and responsibilities. The Act establishes a Statistics Council that consists of members from diverse organizations or interest groups (Statistics South Africa, 2020, 2021).

Stats SA produces and disseminates over 250 statistical products annually, and it has demonstrated its capacity to deliver key projects. The NSO has been able to achieve an average response rate of over 85 percent in household and establishment-based surveys, which is in accordance with international best practice. Over the years, Stats SA has adopted international standards and practices in the production and dissemination of official statistics (Statistics South Africa, 2020).

Stats SA has employees' vacancy rate of 19.2 percent, and 53.0 percent of all workers are women. The NSO continuously provides training to its workforce in order to enhance their skills and capability. During the implementation of 2015–2020 strategic plan, Stats SA was able to consistently achieve 80.0 percent of its target, as outlined in the work plan. User satisfaction survey conducted by Stats SA showed that the department produces credible and trustworthy statistics, and several users access these statistics from its website (Statistics South Africa, 2020). Stats SA has established several initiatives such as ISIbalo Capacity Building Programme aimed at training young statisticians. The NSO uses various channels to disseminate official statistics such as Roambi, an App that was used to disseminate 2011 census data (Statistics South Africa, 2021).

Although Stats SA has been praised internationally for the production of high-quality statistics, some of its statistics have been criticized both locally and globally. For example, it is doubted that the South Africa population census data exceed that of a number of advanced countries (Van Belle, 2017). Findings of a study on the use of available data to inform the COVID-19 outbreak in South Africa pointed out that the methods used by the government to share information on the pandemic were ineffective. This was due to the use of different data-sharing platforms, requiring a user to navigate so as to get accurate data; also, data provided were not in computer readable format, requiring further processing.

This negatively affects the accessibility, simplicity, and readability of the shared information, which has been identified as a major concern of data shared by the South African government. The study notes that there is need for government departments to continuously engage data users and collaborate with other organizations in order to produce and disseminate accurate data (Marivate and Combrink, 2020). Budgetary allocations to Stats SA have been declining, and this has been negatively affecting the NSO's operations (Price, 2021).

3.6.4 Discussion on the Case Studies

The three case studies indicate that the three NSOs are established as independent organizations; follow international standards and methods in production statistics; produce quality statistics that meet needs of different users including policy makers; disseminate official statistics through various platforms, thus making them available and accessible; engage data users through various platforms; are using or exploring use of non-traditional sources of data; have adopted use of modern technologies in data production; receive relatively adequate funding from the government; have relatively adequate human resources; and coordinate NSS. However, some official statistics produced by these three NSOs have been criticized as inaccurate, inconsistent, and impartial. Despite these criticisms, it is evident that these three NSOs have adopted key value chain approach activities in data production, use, and governance, and therefore, other African countries can use them as a benchmark in the process of adopting value chain approach in their statistical operations.

3.7 Recommendations

The following recommendation are suggested to enable improvement of efficiency of data governance at the National Statistical Offices in Africa.

3.7.1 Strengthening the Systems Used in Collection and Processing of Administrative Data

In order to fill the existing data gaps and satisfy their needs, data users such as private sector, civil-society organizations, academia, and other non-state actors resort to independent production of statistical information for use in their operations. Strengthening administrative data systems and using data from other producers will help in filling data gaps, enhance production of quality statistics, and reduce duplication in production of official statistics. This can be achieved by enhancing the capacity of NSOs to coordinate the NSS and formation of partnerships with other data producers outside NSS to supplement official statistics. Efficient and effective administrative data systems (e.g., automated) will reduce data gaps, which will then reduce sample sizes for surveys and the number of variables to be collected in surveys/censuses, thus lowering the cost of data production and improving data quality. This will improve the quality of basic statistical data produced across NSS, which will lead to the production of quality official statistics.

3.7.2 Use of Non-traditional Sources of Data in Production of Official Statistics

African NSOs need to address issues limiting access, invest in technology required to process and analyze the data, and hire and retain staff with technical capacity and skills to process, analyze, and interpret the data from non-traditional sources. Specifically, African countries need to develop and implement legal framework that protects personal data and guarantees data privacy in order for them to gain access and benefit from these non-traditional data sources.

Enacted data protection laws and regulations should ensure that data obtained from all sources and in the custody of NSOs and other organizations across NSS are treated with utmost confidentiality. In addition, organizations across NSS should put in place strong data security mechanisms that protect the data in their custody against the ever-evolving digital attacks (cyberattacks).

African NSOs need to form partnerships with owners of big data, such as businesses, so as to create an enabling environment for sharing the data. In addition, NSOs in Africa need to collaborate with ICT agencies, data protection agencies, monitoring and evaluation agencies, data scientists and data analysts, among others, to facilitate continuous interactions in relation to use of new data sources. African countries need to adequately fund their respective NSOs to enable them acquire modern technologies required to process and analyze data from non-traditional sources.

3.7.3 Identification of Needs for Different Users and Establishing a National Users' Engagement Platform

African NSOs should strengthen the process of identifying data needs for different users, and this process should be part of their routine activities. These offices should create an engagement platform where data users, mainly policymakers across government, and lawmakers are encouraged to make use of available official statistics while making decisions. The statistical agencies in Africa should develop online user engagement platforms and regularly conduct user satisfaction surveys through which all types of data users will continuously provide feedback on the usage of official statistics and their level of satisfaction. This requires NSOs to develop a national user engagement strategy, which will guide the process of engaging different data users and producers, obtaining their feedback and experiences on use of official statistical products and services and incorporating this feedback in production of official statistics.

3.7.4 Coordinating the National Statistical System and Collaboration Across African Statistical System

As outlined in Fundamental Principles of Official Statistics and African Charter on Statistics, NSOs across Africa should coordinate the national statistical system in order to achieve an efficient and consistent statistical system and ensure production of quality and comparable statistical information. The legal framework establishing and governing operations of African NSOs should give these offices legal authority to coordinate NSS through monitoring and evaluating the quality of data collected by different government agencies in their respective countries. Under the established legal framework, African NSOs will be mandated to ensure that there is collaboration, harmonization, and coordination across NSS.

Organizations across African Statistical Systems need to collaborate among themselves in order to improve the production and use of statistical information, and there is a need to establish peer-review mechanism. At regional level, collaboration can be achieved through Regional Economic Blocs/Communities with the aim of producing comparable statistics. African NSOs through the African Union need to establish a peer-review mechanism through which periodic peer reviews of statistical processes of an NSO will be conducted by other statistical agencies.

3.7.5 Use of International Standards and Methods in Production of Official Statistics

African countries should continuously update sampling frames or create new sampling frames in order to ensure that samples drawn are representative of the population. African NSOs should domesticate and use international accepted standards and methods in the production of different statistics; this will improve quality and comparability of official statistics at regional, continental, and international levels.

African NSOs need to leverage their adoption of modern technologies in the production of statistical information to develop and adopt an integrated data production system, where processes for different statistical activities are standardized and linked together through development and application of similar standards, methods, and modern technologies. This will achieve efficiency in production of official statistics, lower costs, and reduce the time taken to produce statistical outputs.

3.7.6 Funding of Activities for Production of Official Statistics in Africa

Funding the production of official statistics in Africa needs to be prioritized in national budgets with the aim of ensuring allocation of adequate funds for the production of statistics across NSS. Donors/development partners funding for statistical activities in Africa should be coordinated and managed centrally to ensure efficiency, through a committee or multi-donor basket fund. This will ensure that statistical activities funded by donors/development partners are aligned to national priorities.

African countries' statistical requirements need to be anchored in National Strategies for Development Statistics and synchronized with national development plans and policies, and funding including that from donors needs to be directed toward these strategies. This will increase the demand for official statistics to monitor and evaluate both national and international priorities, which is key in attracting and sustaining funding of statistical activities. Legal autonomy of NSOs, political support for statistics, and coordination of NSS by NSOs will also improve funding for national statistical systems in Africa.

African NSOs should develop resource mobilization strategy and create an office purely responsible for the mobilization of resources for carrying out statistical activities. The mandates of African NSOs need to be enhanced to enable them provide technical and consultancy services at a fee in order to increase internally generated revenues.

3.7.7 Addressing Human Capital Skills Gaps Across African NSOs

African NSOs need to implement the following strategies in order to resolve challenges of understaffing and inadequate technical skills. First, NSOs in Africa need to hire additional employees to address shortages. New employees need to include data analysts, data scientists, data engineers, and data miners who have the capability to extract and process non-traditional sources of data and also improve overall production of official statistics. Second, African statistical agencies need to come up with appropriate career progression framework which will enable them to hire and retain young professionals, thereby reducing staff turnover. Under this strategy, NSOs should develop and implement performance management system where performance of employees is evaluated and rewards are given based on performance.

Third, development of succession strategy will enable transfer of managerial and technical skills from experienced staff to newly employed employees and ensure smooth running of operations. Fourth, African NSOs should create knowledge management systems where employees who have undergone technical training in various fields share their knowledge and skills with other employees.

Fifth, African NSOs should come up with training and development strategies where technical skills and competencies of existing workforce are identified and articulated, required skills are identified, and strategies of addressing these skills gaps are developed. Sixth, African NSOs should form and build partnerships with academic institutions and other data communities including private sector in order to share technical knowledge and skills in new data production methodologies. These statistical agencies should also aim at establishing technical statistical training institutes where trainees will be offered technical training in the production of official statistics.

3.7.8 Investment in Modern Technologies for Use in Production and Dissemination of Official Statistics

Information and communication technology (ICT) plays a significant role in all aspects of statistical production, starting from data collection up to dissemination. ICT landscape is dynamic and is rapidly changing, and therefore, NSOs need to keep up with these rapid changes by investing in modern technologies through acquisition of modern ICT hardware and software. Acquisition and use of modern technologies will enable NSOs to develop new products and processes, thus enabling them to overcome challenges encountered during the production and dissemination of official statistics. This will, in turn, reduce costs of producing data and improve quality of statistics produced. The investment in modern technologies needs to be coupled with the hiring of ICT experts (data analysts, data scientists, data engineers, data miners) and young professionals so as to reduce shortage of employees and fill the existing skills gap, especially in modern technologies.

3.7.9 Enhancing the Use of Official Statistics

African NSOs first need to release statistical information and statistical products in their custody for use by all users. These statistical products need to be released in user-friendly formats, which are easy to use, and that can be integrated with other data sets. This will enable users to access available statistics for use, which will, in turn, assist in identification of data gaps. This will also enhance data openness, thus improving the trust of official statistics in Africa.

Second, African NSOs need to develop an advocacy strategy for use of statistics, whereby all users will be informed on the availability of the various statistical products, how they can be accessed, and how they can be utilized. These offices should also provide support to data users, and details on how this support can be obtained need to be clearly communicated. In addition, African NSOs should develop a dissemination strategy outlining the guidelines for disseminating and communicating official statistics and methods for dealing with errors and making revisions.

Third, African NSOs should create an engagement platform where data users, mainly policymakers across government, and lawmakers are urged to make use of available official statistics while making decisions. Engagement of data users by NSOs in Africa should be one of their core activities, and this can be achieved by developing a user engagement strategy.

Fourth, African countries need to regularly train government employees, especially middle-level managers, on how to analyze, interpret, and communicate data for effective and efficient day-to-day decision-making. Fifth, official statistics should be incorporated in countries' development plans in order to enhance their uptake by government officials and other stakeholders. Development plans in Africa should be synchronized with the production of official statistics, whereby all statistical indicators required to monitor and evaluate the development goals are clearly mapped and produced by national statistical systems.

3.7.10 Improving the Trust of Official Statistics in Africa

There is a need to enhance the trust of official statistics in Africa in order to increase uptake of these statistics for evidence-based decision-making. First, African NSOs need to ensure adherence to Fundamental Principles of Official Statistics in the production of official statistics, which will guarantee the production of quality statistical information through exercising professional independence and impartiality.

Second, African NSOs need to come with release dates and release calendar for their regular and ad hoc statistical products. This will enable data users to know the exact date when they expect different statistical products to be released and made available for use by all users. NSOs should promptly communicate any changes made to the release schedule and reasons for the changes.

Third, dissemination of official statistics needs to be done by the chief executive officers of NSOs, thus enabling these offices to exercise their independence. Fourth, there is a need to build the capacity of African NSOs leadership on the best practices of producing and disseminating official statistics. This will improve the capacity of NSOs management to engage policymakers and advocate for use of official statistics in formulation and evaluation of policies. NSOs should develop strategies to promote usage of official statistics at policy level, which will, in the long run, facilitate allocation of adequate resources for statistical activities.

3.7.11 Improving Data Governance Across NSOs in Africa

The legal framework governing the production of official statistics in Africa should ensure that produced statistics are accurate, relevant, timely, comparable, consistent, and impartial as per Fundamental Principles of Official Statistics and African Charter on Statistics. African NSOs should be autonomous to enable them produce official statistics without any interference, and this will ensure that they exercise professional independence in data production, thus improving the credibility and trust of official statistics. Legislation on production of official statistics should therefore be anchored on the Fundamental Principles of Official Statistics and African Charter on Statistics, which guarantees production of quality official statistics.

On succession planning, African NSOs need to develop and implement a succession strategy, which will enable transfer of managerial and technical skills from experienced staff to newly employed employees, and ensure smooth running of operations even after knowledgeable staff exit the service. On corporate strategy, African NSOs should develop and implement strategic plans and national strategies for development statistics that will guide the implementation of different statistical activities and related support activities. African NSOs should develop and implement procurement plans for different statistical activities so as to ensure timely procurement of required materials and services. African NSOs should also strive to get adequate office space in order to provide favorable work environment for their workers. National statistical agencies in Africa that have not developed DQAF should develop and implement these frameworks to ensure generation of quality statistics.

3.8 Conclusion

Official statistics have been identified as part of the challenges linked to slow development in Africa as they are largely viewed as of low quality. There is a need, therefore, to strengthen the state of official statistics in Africa by relooking at the way these statistics are produced in Africa with the aim of informing policies and monitoring and evaluating development goals at national, regional, continental, and international levels.

The study focused on the need for NSOs to encompass value chain approach in data production, increasing the uptake of statistics and improving data governance in order to improve the quality of official statistics for sound policymaking in Africa. Review of existing literature indicated that African NSOs and respective NSSs experience a number of challenges, including weak basic statistical data systems, low use of non-traditional sources of data, lack of a comprehensive data users engagement strategy, weak coordination of the NSS, inadequate data dissemination systems, lack of a strategy to enhance uptake of official statistics, inadequate funding, inadequate human resources, and gaps in the legal framework leading to lack of autonomy, among others. Recommendations have been made with the aim of improving production and use of official statistics as well as governance of official statistics processes.

Studies are normally conducted in different periods; therefore, existing literature may not reflect the current position of the study subject. There is a need, therefore, to conduct further research preferably through a survey covering all African countries where African NSOs and other organizations across NSS would be asked questions on production of official statistics, usage of official statistics, and governance of official statistics processes. In addition, data users such as research institutions would be asked questions on their experiences in relation to the use of official statistics.

References

African Union Commission. (2009). *African charter on statistics*. African Union Commission.

African Union Commission, et al. (2017). *Strategy for the harmonization of statistics in Africa, 2017–2026, SHaSA 2*. African Union Commission.

Ansoms, A., Cottyn, I., Niyonkuru, R. C., & Vrelust, T. (2021). *The disappearance of half a million young people from Rwanda's stats*. Royal African Society. https://africanarguments.org/2021/01/the-disappearance-of-half-a-million-young-people-from-rwandas-stats/ (accessed 02 November, 2021)

Australian Bureau of Statistics. (2019). *Corporate plan 2019–20*. Australian Bureau of Statistics.

Australian Bureau of Statistics. (2021). *ABS Consultation Hub*. https://www.abs.gov.au/about/consultation (accessed 7 September 2021).

Bruno, M., et al. (2020). Census metadata driven data collection monitoring: The Ethiopian experience. *Statistical Journal of the IAOS, 36*, 67–76.

Devarajan, S. (2013). Africa's statistical tragedy. *Review of Income and Wealth*, Series 59, Special Issue. https://doi.org/10.1111/roiw.12013.

European Commission, Statistics Sweden and Eurostat. (1999). *Quality work and quality assurance within statistics*. Statistics Sweden.

Ghana Statistical Service (GSS). (2020). *Corporate plan, 2020–2024*. Ghana Statistical Service.

Government of Malawi, National Statistical Office. (2019). *2018 Malawi population and housing census: Main report*. National Statistical Office.

Howard, C. (2019). The politics of numbers: Explaining recent challenges at the Australian Bureau of Statistics. *Australian Journal of Political Science, 54*(1), 65–81. https://doi.org/10.1080/10361146.2018.1531110

International Telecommunication Union-ITU. (2016). *Big data for measuring the information society: Country Report-Kenya*. International Telecommunication Union.

Kenya National Bureau of Statistics. (2019). *2019 Kenya population and housing census: Volume I on Population by county and sub-county*. Kenya National Bureau of Statistics.

Kenya National Bureau of Statistics (KNBS). (2018). *2018–2022 Strategic plan*. KNBS.

Kenya National Bureau of Statistics (KNBS). (2019). *Kenya strategy for the development of statistics, 2019/20–2022/23*. KNBS.

Marivate, V., & Combrink, H. M. (2020). Use of available data to inform the COVID-19 Outbreak in South Africa: A case study. *Data Science Journal, 19*(19), 1–7. https://doi.org/10.5334/dsj-2020-019

Namibia Statistics Agency (NSA). (2017). *Strategic plan (2017/18–2021/22)*. Namibia Statistics Agency.

National Institute of Statistics of Rwanda (NISR). (2019). *The Third National Strategy for the Development of Statistics, NSDS3*. NISR.

National Institute of Statistics of Rwanda. (2021). *Data revolution*. National Institute of Statistics of Rwanda. https://www.statistics.gov.rw/content/data-revolution (accessed 31 August, 2021)

PARIS21. (2018). *Good practices for sustained financing of national statistics.* PARIS21 Discussion Paper, No. 12, Paris.

Porter, M. E. (1985). *Competitive advantage: Creating and sustaining superior performance*. The Free Press.

Price, A. (2021). *Stats SA 'throttled' by budget cuts, requires extra funding to combat brain drain*. The Mail & Guardian. https://mg.co.za/business/2021-03-19-stats-sa-throttled-by-budget-cuts-requires-extra-funding-to-combat-brain-drain/ (accessed 02 November, 2021)

Republic of Kenya. (2021). *Kenya shares its census technology with other African countries*. Republic of Kenya. https://www.president.go.ke/2021/04/01/kenya-shares-its-census-technology-with-other-african-countries/ (accessed 31 August, 2021)

Statistics Botswana. (2015). *Strategic plan, 2015–2020*. Statistics Botswana.

Statistics South Africa. (2020). *Strategic plan 2020/2021–2024/2025*. Statistics South Africa.

Statistics South Africa. (2021). *ISIbalo capacity building programme*. http://www.statssa.gov.za/?page_id=6454 (accessed 12 October 2021).

Telford, J., Araghi, R., & Samson, P. (2017). *Modernization processes in National Statistical Offices – Transforming the Australian Bureau of Statistics*. Australian Bureau of Statistics.

The Partnership in Statistics for Development in the 21st Century (PARIS21) and The Mo Ibrahim Foundation (MIF). (2021). *Bridging the Data-Policy Gap in Africa: Recommendations to national statistical officers and governments to enhance the production and use of data for evidence-based policymaking*. PARIS21 Secretariat and MIF, Paris.

United Nations. (2021). *the handbook on management and organization of national statistical systems: 4th edition of the handbook of statistical organization*. United Nations.

United Nations Conference on Trade and Development (UNCTAD). (2020). *Data protection and private legislation worldwide*. UNCTAD. https://unctad.

org/page/data-protection-and-privacy-legislation-worldwide (accessed 02 November, 2021)

United Nations Economic and Social Council. (2013). *Fundamental principles of official statistics*. United Nations Economic and Social Council.

United Nations Economic Commission for Africa (UNECA). (2020a). *2020 round of population and housing censuses in Africa*. UNECA.

United Nations Economic Commission for Africa (UNECA). (2020b). *Guidelines for developing and integrated user engagement strategy for national statistical systems*. UNECA.

United Nations Economic Commission for Africa (UNECA). (2020c). *Progress report on the implementation of the 2008 system of national accounts*. UNECA.

United Nations Economic Commission for Africa (UNECA), et al. (2017). *2017 Africa sustainable development report: Tracking progress on agenda 2063 and the sustainable development goals*. UECA.

United Nations Economic Commission for Africa (UNECA), et al. (2016). *The Africa data revolution report 2016: Highlighting developments in African data ecosystems*. ECA Printing and Publishing Unit.

United Nations Economic Commission for Europe (UNECE). (2018). *Recommendations for promoting, measuring and communicating the value of official statistics*. United Nations.

United Nations Economic Commission for Europe (UNECE). (2019). *Generic Statistical Business Process Model (GSBPM)*. United Nations.

United Nations Statistical Commission. (2020). *Report of the global working group on big data for official statistics*. United Nations Statistical Commission.

United Nations Statistical Commission. (2019). *The global statistical geospatial framework*. United Nations Statistical Commission.

United Nations Statistics Division. (2020). *Report on the results of the UNSD survey on 2020 round population and housing censuses*. United Nations Statistics Division.

United Nations Statistics Division and World Bank. (2020). *Monitoring the state of statistical operations under the COVID-19 Pandemic: Highlights from a global COVID19 survey of National Statistical Offices*. United Nations Statistics Division.

United States Census Bureau. (2020). *2020 Census questionnaires and instructions*. United States Census Bureau. https://www.census.gov/programs-surveys/decennial-census/technical-documentation/questionnaires.2020_Census.html (accessed 31 August, 2021)

Van Belle, J. P. (2017). *The South Africa country data report*. Centre for Information Technology and National Development in Africa (CITANDA). University of Cape Town, Cape Town.

Wilson, T., & Blood, D. (2019). *Rwanda: Where even poverty data must toe Kagame's line*. Financial Times. https://www.ft.com/content/683047ac-b857-11e9-96bd-8e884d3ea203 (accessed 02 November, 2021)

World Bank. (2021). *World development report (2021): Data for better lives.* World Bank. https://doi.org/10.1596/978-1-4648-1600-0

CHAPTER 4

Data Protection Legal Regime and Data Governance in Africa: An Overview

Olumide Babalola

4.1 Introduction

Data protection refers to strategies and processes that are applied to provide security to privacy, availability and integrity of data use and production.[1] On the other hand, data governance is defined by the Data

Barrister and Solicitor of the Supreme Court of Nigeria; Managing Partner, Olumide Babalola LP, Nigeria; Member, International Association of Privacy Professionals (IAPP); Member, International Network of Privacy Law Professionals (INPLP).

[1] Cloudian. Data Protection and Privacy: 12 ways to Protect User Data. https://cloudian.com/guides/data-protection/data-protection-and-privacy-12-ways-to-protect-user-data/ (accessed on April 9, 2022).

O. Babalola (✉)
University of Portsmouth, Portsmouth, UK
e-mail: olumide@oblp.org

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8_4

Governance Institute[2] as the 'system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods.' Additionally, within the context of data protection and data governance, digital rights are ingrained in production and utilization of the data through digital gadgets. With increased use of digital space to generate data, protection on the rights of the user is vital when formulating the legal framework.

Data protection reportedly surfaced in Europe in 1970 when the German Federal State of Hesse (Mayer-Schonberger, 1997) enacted its data protection law, which was followed by the Swedish national Data Act in 1973 (Oman, 2010). In Africa, the Republic of Cape Verde led the way in 2001 when it enacted the first data protection law in Africa. The Cape Verdean Data Protection Act was passed on the 22nd day of January 2001, to create a legal framework for protection of personal data in the country (Makulilo, 2012).

Unlike in Europe where member states transposed the provisions of the regional international instruments into their various municipal legislation on data protection, Cape Verde heavily relied on Portuguese Data Protection Law, which itself transposed the EU Data Protection Directive 95/46/EC (Wong, 2012)[3] before it was replaced by the EU General Data Protection Regulation (GDPR) (Traca & Embry, 2011). Between 2001 and 2014 when the first and only pan-African international treaty on data protection was adopted in Equatorial Guinea,[4] fourteen African countries[5] had already enacted their respective data protection laws without the benefit of drawing legislative inspiration from the convention as most of the laws were modelled after European data protection legal framework

[2] Data Governance Institute. Definition of Data Governance. https://datagovernance.com/the-data-governance-basics/definitions-of-data-governance/ (accessed online on 9 April 2022).

[3] The Directive was enacted by the European Union in 1995, to harmonize all the data protection laws within the Union and to regulate the processing of personal date stored in digital database.

[4] The AU Convention on Cybersecurity and Personal Data Protection was adopted in Malabo on the 27th day of June 2014.

[5] Cape Verde (2001), Mauritius (2004), Seychelles (2004), Tunisia (2004), Burkina Faso (2004), Senegal (2008), Morocco (2009), Benin (2009) Angola (2011), Gabon (2011), Lesotho (2011), Ghana (2012), Mali (2013) and Cote d'voire (2013).

(Greenleaf & Coltier, 2018; Schwartz, 2019; Mercer, 2020; Scott & Cerulus, 2018).

In the age of digitalization, legal frameworks for data protection are inevitable for Africa. More businesses and government are conducting meetings, transactions and data storage online. African Union has envisioned a digital single market (DSM), with legislation on it started in Tunisia and Senegal to support the business environment. Other policy frameworks such as Declaration of Principles on Freedom of Expression and Access to Information in Africa and the Declarations on Internet Governance are a basis of protection of digital rights and privacy of citizens. However, digital rights are still being violated in terms of Internet shutdowns and social media taxation (Boakye, 2021).

As these data protection legal frameworks are being adopted across Africa and the globe, the increased data produced and utilized in the digital space cannot be ignored. The digital rights of the producers and users are imperative for effective data use for socio-economic and political growth in Africa.

In a description approach, this paper examines the major international instruments regulating data protection in Africa by briefly chronicling the events that culminated in their adoption vis-à-vis their aims and objective. The paper also analyses the salient provisions of the instruments in the light of their applicability and the extent to which they have sharpened data protection compliance on the continent. This paper then analyses the nexus between data protection and data governance in Africa within the context of the regional instruments, and it then concludes with the necessity of legal framework for data protection in relation to data governance in Africa, with some recommendations that could be adopted to either develop new or strengthen the existing data protection framework especially in relation to data governance.

This paper, in a descriptive and normative manner, poses and analyses a number of questions thus:

1.1 What are the laws or quasi-legal guidelines (legal framework) that regulate or support data protection in Africa?
1.2 How does this legal framework measure up to international standards?
1.3 How enforced or enforceable is the framework on the continent?
1.4 In what manner does this framework influence or ought to influence data governance in Africa?
1.5 What are the incentives for data protection and data governance on the continent?

4.2 Legal Framework on Data Protection in Africa

Unlike what is obtainable in the European Union (EU) where the General Data Protection Regulation (GDPR) provides some sort of formidable harmonization of the erstwhile irregular data protection laws across the union, its African counterpart does not have a pan-African legislation that is immediately enforceable across board without domestication. This is not however to say that Africa does not have an existing legal framework on data protection; let us just say the elephant in the room remains the institutional capacity and political will to enforce the available instruments. Hereunder, the paper discusses the extant legal framework for data protection on the continent.

4.2.1 African Union Convention on Cyber Security and Personal Data Protection 2014 (Malabo Convention)

Conversations around regulation of the cyberspace effectively began in the late nineties when the committee of the United Nations General Assembly contemplated an instrument on 'disarmament and international security' whose deliberations were spearheaded by a draft resolution introduced by Russia in 1998 (Kavanagh, 2017). Upon Russia's proposal, the United Nations subsequently constituted a Group of Government Experts (GGE) engaged in the developments in the field of information and telecommunications in the context of international security.[6]

In one of its reports (UN General Assembly, 2021), the GGE notes that: 'The use of ICTs in future conflicts between states is becoming more likely, the risk of harmful ICT attacks against critical infrastructure is both real and serious and states are rightfully concerned about the danger of destabilizing misperceptions, the potential and economy deriving from the difficulty of attributing the source of an ICT incident' (Tikk & Schia, 2020). However, the UN's activities around cybersecurity did not spur many African countries into the anticipated regulation of the data

[6] The United Nations: Recent development in the field of information telecommunications in the context of international security. https://ccdcoe.org/incyder-articles/united-nations-recent-developments-in-the-field-of-information-and-telecommunications-in-the-context-of-international-security/. Accessed 25 May 2021.

protection as only eleven member states[7] instituted frameworks on data protection as of 2011 (Ball, 2017).

In 2011, the AU took a bold step towards regulating data protection when it published a draft AU Convention on Establishment of a Credible Legal Framework for Cybersecurity in Africa[8] which sought to, among other objectives, harmonize the laws of member states on data protection and sundry matters (Orji, 2012). In 2013, the draft was however reviewed and renamed the African Union Convention on the Confidence and Security in Cyberspace,[9] but it was also reviewed and went through another name change that culminated in the AU Convention on Cybersecurity and Personal Data Protection in 2014, which was preceded by a conference of experts from AU member states' ministry of justice where the content of the convention was thoroughly considered (Abdulrauf & Fombad, 2016).

Ultimately, on the 27th day of June 2014, during the 23rd ordinary session of the AU Summit in Malabo, Equatorial Guinea, the draft Convention on Cybersecurity and Personal Data Protection[10] was adopted by the heads of state to establish a credible framework for cybersecurity in Africa through 'protection of personal data etc.'[11]

The Malabo Convention has a total of 38 articles, preceded by a 20-paragraphed preamble. The Convention seeks to encourage member states to create frameworks and mechanisms to protect personal data and fundamental right as well as easing free flow of data within the continent. The first article defines essential data protection terms like consent, data controller, data subject, direct marketing, encryption, health data, personal data processing, recipient, sensitive data, third party and so forth but surprisingly omitted the definitions of equally important concepts like pseudonymization, data processor, data breach, data protection authority

[7] Cape Verde (2001), Mauritius (2004), Seychelles (2004), Tunisia (2004), Burkina Faso (2004), Senegal (2008), Morocco (2009), Benin (2009) Angola (2011), Gabon (2011) and Lesotho (2011).

[8] Found at https://au.int/en/cyberlegislation

[9] Found at https://ccdcoe.org/uploads/2018/11/AU-130101-DraftCSConvention.pdf

[10] Found at https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection

[11] The 23rd Ordinary Session of African Union ends in Malabo. See https://au.int/fr/newsevents/29258/23rd-ordinary-session-african-union-ends-malabo#:~:text=Malabo%2C%20Equatorial%20Guinea%2030%20June,from%2026%2D27%20June%202,014. Accessed 26 May 2021.

or supervisory authority, cross-border processing and so forth. While one may argue that the omission of such term does not superficially appear far-reaching, the convention is meant to be a compass for data protection laws on the continent as gleaned from its Articles 8(1) and (2) which seek to establish a framework for protection of 'physical' data and mechanism to ensure data processing guarantees the protection of fundamental rights. Yet even this falls short of the status of a legislative model in such material respect. Hence, it is desirable that the Convention is supplemented by relevant instruments to comprehensively define the omitted regular and fundamental data protection clauses, otherwise its enforcement may engender unimaginable conceptual confusion.

The convention applies to automated or non-automated processing[12] of personal data within the territory of a member state.[13] Like the GDPR, the Convention does not provide a definition or description of what constitutes 'automated' or non-automated processing, but the European law defines 'profiling' (Wiedemann, 2018). Automated processing has however been defined as 'a processing operation that is performed without any human intervention; conversely, non-automated processing is such that it is performed partly or wholly with human intervention.[14] Profiling and automated decision-making within the African context is increasing in the banking sector, especially with the rising development of FinTechs and proliferation of automated teller machines (ATM); however, there exists no pan-African legislation on this.

The convention requires member states to establish independent national authorities assigned with the statutory responsibility of ensuring that personal data within their respective territories are processed in accordance with the provision of the convention while keeping faith with the universal role of Data Protection Authorities (DPAs) (Giugiu & Larsen,

[12] Article 9(1), Malabo Convention.

[13] The 55 members of the AU include: Burundi, Cameroon, Central African Republic, Chad, Congo, DR Congo, Equatorial Guinea, Gabon, Sao Tome, and Principe, Comoros, Djibouti, Eritrea Ethiopia, Kenya, Madagascar, Mauritius, Rwanda, Seychelles, Somalia, South Sudan, Republic of Sudan, Tanzania, Uganda, Algeria, Egypt, Libya, Mauritania, Morocco, Sahrawi Arab Democratic, Tunisia, Angola, Botswana, Kingdom of Eswatini, Lesotho, Malawi, Mozambique, Namibia, South Africa, Zambia, Zimbabwe, Benin, Burkina Faso, Cape Verde, Cote d'voire, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Niger, Nigeria, Senegal, Sierra Leone and Togo.

[14] IAPP https://iapp.org/resources/article/automated-processing/. Accessed 29 May 2021.

2016). The Convention expects the respective national DPAs to educate the public on their data protection rights within their respective territories[15] while its membership is insulated from government influence, thereby underpinning their independence and impartiality (Greenleaf, 2012). As of June 2021, out of the 30 countries with proper data protection laws in Africa, only 20 have data protection authorities (DPAs),[16] while others are either yet to establish one or constitute its members. The Convention makes provisions for the duties and powers of the DPAs to include informing the public of their rights, issuing opinions, receiving and resolving complaints, data processing audit, imposing administrative sections, maintaining a data processing directory, regulating transborder transfer, establishing cooperation mechanisms with other national DPAs,[17] authorizing certain processing activities,[18] such as data involving genetic information, information on offences, national identification number, biometric data, historical and statistical data and so forth.

In what appears a renaming and rearrangement of the universally recognized principles of data protection, the Convention groups consent together with legitimate processing,[19] separate from the principle of lawfulness and fairness; it then fuses purpose with storage limitation,[20] accuracy and transparency as a stand-alone principle while confidentiality is grouped with security of personal data.[21] In all, the convention recognizes six re-designated principles, none of which contemplates the principle of data minimization or accountability as recognized under European law, even though it provides for specific principle in the event of processing sensitive personal data.[22] The consequence of such regrouping and muddling of principles would not only be evident in enforcing the clustered

[15] Art. 11(1)(a).

[16] Angola, Benin, Burkina Faso, Cape Verde, Chad, Côte d'Ivoire, Egypt, Gabon, Ghana, Kenya, Lesotho, Mali, Mauritius, Morocco, Niger, Nigeria, Sao Tome and Principe, Senegal, South Africa and Tunisia. See Paradigm Initiative and Olumide Babalola, 'Data Protection Authorities in Africa: A Report on the establishment, independence, impartiality and efficiency of data protection supervisory authorities in the two decades of their existence on the continent' Data protection, cybercrimes and/or cybersecurity laws Africa 2021 (werksmans.com) Accessed 26 October 2021.

[17] Article 12(2) (a)–(o).

[18] Art. 10 (4).

[19] Article 13 (1).

[20] Article 22.

[21] Article 13 (3)–(6).

[22] Article 14.

concept, it is potentially capable of confusing data controllers on their obligations thereunder.

The Convention, like most other data protection laws, recognizes data subject's right to information, right to access, right to object, rectification or erasure, but it again omits right to lodge complaint with regulator, right to data portability, restriction of further processes and so forth.[23] It also mandates data controllers to ensure confidentiality and security of personal data in their custody.[24]

Although the Convention was adopted in 2014, it is yet to enter force by reason of Article 36 makes it enforceable only thirty days after its ratification by fifteenth member states. As of 20th day of June 2021, only Angola, Ghana, Guinea, Mozambique, Mauritius, Namibia, Rwanda, Senegal and Zambia have ratified the Convention.[25] In spite of its limitation, Abdulrauf (2021) however argues in favour of the Convention's perceived expansive provision and authoritative stance, especially as far as they influence subsequent data protection legislation on the continent.

Though the AU Malabo Convention sought to enhance data protection in both the physical and digital data collection, use and storage, there are no articles which stipulated how to enhance digital rights in utilization and production of data online. There is evidence of digital rights violation in Africa in terms of online users' arrests and intimidation, Internet blocking and introduction of counterproductive laws and regulations undermine use technology and data generated to drive growth in Africa (CIPESA, 2019). The legal framework does not cover the sale and utilization of the data by third parties. This has seen infringement on privacy and sometimes adversely affects livelihoods of those affected. Furthermore, the issue of cybercrime is a matter of concern as more countries do private and government businesses online (Interpol, 2021).

[23] Articles 16 and 17.

[24] Articles 20 and 21.

[25] Status List found at https://au.int/sites/default/files/treaties/29560-sl-AFRICAN%20UNION%20CONVENTION%20ON%20CYBER%20SECURITY%20AND%20PERSONAL%20DATA%20PROTECTION.pdf. Accessed 20 June 2021.

4.2.2 Supplementary Act on Personal Data Protection Within the ECOWAS (ECOWAS ACT)

The Economic Community of West African States (ECOWAS) was established for the promotion of regional cooperation among member states especially for economic growth, among other objectives (Terwase et al., 2015). Its consequent ECOWAS Treaty mandates the harmonization and coordination of national policies and promotion of integration programmes in science, technology, legal matters and so forth (Ashiru, 2021).[26]

On the 16th day of February 2010, twelve heads of government within the ECOWAS gathered in Abuja, Nigeria, and adopted the Supplementary Act A/SA.1/01/10 on personal data protection within ECOWAS[27] (the Act) which predominantly seeks to regulate data protection within the member states.

The Act defines data protection terms like consent, data protection authority, personal data, sensitive data health data, data subject, data controller, data processor, third party and recipient[28] but omits important terminologies like processing, profiling, pseudonymization, anonymization, profiling, personal data breach, cross border and so forth. The consequence of the omission may however come to play when the Act is invoked to settle issues relating to transborder processing of data, especially before the regional courts, when faced with questions of conflict of laws and decision on lead national DPA and so forth (Estadella-Yuste, 1991). The Act applies to processing of personal data by public or private bodies by automated or non-automated means carried out within the ECOWAS, with exceptions.[29]

The Act mandates each member state to establish its own independent national DPA with parameters guaranteeing their impartiality and professional secrecy[30] and highlights the responsibilities of DPAs, their secrecy and powers to impose sanctions on erring parties.[31] In its own version of seven data protection principles, the Act states that processing is legitimate

[26] Art. 32(a) ECOWAS Reviewed Treaty 1993 found at https://www.ecowas.int/wp-content/uploads/2015/01/Revised-treaty.pdf

[27] https://www.statewatch.org/media/documents/news/2013/mar/ecowas-dp-act.pdf. Accessed 22 June 2021.

[28] Art. 1.

[29] Art. 3 and 4.

[30] Art.14 and 19.

[31] Art. 20.

where it is done with data subject's consent but gives exception where the requirement of consent can be dispensed with.[32]

The second principle of legality and fairness requires processing to be done in a legal, fair and non-fraudulent manner.[33] In what appears a bifurcation of some sort, the Act separates consent, which is a ground of lawful processing, from the principle of legality and fairness, which is fused under the EU principle of lawfulness, fairness and transparency (Kosta, 2013). As its third principle, the Act fuses purpose limitation, data minimization, storage limitation into one principle styled 'principle of purpose, relevance and preservation'[34] which requires data to be obtained for specific purpose, kept adequate and not kept beyond the required period. The principle also imports an element of the lawfulness principle.[35] Other principles are accuracy, purpose relevance and preservation, transparency, confidentiality and security and choice of data processor.[36]

Taking a cue from the European model on transborder transfer of data to third countries, the Act restricts transfer of personal data outside ECOWAS sub-region to only countries where there is an adequate level of protection (Wagner, 2018)[37] for fundamental rights and freedoms. Although the Act does not provide elaborate mechanisms for regulating such transfers, it simply mandates data controllers to inform DPAs before the transfers.[38] On data subject's rights, the Act recognizes right to be informed,[39] right to access,[40] right to object[41] and right to rectification and destruction.[42] Again, the Act omits vital data subject's right like right to restriction of further processing, right to data portability, right in relation to automated decision-making and so forth. The Act substantially

[32] Art. 23 (1) and (2).

[33] Art. 24.

[34] Art. 24.

[35] Art. 25.

[36] Articles 25, 26, 27, 28 and 29.

[37] This is a European concept which was recognized by the OECD Guidelines but became prominent under the repealed Data Protection Directive 95/46EC by regulating international transfer of personal data.

[38] Art. 36.

[39] Art. 38.

[40] Art. 39.

[41] Art. 40.

[42] Art. 41.

concludes on the obligations of data controller to be confidentiality, security, preservation and durability,[43] which however appear similar to data protection principles in their objectives.

4.2.3 Southern African Development Community (SADC) Model Law on Data Protection

In 2009, the imperativeness of creating a harmonized and uniform set of policies for the information communication technology industry for the sub-Saharan countries in the group of African, Caribbean and Pacific states necessitated the enactment and adoption of the Southern African Development Community (SADC) Model Law on Data Protection[44] which was adopted in 2013. Like many data protection laws, the Model Law defines terminologies such as consent, data controller, processor, data subject, genetic data, child personal data, processing, protection authority, recipient, sensitive data third party and transborder flow. The law however does not define anonymization, pseudonymization, profiling, personal data breach, data subject access request and so forth.

From the wording of Article 2, it appears that the scope of the law is not limited to the SADC sub-region as it only refers to 'given country' or 'territory'—terms which are not even defined therein. Even from the preamble, it does appear that the Model Law is not restricted to any region especially as contained in the concluding paragraph that:

It is with the above in mind that it is acknowledged that the protection of personal data involves the establishment of a specific and adapted regime to the participants of each region as set out in this Model Law.

It is however worthy of note that, in spite of its pan-African scope, the Model Law is a soft law without a legally binding effect on member states, but like the OECD Guidelines in Europe, it only provides a guide to member states on the approach to law-making on the data protection as well as an attempt at harmonizing the laws in the region (Shumba, 2015; Makulilo & Mophethe, 2016).[45]

[43] Art. 42–45.

[44] Found at https://www.itu.int/ITU-D/projects/ITU_EC_ACP/hipssa/docs/SA4docs/data%20protection.pdf. Accessed 19 June 2021.

[45] OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. https://www.oecd.org/digital/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm. Accessed 9 June 2021.

The Law envisages the establishment of independent regulator for member states to be constituted by judges appointed by the executive and non-governmental organizations with competent and requisite knowledge of data protection and the benefit of immunity.[46] Unlike other regional instruments in Africa, the Model Law provides the most comprehensive provisions on the nature, independence duties and powers of national DPAs, but it unfortunately contemplates the DPA reports to an undefined institution instead of the parliament[47] and thereby erodes its independence (Greenleaf, 2012). The Model Law recognizes the principle of data quality, lawfulness and purpose limitation, and it makes copious provisions on processing of sensitive and non-sensitive data, children's data and data relating to litigation,[48] but it however omits principles like data minimization, storage limitation accuracy, accountability, integrity and confidentiality and so forth. The Law outlines the duties of controllers in cases where personal data are collected directly from data subjects and otherwise, duty to ensure data security and accountability for third parties that access data through them, data breach or incident notification.[49]

The law recognizes the following data subject's rights: access, objection, automated decision-making right of representation and right to judicial redress.[50] Under the law, members of DPAs are meant to be administered to oath of secrecy[51] as they are empowered to impose fines on controllers for violation as well as prosecution of offenders in the law court.[52] The law subjects cross-border transfer of data to the relevant provisions of the national law adopted for the implementation of the Model Law, and this appears as the only provision that is fixated on member states of the SADC as it requires adequate level of protection before personal data can be transferred to non-member states.[53] Although the law references adequacy level, unlike the EU GDPR, it does not provide the parameters for determination of such level of protection (Wagner, 2018).

[46] Art. 3(2).
[47] Art. 3(10).
[48] Art. 11–17.
[49] Art. 21–27.
[50] Art. 31–38.
[51] Art. 41(1)
[52] Art. 42.
[53] Art.44 (1)(a).

Despite the laudable provisions of the Model Law, it merely serves as an advisory framework for the enactment of national laws as opposed to a legally binding instrument that can be ratified.[54]

4.2.4 East African Community (EAC) Legal Framework for Cyberlaws 2008

In its strides to deepen East Africa's regional integration via digital interconnectivity for the seamless provision of services, the East African Community constituted a Task Force which recommended a legal framework for cyberlaws[55] with the main objective of developing policies facilitating cooperation between member states (Mwiburi, 2019).

The Framework defines 'data protection' as the obligations assigned to entities processing personal data. It also recognizes that a data protection regime ought to guarantee certain data subjects rights.[56] Thereunder, data controllers are duty-bound to comply with muddled principles of accountability, transparency, fairness, lawfulness data accuracy, data security and processing limitation.[57] The Framework omits data minimization, purpose limitation and accountability but suggests a self-regulatory system to minimize the cost associated with conventional compliance enforcement approach.[58]

Without prejudice to Legal Framework's progressive but brief provisions on data protection, they are mere guides for member states but not legally binding on them until they transpose the provisions into their respective national laws (Greenleaf & Georges, 2014). It is worthy of note that the legal framework remotely or otherwise influenced the data protection legislation in Kenya, Uganda and Rwanda which passed the legislation afterwards.

[54] Parliament of the Republic of South Africa. https://pmg.org.za/files/RNW2764-150825.docx. Accessed 9 June 2021.

[55] Found at http://repository.eac.int:8080/bitstream/handle/11671/1815/EAC%20Framework%20for%20Cyberlaws.pdf?seq. Accessed 11 June 2021.

[56] Clause 2.5.

[57] Ibid.

[58] Ibid.

4.3 Interplay Between Data Governance and Data Protection in Africa

Data governance (DG) is the 'exercise of authority and control over the management of data' (Abraham et al., 2019). It also entails the trust reposed in data and its accountability for any adverse result occasioned by its poor quality. The whole gamut of data governance as a concept speaks to the data processing principle of accountability (Weber et al., 2009). Otto et al. (2007) define the concept as a 'companywide framework for assigning decision related rights and duties to be able to adequately handle data as a company asset'. It is the 'formal orchestration of people, process and technology to enable an organization to leverage data as an enterprise asset' (Zornes, 2006).

DG is concerned with the apportionment of responsibilities and liabilities among the various players in a data management system with respect to the decision-makers' rights and accountability over an entity's data assets. While data governance principally relates to collection and management of data that ensures effective and efficient use for the overall productivity of an entity (Cheong & Chang, 2007), data protection, on the other hand, safeguards the collected personal data[59] from misuse, compromise and/or corruption within the confines of certain principles. It is instructive that data governance is not restricted to personal data, but data protection in this context only protects the personal data managed alongside the big data (Elgendy & Elraga, 2014)[60] under the data governance framework. Hence, certain principles of data processing significantly impart data governance as far as personal data handled by the legal entity is concerned.[61] Unlike in Europe where the principles of data processing are uniformly provided by the GDPR (Bygrave, 2014), in Africa, the only readily binding and enforceable regional instrument is the ECOWAS Supplementary Act on Data Protection (Greenleaf & Coltier, 2020); the

[59] This is defined under the Malabo Convention as: "any information relating to an identified or identifiable natural person by which this person can be identified directly or indirectly in particular by reference to an identification number or to one or more factors specific to his/her physical, physiological, mental, economic, cultural or social identity."

[60] Defined as "Big data refers to datasets that are not only big, but also high in variety and velocity, which makes them difficult to handle using traditional tools and techniques".

[61] The principles of accuracy, storage limitation and accountability will be discussed in detail later in this paper.

AU Convention on Cybersecurity is not yet in force as its commencement provision has not been activated since less than 15 members have signed it.[62]

Notwithstanding its comatose state, the Convention provides for the principles of accuracy and storage limitation,[63] but it does not expressly provide for accountability.[64] However, this principle is an offshoot of the transparency principle (Alhadeff et al., 2021), hence since the convention provides for the latter, accountability can be discussed thereunder in relation to data governance. Since the Malabo Convention is pan-African in its coverage, I will discuss some of its principles which interplay with data governance in Africa albeit in its current unenforceable predicament.

4.3.1 Accuracy Principle

Accuracy is one of the components of data quality (Cong et al., 2017). The principle of accuracy entails the accuracy, completeness and consistency of data, and it goes without saying that organizations require the highest quality of data for them to function optimally (Joshi, 2020). In an entity's use of (personal) data, privacy issues like transparency, security and compromise of (personal data) are always thrown up and sometimes the relevant questions are left unanswered. Bair notes that data quality is defined by 'data type and domain, completeness, uniqueness, and referential integrity, consistency across all data bases, freshness and timeliness and business rules conformance' (Bair, 2004).

On the relationship between the principle of accuracy and data governance, Cohn argues that 'data governance is a catalyst for quality and value is derived from well governed quality data. Relevant, timely, consistent, reliable and accurate data is an expectation, and it is not achieved

[62] List of Countries Which Have Signed, Ratified/Acceded To The African Union Convention On Cyber Security And Personal Data Protection. https://au.int/sites/default/files/treaties/29560-sl-AFRICAN%20UNION%20CONVENTION%20ON%20CYBER%20SECURITY%20AND%20PERSONAL%20DATA%20PROTECTION.pdf. Accessed 22 June 2021.

[63] Article 13 (3).

[64] The principle has been part of the European data protection law since 1980 when it was introduced into the Organization for Economic Cooperation and Development's 'The Recommendations of the Council Concerning Guidelines Governing Protection of Privacy and Transborder flow of Personal Data' on 23 September 1980. https://www.oecd.org/digital/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm. Accessed 21 June 2021.

serendipitously' (Cohn, 2015). In data protection parlance, the principle of data quality (Hoeren, 2018) requires personal data to be effective, fit, relevant and all-embracing for its intended purpose of processing.[65] The principle stipulates that, when organizations collects data for decision-making, they must ensure that such personal information is not only utilized in a manner that is relevant to the purpose of collection, but it must also be accurate, wholesome and regularly updated. This will ensure that personal information used for critical organizational decisions is accurate to prevent undue violation of fundamental rights and freedoms of data subjects (Bygrave, 2002).

Under this principle, organizations (private and public) are duty-bound to ensure the accuracy of information they keep and the opinions that they express regarding data subjects, especially when decisions affecting the latter are made (Hallinan & Borbesius, 2020). It mandates data controllers to take reasonable steps to ascertain aptitude of personal information processed within the context of their organizational activities. As a representation of this principle, Article 13, Principle 3 of the Malabo Convention mandates data collection to be adequate, relevant and not excessive in relation to the purposes for which they there collected.[66] Ultimately, legal entities must set up mechanisms to ensure the validity and quality of personal data in their custody by imbibing a business culture of periodic updates and timely deletion of the outdated or irrelevant ones.

4.3.2 Storage Limitation

Storage of data is one of the main components of data governance. Sometimes, they are stored indefinitely in unregulated and unguarded databases for the controllers' whimsical analyses and/or utility, oftentimes without the consent of data subjects (Pike, 2020).[67] The passage of data protection legislation in Europe, for example, threw many organizations into panic mode, especially when auditing the legal bases for collecting and/or storing online visitors' data through their digital platforms

[65] 'Processing' is a technical word for alteration, use, transmission, collection, storage, destruction, transfer and any operation or sets of operation on personal data as defined by Article 1 of the Malabo Convention.

[66] Art. 13 (4) provides that data collected shall be accurate and, where necessary, kept to date.

[67] Elizabeth R. Pike, 'Defending Data: Towards Ethical Protections and Comprehensive Data Governance' (2020) 69(687) Emory Law Journal, 687.

without necessarily fixing the mechanisms for obtaining informed consent (Francesco et al., 2021). On the other hand, while businesses are not precluded from storing customers' personal data, such storage must be within the confines of the applicable data protection laws and its exceptions (Duceto, 2020). For example, data processing for research purposes constitutes one of the exceptions to the principle of storage limitation, since data can be kept for longer than necessary, especially for verification of research results (Pormeister, 2017).

The indiscriminate and indefinite storage of customers and other data subjects' personal data by organizations poses unimaginable privacy risks attributable to unregulated and, most times, insufficient and inadequate technical and organization security measures (if any) by data controllers (Biega & Finick, 2021). Essentially, personal data must not be kept in a form that identifies data subjects for longer than is justifiable by law. Where personal data are no longer needed or have become irrelevant or out of date, data controllers can either outrightly delete, anonymize or pseudonymize them in certain cases (Mourby et al., 2018).

Under the AU Malabo Convention, the storage limitation is however not a principle but an obligation on the data controllers. Article 22 emphatically prohibits personal data from being kept for longer than necessary for its purpose of collection, but the provision is of exceptions or parameters for the applicable retention period. The principle interplays with data subject's right to be forgotten or deletion or erasure of personal data which is no longer relevant or up to date. Without prejudice to the circumstances surrounding an organization's collection of personal data, this principle still operates to provide them from keeping and/or storing the data for longer period than reasonably necessary. Hence, once data have been used for the purpose of collection, it behoves the organization to immediately delete or anonymize such personal data to reduce the risk of violating the principles of data minimization and accuracy when they become irrelevant, surplus to requirement, inaccurate or outdated. There are no specific retention periods in the regional instruments; however, resort should be had to the relevant national laws on data retention limits, but ultimately a formidable data governance policy ought to be devised to plug the legislative gaps in this regard.

Legislative and stakeholder's engagement for data governance however becomes very important when it is considered that out of 55 African countries, at least 49 have (or about to) enacted laws or regulations requiring prospective subscribers to provide personal data as conditions to own telephone lines (Donovan & Martyin, 2013), but sadly, only about 19 of those countries have established Data Protection Authorities[68] to enforce compliance with relevant data protection laws.

4.3.3 Accountability

This principle originated from the OECD Guidelines[69] of 1980 and repeated in its revised version of 2013. The principle principally requires legal entities to acknowledge and assume liability for their operations on personal data in the course of the organizational activities. Data controllers have the bounden duty of demonstrating adequate technical and organizational measures to secure data in compliance with the relevant data protection legislation for the ultimate protection of data subjects' rights (de Hert et al., 2012). In compliance with this principle, legal entities are obliged to document the observance of their obligations under the relevant data protection legislation (Becker, 2019).

Accountability is not expressly provided under the Malabo Convention, but the principle is closely linked to the principle of transparency, and it has been regarded as a privacy and data protection—enhancing principle (Guagnin & Leon, 2012; Zimnerman & Cabinakova, 2015). In demonstrating their accountability, organizations must take hands-on approach to data protection and privacy issues by adopting effective and

[68] Angola (Agência de Protecção de Dados); Benin (Autorité de Protection des Données à caractère Personnel); Burkina Faso (Commission de l'Informatique et des Libertés); Cape Verde (Agência de Protecção de Dados); Chad (Agence Nationale de Sécurité Informatique et de Certification Électronique); Côte d'Ivoire (Autorité de Régulation des Télécommunications de Côte d'Ivoire); Egypt (Personal Data Protection Centre); Gabon (Commission nationale pour la protection des données à caractère personnel); Ghana (Data Protection Commission); Kenya (Office of Data Protection Commissioner); Madagascar (Commission Malagasy de l'Informatique et des libertés); Mali (Autorité de Protection des Données à Caractère Personnelles); Mauritius (Data Protection Office); Morocco (Commission Nationale de Contrôle de la Protection des Données à Caractère Personnel); Niger (Haute Autorité de Protection des Données à caractère Personnel); Nigeria (National Information Technology Development Agency); Sao Tome and Principe (Agência Nacional de Protecção de Dados Pessoais); Senegal; (Commission des Données Personnelles); South Africa (Information Regulator); and Tunisia (Instance nationale de protection des données personnelles).

[69] Article 14 provided that: "A data controller should be accountable for complying with measures which give effect to the principles stated above.'

contemporary measures, which are not only discernable at a glance but transparently demonstrable upon regulatory request or audit (Falk, 2016).

Data controllers must take full responsibility for how they directly or indirectly deal with data and implement appropriate measures and documentation in proof of their compliance with applicable laws (Bennet, 2021). They are responsible and must demonstrate data quality.

4.3.4 Confidentiality and Integrity

This is recognized under Principle 6 of the Malabo Convention. This principle simply mandates organizations processing personal data to employ appropriate organizational and technical measures to protect such personal information from misappropriation, corruption, theft and/or destruction. Confidentiality in this sense speaks to the duty of the organization handling data to ensure that such information is not shared or exposed to unintended persons while keeping it as safe and secret as technically possible.

4.4 Incentives of Legal Framework for Data Protection/Data Governance in Africa

The benefits of data protection to data governance are numerous. However, for the purpose of this paper, I shall briefly discuss the incentives from the rights protection and economic gains for organizations and governments.

4.4.1 Privacy Right Guarantees

Even though the African Charter does not expressly recognize privacy as a fundamental right, it does not rule out African's entitlement to enjoy private family life.[70] This idea of a privacy entitlement for individuals is what also underpins the notion of data protection. In fact, data protection originated from the right to privacy, hence a proper and formidable legal framework for data protection would not only guarantee certain data subjects rights, but it would also ensure considerable control over their personal information and ultimately repose consumers' trust in the processing activities.

[70] Article 18, African Charter of Human and Peoples Rights enjoins member states to protect the family—'natural unit and basis of society', African Union (1981).

4.4.2 Healthy Democracy

A healthy democratic state is one in which its citizens can make informed and autonomous choices (Forde, 2016). Yet processing data without consideration for the impact it may have on individuals may have the effect of limiting the ability of individuals to make choices or limit the choices available to such individuals in a way that limits their autonomy (Feldman, 1994). This is even more crucial in today's world of technological reliance where automated processing and digital identities are gradually becoming more significant determinants of an individual's real-life choices. Data protection laws militate against this. The idea that when consent is relied upon as the legal basis for processing, it must be informed and must not be obtained using coercive tactics echoes these concerns for democratic autonomy. Furthermore, even in instances where data are processed without consent, the notion of data subjects' rights and the transparency, fairness and accountability obligations grant the much-needed controls individuals need to maintain their ability to make truly free choices. Thus, it is safe to conclude that a world where data protection is respected is one in which the seeds of corporate or governmental totalitarianism cannot flourish.

4.4.3 Economic Gains from Free Flow of Data

The concept of free flow is not merely one where there are no legal barriers to cross-jurisdictional data transfers. Instead, it entails that where these legal barriers exist, they do not impose data localization requirements Data localization requirements have the direct effect of raising the costs for doing business across jurisdictions. Particularly for data-driven businesses such as cloud service providers, these costs have the added effects of posing significant barriers for entry into new markets within the continent. This disincentivizes the creation of such businesses and creates an environment that limits the growth of African start-ups and SMEs. Furthermore, the free flow of personal data would ease information dissemination and beneficial collaboration of businesses and corporate entities within the region. However, these benefits extend to collaboration opportunities outside the region. The European Data Protection framework is setting the global trend for technological collaborations across the world.

Implementing and initiating an African data protection framework may create an opportunity for the recognition of African countries as having an adequate level of data protection. This has the potential to facilitate more cross-border collaboration and even non-data-driven businesses looking to partner with African businesses in a capacity.

4.5 Conclusion

Data governance predominantly speaks to the management of data for organizational growth. Experience has shown that the management of big data would always involve the handling of personal data, hence the activation of data protection principles.

African currently has only one binding regional instrument—ECOWAs Supplementary Act on Data Protection—among other international instruments with provisions on various principles that persuasively impart data governance on the continent. Out of 55 African countries, only 30 have fully dedicated data protection laws and 19 of them have established DPAs to enforce compliance with the laws; hence, it is crystal clear that data governance on the continent remains largely unsupported by legislative and enforcement framework.

With the magnitude of personal data exchanged, stored or transmitted within the data governance system in Africa, an appropriate and formidable data protection legal framework becomes essential to guarantee users of control over their personal information, on the one hand, and to regulate the data controllers' processing/management of such information against misuse, compromise, theft or other untoward dealings with personal data, on the other hand.

For a properly regulated data governance, it is hoped that African countries would evenly ratify the Malabo Convention and strengthen their respective municipal data protection legal frameworks to complement public and private organizational management of personal data not only within their respective territories but also across the continent.

The transborder cooperation of national DPAs envisaged by the regional treaties ought to be encouraged and strengthened to boost enforcement of regional and municipal data protection laws with the aim of enhancing trans-border flow of data and international data governance within the confines of uniform cross-border data protection rules.

References

Abdulrauf, L. A. (2021). Giving ‘teeth’ to the African Union towards advancing compliance with data privacy norm. *Information and Communication Technology Law, 30*(2), 1.

Abdulrauf, L. A., & Fombad, C. M. (2016). ‘The African Union’ data protection convention 2014: A possible cause for celebration of human rights in Africa. *Journal of Media Law, 8*(1), 1–8.

Abraham, R., vom Brocke, J., & Schneider, J. (2019). Data governance: A conceptual framework, structured review and research agenda. *International Journal of Information Management, 49*, 1.

African Union. (1981). *Article 18, African Charter of Human and Peoples Rights enjoins member states to protect the family – ‘natural unit and basis of society’.*

Alhadeff, J., van Alsenoy, B., & Dumorhier, J. (2021). The accountability principle in data protection regulation: Origin, development and future directions. In D. Guagnin, C. Iiten, D. Neyland, L. Hempel, I. Kroener, & H. Postigo (Eds.), *Managing privacy through accountability* (p. 15). Palgrave Macmillan.

Ashiru, A. (2021). A comparative analysis of the legal framework for the criminalization of cyberterrorism in Nigeria, England and the United States. *NAUJILJ*, *12*(1), 99, 107.

Bair, J. (2004). *Practical data quality: Sophistication levels?* http://www.knightsbridge.com/pdfs/in_the_news/. Accessed 21 June 2021.

Ball, B. M. (2017). Introductory note to African union convention on cyber security and personal data protection. *The American Society of International Law*, *165*.

Becker, R. (2019). A data information system for accountability under the general data protection regulation. *Giga Science, 8*(12), 122.

Bennet, C. J. (2021). The accountability approach to privacy and data protection: Assumptions and caveats. In D. Guagnin, C. Iiten, D. Neyland, L. Hempel, I. Kroener, & H. Postigo (Eds.), *Managing privacy through accountability* (p. 15). Palgrave Macmillan.

Biega, A., & Finick, M. (2021). *Reviving purpose limitation and data minimization in personalization, profiling and decision-making system*. Max Planck Institute for Innovation and Competition Research Paper No. 21.04, 1, 5.

Boakye, B. (2021). *Tech policy in Africa. Emerging trends in internet law and policy.*

Bygrave, L. A. (2002). *Data protection law: Approaching its rationale, logic and limits* (p. 1). Kluwer International.

Bygrave, L. A. (2014). *Data privacy law. An international perspective* (p. 145). Oxford University Press.

Cheong, L. K., & Chang, V. (2007). The Need for Data Governance: A case study. *18th Australasian conference on information system*, 999.

CIPESA. (2019). *Digital rights in Africa: Challenges and policy options.* Uganda.

Cloudian. *Data protection and privacy: 12 ways to protect user data*. https://cloudian.com/guides/data-protection/data-protection-and-privacy-12-ways-to-protect-user-data/. Accessed on April 9, 2022.

Cohn, B. L. (2015). Data governance: A quality imperative in the era of big data, open data and beyond. *Journal of Law and Policy for the Information Society, 10*(3), 812.

Cong, G., Fan, W., Geerts, F., & Ma, S. (2017). Improving data quality: Consistency and accuracy. *Conference: Proceedings of the 33rd international conference on very large data bases*. University of Vienna, Austria, September 23–27, 1.

de Hert, P., Konstantinou, V. P., Wright, D., & Gutwirth, S. (2012). The proposed regulation and the construction of a principles driven system for individual data protection. *The European Journal of Social Science Research, 26*(1), 1.

Donovan, K. P., & Martyin, A. K. (2013). *The rise of African SIM registration: The emerging dynamics of regulatory change*. https://firstmonday.org/ojs/index.php/fm/article/view/4351/3820. Accessed 15 June 2021.

Duceto, R. (2020). Data protection, scientific research and the role of information. *Computer & Security Review, 37*, 1–5.

Elgendy, N., & Elraga, A. (2014). Big data analytics: A literature review paper. *Lecture Notes in Computer Science, 8557*, 214–227.

Estadella-Yuste, O. (1991). Transborder data flows and the sources of public international law. *North Carolina Journal of International Law and Commercial Regulation, 16*(2), 380–412.

Falk, T. T. (2016). *The concept of accountability as a privacy and data protection principle*. https://www.cpomagazine.com/data-privacy/concept-accountability-privacy-data-protection-principle/. Accessed 17 June 2021.

Feldman, D. (1994). *Secrecy;* dignity or autonomy? Views of privacy as a civil liberty. *Current Legal Problems, 47*(2), 42, 54.

Forde, A. (2016). The conceptual relationship between privacy and data protection. *Cambridge Law Review, 1*, 135–137.

Francesco, G., Palazzani, L., Dimitiou, D., Domingo, J. D., Fons-Martinez, J., Jackson, S., Vignally, P., Tozzi, A., & Rizzo, C. (2021). *Digital tools in the informal consent process: A systematic view*. https://www.researchsquare.com/article/rs-1273/v2. Accessed 13 June 2021.

Giugiu, A., & Larsen, T. A. (2016). Role and power of national data protection authorities. *EDPL, 3*, 342.

Greenleaf, G. (2012). Independence of data privacy authorities: International standards and Asia-Pacific experience. *Computer Law & Security Review, 28*(1), 1.

Greenleaf, G., & Coltier, B. (2018). Data privacy laws and bills: Growth in Africa, GDPR influence. *152 Privacy Laws & Business International Report*, 11.

Greenleaf, G., & Coltier, B. (2020). Comparing African data privacy laws: International, African and regional commitments. *University of New South Wales Law Research Series, 1*, 21.

Greenleaf, G., & Georges, M. (2014). African regional privacy instruments: Their effect on harmonization. *132 Privacy Law and Business International Report*, 19–21.

Guagnin, D., & Leon, H. (2012). *Managing privacy through accountability* (p. 15). Palgrave Macmillan.

Hallinan, D., & Borbesius, F. Z. (2020). Opinions can be incorrect (in our opinion) on data protection law's accuracy principle. *International Data Privacy Law, 10*(1), 1.

Hoeren, T. (2018). Big data and data quality. In T. Hoeren & B. Kolany-Raiser (Eds.), *Big data in context legal, social and technological insights* (p. 2). Springer.

Interpol. (2021). *African Cyberthreat Assessment report: Interpol's Key Insight into Cybercrime in Africa*.

Joshi, A. (2020). *Data quality and data governance, where to begin*. https://www.collibra.com/blog/data-quality-vs-data-governance. Accessed 11 June 2021.

Kavanagh, C. (2017). *The United Nations, cyberspace and international peace and security: Responding to complexity in the 21st century*. https://www.unidir.org/files/publications/pdfs/the-united-nations-cyberspace-and-international-peace-and-security-en-691.pdf. Accessed 25 May 2021.

Kosta, E. (2013). *Consent in European data protection law* (p. 195). Nijhoff Publishers.

Makulilo, A. B. (2012). Privacy and data protection in Africa: A state of the art. *International Data Privacy Law, 2*(3), 163.

Makulilo, A. B., & Mophethe, K. (2016). Privacy and data protection in Lesotho. In *African data privacy laws* (pp. 337–347). Springer International.

Mayer-Schonberger, V. (1997). Generational development of data protection in Europe. In P. Agre & M. Rotenberg (Eds.), *Technology and privacy: The new landscape* (pp. 219–242). The MIT Press.

Mercer, S. T. (2020). The limitations of European data protection as a model for global privacy regulation. *AJIL Unbound, 114*, 20–25.

Mourby, M., Mackey, E., Elliot, M., Gowans, H., Wallace, S. E., Bell, J., Smith, H., Aidinlis, S., & Kaye, J. (2018). Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK. *Computer Law & Security Review, 34*, 222–233.

Mwiburi, A. J. (2019). *Preventing and combating cybercrime in East Africa. Lessons from Europe's cybercrime frameworks* (p. 141). Duncker & Humblot.

Oman, S. (2010). *Implementing data protection in law*. Stockholm Institute for Scandinavian Law.

Orji, U. J. (2012). The defects of the draft African Union convention on the establishment of a credible legal framework for cybersecurity. *IEEE*, 1.

Otto, B., Wende, K., Schmidt, A., & Osl, P. (2007). Towards a framework for corporate data quality management. *16th Australasia conference on information systems University of Southern Queensland, Toowoomba Australia* (pp. 916–926).

Pike, E. R. (2020). Defending data: Towards ethical protections and comprehensive data governance. *Emory Law Journal, 69*, 687.

Pormeister, K. (2017). Genetic data and the research exemption: Is the GDPR Going too far? *International Data Privacy Law, 7*(2), 137–140.

Schwartz, P. M. (2019). Global data privacy: The E.U. way. *New York University Law Review, 94*, 771.

Scott, M., & Cerulus, L. (Jan. 31, 2018). Europe's new data protection rules export privacy standards worldwide. *POLITICO.*

Shumba, T. (2015). Revisiting legal harmonization under the Southern African development community treaty: The need to amend the treaty. *Law Democracy Development, 19*, 1.

Terwase, I. T., Abdul-Talib, A.-N., & Zengeni, K. T. (2015). The role of ECOWAS on economic governance, peace and security perspectives in West Africa. *Mediterranean Journal of Social Sciences, 6*(3), 257.

Tikk, E., & Schia, N. N. (2020). The role of the UN security council in cybersecurity. In E. Tikk & M. Kerttunen (Eds.), *Routledge handbook of international cybersecurity* (p. 354). Routledge.

Traca, J. L., & Embry, B. (2011). An overview of the legal regime for data protection in Cape Verde. *International Data Privacy Law, 1*, 3.

UN General Assembly. (2021). *Final substantive report.* https://front.un-arm.org/wp-content/uploads/2021/03/Final-report-A-AC.290-2021-CRP.2.pdf. Accessed 26 October 2021.

Wagner, J. (2018). The transfer of personal data to third countries under the GDPR: When does a recipient country provide an adequate level of protection? *International Data Privacy Law, 8*(4), 318–337.

Weber, K., Otto, B., & Osterle, H. (2009). One size fits all – A contingency approach to data governance. *Journal of Data and Information Quality (JDIQ), 1*(1), 1–27.

Wiedemann, K. (2018). *Automated processing of personal data for the evaluation of personally traits: Legal and ethical issues.* Max Plank Institute for Innovation and Competition Research Paper No. 18-04, 3.

Wong, R. (2012). The Data Protection Directive 95/46/EC: Idealisms and Realisms. *International Review of Law Computers & Technology, 1.*

Zimnerman, C., & Cabinakova, J. (2015). A conceptualizing of accountability as a privacy principle. In W. Abramowicz (Ed.), *BIS* (p. 1). Springer.

Zornes, A. (2006). Corporate data governance best practice. *The CDI Institute Market Plus TM Depth Report*, 1.

CHAPTER 5

Data Regulation in Africa: Free Flow of Data, Open Data Regimes and Cybersecurity

Hanani Hlomani and Caroline B. Ncube

5.1 Introduction

Fast-paced technological advancements have made it difficult for legal scholars, policymakers, and legislators to stay abreast of all the considerations and important policy debates that are necessary to ensure that the law is not outpaced and eventually invalidated by technology. The Covid 19 pandemic period (2020–2022), evidenced that it is possible for the world to move to a partially or fully digitised ecosystem. It has already been said that several global corporations are contemplating whether reverting to the old way of working in the post pandemic period would make sense given how efficiently the world has been able to adapt, collaborate and produce results in a digital ecosystem. At the centre of it all, has been the need to move data from device to device, from one location to another and from one person to the next. This creates a legal conundrum for those tasked with legislating and policymaking, who have the task of formulating

H. Hlomani (✉) • C. B. Ncube
Department of Commercial Law, University of Cape Town,
Cape Town, South Africa
e-mail: HLMHAN003@myuct.ac.za

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8_5

sound policies and legislative instruments that ensure that such data can move freely, lawfully and without impeding on any personal or commercial interests in a safe digital environment that is protected against cyberattacks.

Data can be defined as pieces of information that can either be qualitative or quantitative (OECD, 2021). Such information can be abstract or about one or more persons. Data can also be defined as a collection of facts such as words, numbers or observations or a way of describing things. The term "data" is not to be confused or used interchangeably with the term "information". This is because data is a collection of facts that are unstructured and unorganised whereas information relates to how one understands those unorganised facts contextually (Diffen, 2021). Because of the nature of the digital ecosystem, that is heavily reliant on decentralisation, data has been coined as the "new frontier for economy" after gold (Manzo, 2019). This is principally because data is the means through which devices communicate with one another and the main asset on which markets, research, governments and corporate companies rely daily. Given the decentralisation of modes of communication and the flood in web connectivity, a lot of data is exchanged amongst users of the internet and may include information ranging from personal details of people to anything in between of a non-personal nature. As more data becomes available and accessible, the practice of data analytics becomes increasingly important. Data analytics is a process that uses advanced analytic techniques such as predictive analysis, statistical analysis and data mining on sets of data to discover new facts, to predict future events or behaviours or to explain past phenomena that previously had no logical explanations (Russom, 2013). Such analytic techniques have the potential to impact a number of economic sectors across the African continent positively mostly by equipping various stakeholders with information that they previously did not have and from a wealth of resources that were previously blocked by legislation or as a result of geographical barriers.

Currently, debates around data regulation hinge on whether the data in question is personal or non-personal because of the starting premise that from a regulatory approach, personal and non-personal data should not be subjected to the same scrutiny. Personal data can be defined as any information which are related to an identified or identifiable natural person.[1] This means that a data subject is identifiable if it is possible to directly/indirectly identify the subject through identifiers such as name,

[1] Art 4(1) of the GDPR.

identification number, location data, physical, genetic data, cultural data, etc.[2] Practically, this can also include all data which are or can be assigned to a person such as the telephone, credit card or personal numbers of a person. The opposite and therefore the definition of non-personal data is electronic data that does not contain any information that can be used to identify a natural person. Examples include data that is non-personal to begin with, for example, weather data and stock prices, or it can be data that was previously personal in nature but has become anonymised (void of all personal data) (Indian Express, 2020). As has been stated, the rate at which technology is advancing makes it increasingly difficult to keep up with how best to regulate data. The African continent has seemed to be more concerned with the protection of personal data of its citizens. It is estimated that approximately 24 of Africa's 55 countries have enacted or embraced some form of regulation, with the chief aim of protecting personal data, while there are currently about 4 draft laws (Alt Advisory, 2023). This has largely been attributed to the enactment of the European General Data Protection Regulation (GDPR)[3] which was adopted in 2016 and is very influential due to its regulation of cross border data flows, which has impacted a number of countries data protection models globally (Daniel, 2021). However, most innovation-driven countries have realised the value in formulating regulatory regimes that protect personal data, while ensuring, at the same time, that non-personal data can be extracted from personal data so that innovation is fostered. In other words, in addition to the value of naturally non-personal data, there is also value in data that was locked away in data sets that have personally identifiable information.

As data regulatory framework take root in Africa and robust digital space, enabling production and consumption of data, digital rights cannot be understated. Digital rights encompass online privacy, freedom of expression, access to internet without arbitrary shutdown and costly internet prices (Hutt, 2015). A big part of upholding digital rights entails being able to use digital spaces without social, political or economic hindrances and being victims of criminal activities through cybercrime. Cybercrime is a concern in Africa, Interpol (2021). The incidences of cybercrime include online scams, digital extortions, business email compromise, ransomware

[2] 'Personal Data' *General Data Protection Regulation (GDPR)* <https://gdpr-info.eu/issues/personal-data/> accessed 17 January 2022.

[3] The General Data Protection Regulation 2016/679.

and botnets. Lack of standard and protocols to mitigate against the cybercrime is prevalent in Africa where approximately 570 million people are online as at 2022 (Statista, 2023).

In view of the above, this paper will seek to address the following issues. Firstly, it will interrogate the manner in which data (both personal and non-personal) has been dealt with by the EU as a yardstick, with the aim of imagining how a harmonised data regulation system that encompasses both personal and non-personal data would look either on a continental scale or on a regional scale within Africa. Emphasis will be placed on determining how legal instruments guide or mandate the identification of non-personal data. This is because, in addition to the already existing efforts to protect personal data, it is necessary to regulate how non-personal data is used to ensure that both personal and non-personal data are able to move freely across the continent and across the globe underpinned by sound regulation and policies. Such free movement is necessary to reap the potential economic benefits and may aid in realising Africa's development agenda and the sustainable development goals (SDGs). With the central theme being that of the free movement of data, the paper will then delve deeper into aspects such as the use of and liberalisation of open data policies (the notion that specific data should be freely available for use and reuse, especially public sector information). The paper will then discuss, in addition to the potential legal conundrums that come with the liberalisation of data and its movement, the cybersecurity concerns that come with such a regulatory regime. This would be done by considering how the Malabo Convention protects personal data as well as Regional Economic Communities (RECs) and individual approaches by AU member states to cybersecurity.

5.2 Defining Personal Data

Demarcating what falls within the boundaries of personal data has befuddled scholars for a while. While obvious data such as names, ID numbers etc. are unmistakably personal data, the EU's definition of personal data in Article 4.1 of the GDPR defines personal data as any information which are related to an identified or identifiable natural person. Since the definition includes "any information", one must assume that the term "personal data" should be as broadly interpreted as possible. It is for this reason that the bounds of what data is personal or not are constantly being debated.

In the *Breyer* case,[4] the Court of Justice of the European Union (CJEU) in attempting to define what "personal data" is, held that any piece of information, that when additional information is sought from a third party, is able to identify a data subject shall constitute personal data (Bird & Bird, 2019). As such, if one were to apply the principles of *Breyer* practically, the likelihood of data which initially presented itself as non-personal data may eventually fall within the ambit of the GDPR's definition of personal data. As that is the case, failing to account for non-personal data may mean that it is subjected to the same restrictions as personal data or other data localisation requirements.

In addition to the blurred lines on what constitutes personal data, data localisation requirements also pose a threat to radical economic transformation on the African continent on the back of the data revolution. Data localisation requirements are typically restrictions on the flow of data from one country to another. For example, it may be required by law that all processing of data relating to a certain country's citizens be carried out using servers located within such a country's borders and thus making it illegal to process such data anywhere but within that territory (Bird & Bird, 2019). Such restrictions raise the cost of doing business across borders and in a digital ecosystem, the threat to efficiency is real. Further, they stifle the access of businesses and public sector bodies to cheaper and more innovative services or force companies operating in multiple countries to contract excess data storage and processing capabilities.[5] For start-ups and SMEs, this constitutes a serious obstacle to growth, to entering new markets and to the development of new products and services. The EU has adopted a regulatory framework for the free flow of non-personal data in the EU (European Union, 2018) which lists some of the non-personal data as being data generated by artificial intelligence, the Internet of Things and machine learning as potential sources of non-personal data along with a few very specific examples.[6]

[4] *Patrick Breyer v Bundesrepublik Deutschland* Case C-582/12.

[5] Ibid.

[6] Bird and Bird supra note 16.

5.3 Current African States' Approaches and Significance for the African Continental Free Trade Area (AfCFTA)

The African Union (AU) in 2014 adopted the Convention on Cyber Security and Personal Data Protection at the Twenty-third Ordinary Session of the Assembly, held in Malabo, Equatorial Guinea (known as the Malabo Convention), which has only recently come into force (May 2023) following the last necessary ratification by Mauritania. This convention much like the GDPR focuses on personal data and cybersecurity. At the time of writing, October 2021, the AU Commission has also formulated the Africa Data Policy Framework, which is informed, in part, by the Malabo Convention (African Union, 2022). Several regional economic communities (RECs) have also adopted regulatory instruments which will be summarised in Sect. 5.4. Outside of this concerted effort, very little has been done in terms of a collective continental/regional legislative instrument on data protection with most countries opting to attempt protection individually.

As the continent moves to realise the promises of the AfCFTA, it will be important to have a measure of harmonisation in regulatory frameworks so that inter-Africa trade is enhanced. Businesses and individual entrepreneurs trading in different countries across Africa would benefit if they had some assurance that similar principles of data protection and data governance models are aligned across the continent. E-commerce and digital trade grew exponentially during the last two years due to restrictions on physical interactions between persons to curb the spread of the COVID-19 pandemic. The AU is spearheading the growth of digital trade through the Digital Transformation Strategy for Africa 2020–2030 (African Union, 2020a) and in this context, an e-commerce protocol is being negotiated under the AfCFTA Agreement. (African Union, 2020b) It is envisaged that the AfCFTA and the Digital Transformation Strategy for Africa will catapult Africa's digital economy (Chaytor, 2020). As elaborated in Sect. 5.3, the free movement of data is a core element of promoting inter-Africa trade and cross-border data flows to bring significant benefits. Beyond the AfCFTA e-commerce protocol negotiations, there is global momentum on similar negotiations. Specifically, the World Trade Organization (WTO) began negotiations on trade-related aspects of e-commerce in January 2019 and these are continuing (WTO, 2021). The development of common African approaches will therefore be instrumental in shaping the global WTO agenda.

5.4 The European Union's Approach to Data

The EU has been at the fore of establishing a comprehensive regulatory framework on data of both a personal and non-personal nature. Since 2014, the European Commission has developed a number of directives and laws to facilitate the development of a data-agile economy. Examples include the Regulation on the free flow of non-personal data, the Open Data Directive, the GDPR[7] and the Cybersecurity Act. (European Union, 2018, 2019). The recently adopted EU Data Strategy[8] takes on an interdisciplinary approach to regulation of the data economy. The strategy is chiefly rooted in the need to expand the responsible use, demand and development of digital products and services within the European Single Market for the period 2020 to 2025 and is backed by the intention to make the EU a leader in a data-driven society. Therefore, by creating a single market for data, this will allow it to flow freely within the EU and across sectors for the benefit of businesses, researchers and public administrations.

As has been stated before, the GDPR principally applies to the processing of personal data.[9] This extends to both an identified person as well as an identifiable natural person.[10] If this is applied practically, it therefore means that the GDPR and by extension, data protection does not apply to anonymous information or information which does not relate to an identified or identifiable natural person. The same can be said for personal data which has been so diluted or encrypted that it is rendered anonymous because the data subject is no longer identifiable. It is on this background that the EU adopted the Regulation on a framework for the free flow of

[7] The General Data Protection Regulation Supra note 11.

[8] https://op.europa.eu/en/publication-detail/-/publication/ac9cd214-53c6-11ea-aece-01aa75ed71a1/language-en

[9] Article 2.1. GDPR Supra note 11.

[10] "[that is] one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" (Article 4.1 of the GDPR) including quasi-identifiers and metadata (Article 4.1).

non-personal data in the EU,[11] also known as the FFD Regulation. The regulation makes it clear that it "applies to the processing of electronic data other than personal."[12]

The formulation of this regulation was rooted in the realisation that the expanding Internet of Things (IoT), artificial intelligence and machine learning, which are major sources of non-personal data, continuously presented legal problems for legislators and the courts alike because there was no precedence on how to deal with such data. In a competitive environment where practices such as data analytics may establish a competitive advantage, non-personal data such as real-time traffic avoidance navigation, has the potential to save corporations up to 730 million hours in transit time and up to €20 billion in labour costs among many other examples.[13]

Having realised the value and utility of non-personal data, the FFD regulation seeks to ensure four main objectives.

- The Free movement of non-personal data across borders within the EU. In the same breath, it seeks to ensure that any interested organisation, which has the capacity and means to do so, should be able to store and process data anywhere in the EU.
- The availability of data for regulatory control. In this sense, it aims to ensure that public authorities retain access to data, even when it is located in another EU country, or when it is stored or processed in the cloud.
- The ability to effectively and easily switch between cloud service providers for professional users. The Commission has started facilitating self-regulation in this area, encouraging providers to develop codes of conduct regarding the conditions under which users can move data between cloud service providers and back into their own IT environments.
- Full consistency and synergies with the cybersecurity package, and clarification that any security requirements that already apply to businesses storing and processing data will continue to do so when they store or process data across borders in the EU or in the cloud.

[11] The FFD Regulation supra note 23.

[12] Article 2.1 of the FFD Regulation supra note 23.

[13] https://medium.datadriveninvestor.com/digital-europe-200-billion-investment-strategies-for-artificial-intelligence-data-and-blockchain-f7f656e66603

The GDPR already provides for the free movement of personal data within the EU[14] subject to compliance with/the provision of certain guarantees.[15] In this way, an amalgamation of all laws connected to data regulation in the EU ensures that there is a comprehensive and coherent approach to the free movement of all data in the EU.

5.4.1 Lessons from the EU Approach

The key lessons to be taken from the European approach are that firstly, the EU has realised the fact that various kinds of data exist, and it is not only personal data that is of value or worthy of protection/regulation. Secondly, that data is not valuable when it is stagnant. Rather, any value to be derived from data only emerges when data is allowed to flow/move freely across the EU. Thirdly, the EU's approach to data regulation takes the form of a multi-faceted approach. In addition to having adopted a data strategy which serves as a guide against which all legislative efforts must aspire to effectuate, there was the realisation that to achieve the goals of the strategy, different legislation would have to be enacted albeit with the same goal in mind. Because there are so many aspects to regulating data and technology is constantly evolving, not only does the multi-faceted approach give more legal clarity and certainty on different regulatory issues, but it also makes it easier to amend and develop distinct parts of the law without disrupting ancillary legislative pieces. The effectiveness of this strategy, although still unfolding, has thus far yielded positive results and admiration from the rest of the world as evidenced by how many countries have "borrowed" various provisions from the different EU laws for implementation within their domestic and regional regulatory efforts.

[14] Article 1(3) of the GDPR states that the overarching goal of the GDPR is to ensure that" the free movement of personal data within the Union shall be neither restricted nor prohibited for reasons connected with the protection of natural persons with regard to the processing of personal data."

[15] See Chap. 2 (Art 5 to 11) of the GDPR.

5.5 African Regulatory Approaches to the Liberalisation of Data and Its Movement

5.5.1 The Free Movement of Data

The advent of the internet presented the world with a means to send copious amounts of data to almost any part of the world with minimal cost. In fact, it can be said that the costs of sending data across borders costs no more than sending data within the same borders. The COVID-19 pandemic has highlighted just how important data flows are important for the global economy. Data flows have been shown to have influence in areas such as healthcare (contact tracing/medical research/vaccine production), business/ecommerce (online shopping, streaming services, virtual meetings/conferences) as well as socially (family video calls, online concerts). The extent to which these fields have been influenced (positively) is an indicator that in future, data flows will only continue to rise as more countries and sectors embrace digital transformation. It has already been determined that between 2007 and 2017, global data flows multiplied more than 20-fold and was anticipated that by 2022, the situation as of 2017 would have quadrupled (World Bank, 2021).

Within the African context, international and regional frameworks that facilitate cross-border data flows will be essential for the facilitation of a common market and by extension, the realisation of the continental developmental goals such as the realisation of the African Free Continental Free Trade Area (AfCFTA) (World Bank, 2020) and the African Union's Agenda 2063 (Africa Union, 2015). While some countries allow data to freely flow in and out of their borders, many others have enacted legislative frameworks that speak to the protection of personal data and which contain, in most instances, data localisation clauses. Generally, data localisation laws require that data (mostly personal) about a nation's citizens or residents be collected, processed and stored within the borders of the country (Bowman, 2017). Where a request is made that such data is transferred internationally, several approvals and a lot of bureaucracy must be observed. Data localisation laws are often necessitated by concerns relating to data security. Such laws aim to ensure, through surveillance and other supervisory methods, that where data must be exchanged, that such data is lawfully obtained (through freely given consent), that the data is being used/exchanged for a specific purpose and that the data is not being used for unauthorised activity such as profiling or surveillance by governments or any other third parties without consent (unless otherwise required under the law) (World Bank, 2019). While it is understood that it is essential for digital transactions to be supported by formidable regulatory frameworks in privacy, security and consumer protection; such frameworks can impede the cross-border transfer and use of data by imposing substantial effort and costs on businesses, especially micro, small and medium enterprises (MSMEs), thereby deterring international exchanges (World Bank, 2021).

In today's digital and physical economies, the freedom to move data of both a personal and non-personal nature without restriction between countries generates positive outcomes for organisations, individuals and countries.

5.5.2 *Benefits of Cross-Data Flows*

Benefits of cross-data flows are discussed with a focus on individuals, countries and organisations.

5.5.3 *Benefits for Individuals*

For individuals, the reach and influence of the internet has already enabled their seamless interaction with people and organisations from all across the world. In the same breath, individuals have also been exposed to goods and services from foreign markets that are available online and may be delivered in short periods of time where such products are physical (GSMA, 2018). As has been mentioned before, the practice of data analytics has enabled organisations to cater for more geographic markets, giving those customers access to a wider range of goods and services based on their interests, wants and needs which further improves competition in the markets and overall customer satisfaction. Additionally, cross-data flows also enable individuals to carry out remote work from wherever they are in the world. The surge in remote work has come to be known as the "human cloud". The human cloud is defined as a budding set of online or digital marketplaces for labour where competent professionals and those looking to hire professionals can locate and engage one another in employment/work arrangements (Staffing Industry, 2017). By the end of 2018, it was estimated that the money spent on using the human cloud spend was estimated to generate around $82 billion globally, a figure that was expected to grow exponentially. The facilitation of Cross border flows will not only allow the host country the opportunity to export talent, but also the chance to reduce unemployment rates and generate foreign currency.

5.5.4 *Benefits to the Country*

Free cross-border flows have enabled more national businesses and consumers to enter the digital commerce sphere, thereby encouraging the endorsement of data-driven business strategies and stimulating the national economy (GSMA, 2021). Public-sector bodies and government departments also benefit from cross-border data flows allowing them to deliver better quality public services at a lower cost and pursue public policy objectives that might not otherwise be achievable.

5.5.5 *Benefits to Organisations*

The free movement of personal data delivers social and economic benefits much faster than the alternative, which would require businesses to actively construct their back-offices and to streamline their processes and storage functions to serve multiple individual markets.[16] Countries that adopt regulatory regimes that support the free international transfer of data allow small, specialised organisations to establish an internet presence that is simultaneously national and international.[17] In this way, it is possible to have services successfully adopted in one national market, then expanded to other markets, bringing rapid benefits for second and subsequent countries.[18]

A key advantage of the internet is that it allows any organisation, no matter how small, to use the internet to market and deliver its ideas, goods and services, wherever data is allowed to flow. In this sense, if there are restrictions on the movement of data, organisations would not be able to provide information and products in response to individuals' requests.[19] Multinational organisations are also able to become more efficient by centralising and virtualising their internal operations. Examples of improved efficiency include the cost-effective expansion of business by utilising

[16] Ibid.

[17] Ibid.

[18] Ibid.

[19] Ibid.

flexible, cloud-based infrastructure and specialist application service providers and minimising investment in additional IT equipment.[20]

The COVID-19 pandemic's unseen benefit amid all the negatives, is that more and more international businesses have seen the importance of adopting data-driven digital transformation strategies to secure their future. Such strategies tend to depend on being able to collect, analyse, process and store data across multi-country operations. The Practice of data analytics becomes even more amplified as organisations seek to generate new customer insights and the performance of their operations and products.[21]

5.6 Data Localisation Laws in Africa

This section gives an overview of some African states' data localisation laws. In lieu of the views expressed in Sect. 5.5.1 of this paper, that data localisation laws do not support the free flow of data; the enactment and implementation of these laws, for example through hefty fines for using remote data storage, has detrimental effects. However, each state is entitled to enact and implement its domestic laws. The contribution of this paper is that in legislative and policy making processes, African states ought to consider the impact of data localisation on data flows.

5.6.1 Cote-d'Ivoire

In 2013, The Ivory Coast/Cote-d'Ivoire enacted privacy laws which required firms to get pre-approval from the regulator before processing personal data outside of the Economic Community of West African States (ECOWAS).

5.6.2 Ghana

In 2019 Ghana enacted the Ghana Payment Systems Bill & Guidelines, which among other things, set out the requirements to obtain a payment systems operator license which pertain to local ownership and the appointment of Ghanaian directors. Prior to this, in July 2018, Ghana issued draft regulation that required all domestic transactions to be processed by the

[20] Ibid.

[21] Ibid.

Ghana Interbank Payment and Settlement Systems Limited (GhiPPS, which is wholly owned by the Central Bank of Ghana).

5.6.3 Kenya

Kenya's 2019 Data Protection Act does not contain the explicit data localisation provisions which appeared in earlier drafts of the law. However, it still includes restrictive provisions governing personal data which require explicit consent for transfers of "sensitive personal data"[22] and that data controllers ensure and provide proof that personal data transferred abroad receives the same protection as if stored within the borders of Kenya.[23] Regulations implementing these provisions are still being developed.

Proposed Measures (2021): Following the enactment of the 2019 Data Protection Act, Kenya has released three draft data protection regulations to aid in the implementation of the Data Protection Act.[24] These are the Data Protection (General) Regulations (ODPC, 2021), the Data Protection (Registration of Data Controllers and Data Processors) Regulations and the Data Protection (Compliance and Enforcement Regulations). Under the proposed measures, the General regulation requires that where data processing is done for the purpose of producing a public good, the processing should be carried out through a server and data centre located within Kenya's borders[25] and that at least one serving copy of the personal data should be stored in a data centre located in Kenya.[26] The regulations also include provisions on the cross-border transfer of personal data. Under the General Regulations, it is required that before transferring personal data outside of Kenya, the recipient ought to know they are bound by legally enforceable obligations to ensure the same level of protection to the transferred personal data as that provided for under the Data Protection Act in Kenya and the General Regulations[27]; that the data subject is informed of the safeguards and the implications and risks involved in the cross-border transfer[28]; that the data subject has

[22] Section 44 of the Act read with Section 25.

[23] Section 25(h) of the Act.

[24] Supra note 56.

[25] Regulation 25(1)(a) of the Data Protection General Regulations.

[26] Regulation 25(1)(b) of the Data Protection General Regulations.

[27] Regulation 38(1)(a) of the Data Protection General Regulations.

[28] Regulation 38(2) of the Data Protection General Regulations.

consented to the transfer of their data to that recipient[29]; that the transferring entity has taken reasonable steps to ensure that transferred personal data is not used for any unintended purposes[30]; and that the data subject's rights are safeguarded.[31] The General regulations also provide that cross-border transfers of data may be allowed without restrictions where the transfer is "necessary" as provided under Section 48(c) of the Data Protection Act[32]; where the requirements arbitrarily or unjustifiably discriminate against any person[33]; where the requirements impose a restriction on trade[34]; and where the restrictions on transfers of personal data are greater than are required to achieve the objectives of the Data Protection Act.[35] The General Regulations also prescribes the terms that are to be contained in cross border transfer agreements between transferring entities and the recipients of personal data albeit without prescribing the template model standard clauses as is seen in the European Union.[36]

5.6.4 Nigeria

In 2015, Nigeria enacted broad data localisation requirements as part of the Guidelines for the Nigerian Content Development in ICT (NITDA, 2019). In the guidelines, it is required that all telecommunication companies interested in hosting subscriber and consumer data within Nigeria, should host such data within the country and in line with existing legislation.[37] The same applies to Networking Service Companies[38] and Data and Information Management Companies.[39]

In 2011, The Central Bank of Nigeria also introduced a local storage and processing requirement for entities engaging in point of sale (POS) card services (Central Bank of Nigeria, 2011). Under guideline 4.4.8, All domestic transactions including but not limited to POS and ATM

[29] Regulation 38(1)(b) of the Data Protection General regulations.

[30] Regulation 38(1)(c) of the Data Protection General regulations.

[31] Section 38(1)(d) of the Data Protection General regulations.

[32] Regulation 40(1)(a) of the of the Data Protection General regulations.

[33] Regulation 40(1)(b) of the Data Protection General regulations.

[34] Regulation 40(1)(c) of the Data Protection General regulations.

[35] Regulation 40(1)(d) of the Data Protection General regulations.

[36] Regulation 39 of the Data Protection General regulations

[37] Section 11.1.3 of the Guidelines for the Nigerian Content Development in ICT.

[38] Section 12.1.4 of the Guidelines for the Nigerian Content Development in ICT.

[39] Section 13.1.2 and 13.2.3 of the Guidelines for the Nigerian Content Development in ICT.

transactions in Nigeria must be switched using the services of a local switch and shall not under any circumstance be routed outside Nigeria for switching between Nigerian Issuers and Acquirers.[40]

5.6.5 Rwanda

In 2012, Rwanda enacted a regulation that all critical information data within government should be hosted in their national data centre (MINICT, 2012). In terms of indirect application of data localisation laws, in 2017 Rwanda's telecommunications regulator fined MTN the sum of US$8.5 million for maintaining Rwandan customer data in Uganda and for running its IT services outside the country in breach of its license (CNBC Africa, 2017). Comments have already been made in the introductory Sect. 5.5.3 on the enactment and implementation of data localisation laws. Further, it could be argued that the imposition of such large fines may chill investment by firms that wish to use remote data storage facilities.

5.6.6 Senegal

In 2021 and in the light of the new Government data centre being built in Senegal, President Macky Sall announced that all government data and applications will be hosted at the centre and the repatriated from foreign servers in hopes of strengthening Senegal's digital sovereignty (Swinhoe, 2021).

5.6.7 South Africa

In 2018, following the realisation that domestic South African banks intended to move more of their transactions to global payment service networks, the South African Reserve Bank suspended the migration of all domestic transaction volumes from Bankserv (South Africa's bank-owned domestic payment switch) to international payment schemes (Cory & Dascoli, 2021). The suspension was to remain in place until a new policy was developed and enacted. Such a policy has not yet been developed and enacted at the time of writing.

In 2013, South Africa enacted the Protection of Personal Information Act (the POPI Act),[41] but which only came into full force on the 1st of

[40] Guideline 4.4.8 of The Central Bank of Nigeria's mandatory 2011 Guidelines on Point of Sale (POS) Card Acceptance Services.

[41] The Protection of Personal Information Act No 4 of 2013 available at https://popia.co.za/

July 2021, makes the transfer of personal information outside of South Africa subject to certain exceptions. These include the requirement that the recipient of the data be able to offer complimentary protection of the data,[42] that the data subject consents to the data transfer,[43] that the transfer is necessary for the performance of a contract between the data subject and the responsible party[44] or for the conclusion/performance of a contract in the interest of the data subject[45] and if the transfer is for the benefit of the data subject.[46] While these are not explicit localisation laws, there is concern as to how they will be interpreted and enforced, as they could become *de facto* data localisation tools.

Proposed Measures

More recently, South Africa's "Draft National Policy on Data and Cloud" of 2021 recommends the adoption of data localisation standards and local data processing for all data incidental to "critical information infrastructure"[47] and data mirroring for personal data.[48] It also states that all data generated in South Africa shall be the property of South Africa, regardless of the nationality of the firm involved in collecting it.[49]

5.6.8 Egypt

In Egypt, President Abdel Fattah el-Sisi ratified the Personal Data Protection Law[50] on the 13th of July 2020. The law aims to protect and regulate the collection and processing of personal data of Egypt's citizens and residents. In relation to data localisation, the law prohibits the transfer or retention of personal data to a foreign country or territory without the permission of the Egyptian Data Protection Centre and unless that country or territory has adequate levels of personal data protection.[51] Egyptian Minister of

[42] Section 72(1)(a) of the POPI Act.

[43] Section 72(1)(b) of the POPI Act.

[44] Section 72(1)(c) of the POPI Act.

[45] Section 72(1)(d) of the POPI Act.

[46] Section 72(1)(e) of the POPI Act.

[47] "National critical information infrastructure" as defined in section 9 of the Policy means all ICT systems, data systems, databases, networks (including people, buildings, facilities and processes), that are fundamental to the effective operation of the Republic of South Africa.

[48] Policy Intervention 10.4.1.

[49] Policy Intervention 10.4.4. of the Draft National Policy on Data and Cloud.

[50] The Personal Data Protection Law (Law No. 151 of 2020).

[51] Article 14 of the Data Protection Law of 2020.

Communications and Information Technology, Amr Talaat, was also quoted stating that the data protection law was formulated in support of the Ministry's efforts to localise the data centre industry and create a safe environment for the circulation of information within the cyberspace (Data Centre Planet, 2020). Egypt also belongs to the Arab Maghreb Union who have so far not attempted to regulate data collectively as a union.

5.6.9 Angola

The Data Protection Law[52] draws inspiration from provisions found in the EU and Portuguese legal regimes for the protection of personal data. The enforcement authority, known as the Agência de Proteção de Dados (APD), was only created in October 2019 despite the law being created in 2011; and there is presently no significant level of enforcement. The law requires that the APD be notified prior to any international transfers of personal data to countries deemed to have an adequate level of protection[53] in addition to specific requirements that must be met such as consent of the data subject.[54] Angola also belongs to the Economic Community of Central African States (ECCAS) which, in 2016, adopted a model law (with the support of ITU and EU). However, because ECCAS does not have binding community law instruments, only three member states out of ten have adopted a national privacy law (Le Bihan, 2018).

5.7 Open Data Policies/Standards

Open data policies are a new phenomenon in Africa, with a track of less than ten years of implementation. The drive for the open data is largely driven by civil societies to enhance citizenry engagement with government's service delivery (Mutuku & Tinto, 2019). Most of the open data in Africa are anchored by the government's information system. There are more than 20 countries, regional and international organisations that initiate open data drives specifically for Africa. In SSA, governments have adopted the Open Government Partnership (OGP) and the implementation African Peer Review mechanism by AU members, which has strengthened the creation of open data initiatives.

[52] The Data Protection Law (Law 22/11).

[53] Section VI of the Data Protection law.

[54] Article 34 of the Data Protection Law.

Open standards/policies for data can also be particularly useful tools that make it easier for individuals and organisations to access, use, publish and share better quality data while simultaneously addressing cybersecurity concerns. Open standards for data are reusable agreements that necessitate the access, use, publication and the sharing of better-quality data (Open Data Institute, 2018). Open data standards can also be defined as sets of specifications or requirements for how specific sets of data should be made publicly available (Data Standards, 2017).

They are particularly helpful because:

1. They increase interoperability: Data interoperability is a feature of datasets where data can be easily retrieved, processed, re-used and re-packaged ("operated") by other systems with little to no effort.[55]
2. They improve comparability of data: Because open data standards enable easy access to datasets, they make it easier to compare data from different sources and to draw more concrete conclusions by drawing from a pool of like data sets.
3. They enable aggregation: By lowering the barriers to access to data, open standards for data encourage the publication of new data and better-quality data that is structured in a similar way, making it easier to combine them. In the process, the cost and complexity of combining similar data from multiple sources is significantly decreased (Open Data Institute, 2018).
4. They Enable linkability: Open standards make it easy to combine diverse data sets to give useful insights.

5.7.1 Common Uses of Open Standards for Data

As has been stated, open standards are essential in aiding the creation of a strong data ecosystem. Within this ecosystem, there are data assets,[56] the organisations responsible for the operation and maintenance of the data

[55] https://aims.gitbook.io/open-data-mooc/unit-4-sharing-open-data/lesson-4.2-introduction-to-data-interoperability

[56] The term data asset is used to refer to data that is expected to generate future revenues. This differs from one industry to the other and ultimately what is considered a data asset depends on the relevant business model. Examples include Design and methodology, knowledge/know-how, user input, sensor data, calculated data etc. See 7 examples of a data asset available at https://simplicable.com/new/data-asset

assets, and guides that set out how to use, store and manage the data.[57] A strong data infrastructure is critical to fostering business innovation, driving better public services and creating healthy, sustainable communities.[58]

To Promote Common Understanding
Many open standards exist today for different purposes and in different sectors. The commonality across all successful open standards is that they focus on tackling specific issues with reusable agreements that support better quality data. Therefore, where there is a need for people and organisations to agree on common guidance, a shared language or common models when solving problems, open standards are ideal.[59]

To Support Policy and Legislation
When implementing policies and substantiating legislation adopted or developed by governments and other public bodies, open standards for data can be useful support tools. By establishing standards on how to disclose data, how to automate compliance checks, how to aggregate or report on data and in the process, this can produce better quality data and strengthen a data infrastructure.[60]

To Fill Gaps in a Data Infrastructure
A strong data infrastructure[61] is grounded on principles that promote accountability, transparency, business innovation, civil society and public services. Within the infrastructure are data assets, the organisations that operate and maintain them, and the regulations that describe how to use and manage the data.[62] It is therefore important that a strong data infrastructure is supported by open data standards. The identification of gaps is made easier by lessening the barriers to entry in data pools as well as the participation of more stakeholders.

5.7.2 Benefits of Open Data Standards

The benefits of open data standards can be summarised by the image below (Fig. 5.1).

[57] "When to use open standards for data" Supra note 99.
[58] Ibid.
[59] Ibid.
[60] Ibid.
[61] Data infrastructure consists of data assets supported by people, processes and technology.
[62] When to use open standards for data" Supra note 99.

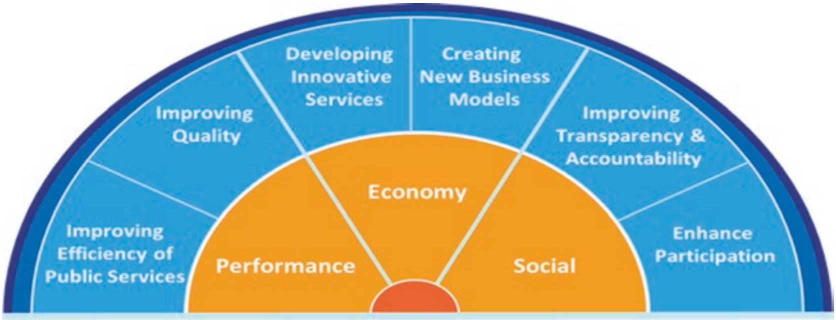

Developing Innovative Services
Creating New Business Models
Improving Quality
Improving Transparency & Accountability
Economy
Improving Efficiency of Public Services
Performance
Social
Enhance Participation

Fig. 5.1 Benefits of Open Data Standards. Image from data europa.eu (European Union, 2017)

Economic Benefits

The economic benefits of Open Data Standards are of greater importance to this discussion. The crux of the benefits presented by open data standards are that standards create new commercial opportunities and ecosystems that encourage competition. First, standards help to deconcentrate authority. Well established market leaders and authorities are discouraged from using custom and proprietary formats and opt instead to make use of cooperatively produced and shared standards (Open Data Institute, 2018). This effectively levels the playing field for data production and data use, allowing new uses of data and new entries to the market.[63]

Therefore, by effectively reducing barriers to entry and the costs associated with the collection and aggregation of data in a particular sector, standards also allow more organisations to enter the ecosystem to provide more diverse products and services within the data ecosystem.[64] Examples include translation, conversion, combination, reporting, training, analytics, consumer products, business-to-business services and more. Open standards for data mean that an organisation can focus on providing value at any stage of the data pipeline.

Social Benefits

Open data standards encourage multi-stakeholder collaboration. Essentially, developing a standard that is useful to the community and used

[63] Ibid.

[64] Ibid

by stakeholders needs multi-stakeholder collaboration. Multi-stakeholder collaboration connects people and organisations working within a sector. Data publishers are interested in who else publishes data using standards so they can understand how issues were overcome and improve their processes. Data users are interested in connecting with other data users with similar goals or issues. In the process a focus for shared vision may be developed (Open Data Institute, 2018). When people and organisations with a common problem or an unmet need work together to reach an agreement about producing or using better-quality data, the people and organisations involved need a shared vision of the open standard including a common understanding of the problem they are trying to solve and agreement on how they will solve it.[65]

In the process, an open standard for data can aid in coordinating activities to understand the problem or unmet need; agreeing on the current ecosystem, data assets, concepts and language in use; agreeing on the data and models needed to solve the problem or meet the need; pooling resources to work towards clearly defined goals for the standard, leading to mutually reinforcing activities; forming connections across sectors to support the standard's goals, which can help to build trust, peer learning and peer support; and producing and reusing tools that strengthen a data infrastructure, including supporting data publishers, providing data users with insight and making it easier for developers to create tools and services.[66]

Policy Impacts

From a policy perspective, open standards can support implementation of policy. In the past, policymakers requiring organisations to publish data have focused on what data must be published but not on how. This leads to situations where disclosure is widespread but the data is difficult to collate and use. By adopting open standards for data and linking them to policy and regulation, policymakers can make data more usable, provide clear guidance on how to disclose data, automate compliance checks, data aggregation and reporting open standards for data provide clarity to data publishers, the opportunity for stakeholder engagement and help ensure consistent and comparable results (Ibid).

[65] Ibid.

[66] Ibid.

Technological Benefits

The key technological benefits of open data are that standards produce better quality data (McGilvray, 2008). Open standards encourage the development of tools and services to help data publishers produce good quality data, including tools to validate, preview and compare data (Open Data Institute, 2018).

Open standards can advise data publishers how often data should be published. Some standards include ways to share publication schedules, publication dates, location and methods of accessing data. Sharing this information makes it easier to trust published data.[67]

In addition, when data is published consistently, the time, cost and processes involved in using it are reduced. Consistent publication encourages the creation of new tools and services that are designed to take advantage of data that conforms to a standard.[68]

Example of Open Data Standards in Use

Probably the most famous open data standard is the General Transit Feed Specification (GTFS) which is a standard which was developed by tech giant Google. The GTFS allows public transit agencies to publish their transit data in a format that can be interpreted and used by a variety of software applications.[69] Because of the interoperability of open data standards, GTFS data can be used by many other third-party software applications for a variety of purposes. Examples include trip planning, timetable creation, mobile data, data visualisation, accessibility, analysis tools for planning and real-time information systems.[70] Among public transportation data formats, GTFS stands out because it was conceived to meet specific, practical needs in communicating service information to passengers. It is designed to be relatively simple to create and read for both people and machines.[71] The value of an efficient transport system has real implications on the economy of a country, but because it is so easy to access and share data in this manner, efficiency is amplified even beyond borders.

[67] Ibid.

[68] Ibid.

[69] GTFS: Making Public Transit Data Universally Accessible available at https://gtfs.org/ accessed 10/10/2021.

[70] GTFS Background available at https://gtfs.org/gtfs-background accessed on 10/10/2021.

[71] Ibid.

5.8 Cybersecurity Concerns

The movement of data entails considerable security risks, hence the need for cybersecurity and the protection of both personal and non-personal data. As indicated above, the Malabo Convention is the only current continental legal instrument that focuses on the protection of personal data and cybersecurity. It is relevant to data governance to the extent that it pertains to these two aspects, which are integral to data governance. Indeed, as noted above, the AUC is shepherding the development and formulation of the Africa Data Policy Framework, which is informed, in part, by the Malabo Convention. A well-crafted data governance framework ought to include both aspects because "security and privacy have become one of the crucial concerns related to data storage and usage within organizations" (Yang, 2019). Leading up to the adoption of the Malabo Convention in 2014, several RECs adopted regulatory instruments on privacy and cybersecurity (Ncube, 2016). These are ECOWAS' Supplementary Act on Personal Data Protection within ECOWAS (2010); the ECOWAS Directive on Fighting Cybercrime (2011); the Common Market for Eastern and Southern Africa (COMESA)'s Model Cybercrime Bill (2011); the Southern African Development Community (SADC)'s Model Law on Data Protection and a Model Law on Computer Crime and Cybercrime (2012). Of these, only the SADC Model covers both privacy and security, however, as it a non-binding instrument, and consequently the Malabo Convention stands out as the only binding instrument regulating both privacy and security. Further, according to its preamble, it "embodies the existing commitments of AU Member States at subregional, regional and international levels to build the Information Society", making it the continental blueprint. Accordingly, this section reprises the Malabo Convention's provisions on cybersecurity and privacy. As already noted above, this section is succinct, due to the coverage of the same content, in greater detail, by another paper which constitutes part of the project.

5.8.1 Privacy

As indicated above, the Malabo Convention focuses on the protection of personal data (privacy) rather than non-personal data. Its definition provision sets out the following fundamental definitions:

> Personal data means any information relating to an identified or identifiable natural person by which this person can be identified, directly or indirectly in particular by reference to an identification number or to one or more factors specific to his/her physical, physiological, mental, economic, cultural or social identity.
>
> Personal data file means all structured package of data accessible in accordance with set criteria, regardless of whether or not such data are centralized, decentralized or distributed functionally or geographically.
>
> Sensitive data means all personal data relating to religious, philosophical, political and trade-union opinions and activities, as well as to sex life or race, health, social measures, legal proceedings and penal or administrative sanctions.

It then turns to the regulation of the processing of personal data which is defined as

> any operation or set of operations which is performed upon personal data, whether or not by automatic means such as the collection, recording, organization, storage, adaptation, alteration, retrieval, backup, copy, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination and locking, encryption, erasure or destruction of personal data.

The core of such regulation consists of the basic principles governing the processing of personal data as set out in Article 13. These are:

Principle 1: Principle of consent and legitimacy of personal data processing.
Principle 2: Principle of lawfulness and fairness of personal data processing.
Principle 3: Principle of purpose, relevance and storage of processed personal data.
Principle 4: Principle of accuracy of personal data.
Principle 5: Principle of transparency of personal data processing.
Principle 6: Principle of confidentiality and security of personal data processing.

Their meaning is the same as of the GDPR's principles as set out at Sect. 5.2 above. They are supplemented by Article 14 which sets out specific principles for the processing of sensitive data. Another core component of privacy in the Malabo Convention is its Section IV on the Data Subjects' following rights: Right to information (Article 16); Right of access (Article

17); Right to object (Article 18) and Right of rectification or erasure (Article 19). Personal Data Controllers have the following obligations: Confidentiality obligations (Article 20); Security obligations (Article 21); Storage obligations (Article 22) and Sustainability obligations (Article 23).

5.8.2 Cybersecurity

The Malabo Convention does not contain a definition of cybersecurity, which would have been useful to underpin a significant aspect that it regulates. Yang et al. define cybersecurity as "the practice of protecting computer and network infrastructures, the operating systems, software programmes run on the infrastructures, and all the data stored or transmitted through the infrastructures from digital attacks and any other misuse" Chap. 3 of the Convention is intended to promote cybersecurity and prevent cybercrime. Article 24 addresses national cybersecurity frameworks, specifically national policies and strategies relating to the Critical Information Infrastructure (CII). The Malabo Convention defines CII as "the cyber infrastructure that is essential to vital services for public safety, economic stability, national security, international stability and for the sustainability and restoration of critical cyberspace."

Article 25 then proceeds to address legal measures, namely (1) cybercrime national legislation; (2) regulatory authorities; (3) citizens' rights and (4) protection of critical infrastructure. According to Article 25.1 national cybercrime legislation is required to effectively sanction "criminal offences acts which affect the confidentiality, integrity, availability and survival of information and communication technology systems, the data they process and the underlying network infrastructure". Data is expressly mentioned here, so it is clear that some national cybersecurity measures in relation to data is mandated. Further, Article 25.2 requires "effective procedural measures to pursue and prosecute offenders." Article 26 then proceeds to require state parties to establish a national cybersecurity system comprising of the necessary institutions, which are appropriately staffed, to oversee implementation of the legal measures through actions including responding to cybersecurity incidents, and coordination and cooperation in forensic investigations and prosecution, amongst others. Such legal measures and their implementation must have due regard to the human rights of citizens.[72] State parties are also required to establish

[72] Article 25.3.

legislative or regulatory measures to protect priority sectors that are important for national security by, for instance, introducing more severe sanctions for offences in these sectors.[73]

Article 26 provides for some further detail regarding the national cybersecurity system through mandating each state "to promote the culture of cyber security" and suggests measures which may include cybersecurity plans and awareness campaigns. Article 27 proceeds to deal with national cybersecurity monitoring structures which state parties are required to adopt for cybersecurity governance within a national framework. Article 28 provides for international cooperation through harmonisation, encouraging states to offer each other mutual legal assistance and the exchange of information, along with the use of existing means for international cooperation. Article 29 then provides for offences that are specific to ICTs. It requires state parties to create offences relating to attacks on computer systems for instance to gain unauthorised access and data breaches such as the interception or attempted interception of computerised data. There are also provisions relating to content related offences in Article 29 and the adaptation of property offences and sanctions to ICTs in Articles 30–31, but these are not pertinent to the chapter's area of focus.

The Malabo Convention's provisions provide a baseline, but more is needed for a robust approach to privacy and security for non-personal data because its privacy provisions are primarily for personal data and its cybersecurity provisions place emphasis on national infrastructure or the CII.

In summary, this section shows that most African states need to create, enhance or strengthen their privacy and cybersecurity frameworks. In view of the aims of both the AfCFTA and the Digital Transformation Strategy for Africa, to facilitate and grow e-commerce and digital trade in Africa, it will be important to align domestic frameworks. This gives a measure of certainty for entrepreneurs trading in multiple jurisdictions. As also indicated above, and reinforced below, the negotiations of the AfCFTA e-commerce protocol will provide a platform to agree on fundamental data governance principles.

[73] Article 25.4.

5.9 Conclusion

What this paper has managed to expose in part is the fact that one of two things are happening on the continent. On the one hand, concerted continental efforts may be unrolling sluggishly while the data revolution is unfolding at a much faster rate. Because this is the case, progressive nations, in a bid to compete within the data economy, have elected to attempt data governance on their own, thereby proffering the present situation of discordant and possibly conflicting data regulation laws. While on the other hand, what we may be witnessing is a lack of trust and confidence amongst African states in unified regulatory efforts. In some instances, because the data that is of the highest value is personal, such a lack of trust may be coupled with paranoia and suspicion by mostly individuals. It will therefore be imperative that a trusted data environment grounded in the rule of law; comprehensive institutional arrangements and regulations; and competent institutions responsible for overseeing the use of public and private data is established as soon as possible.

Such an environment can be created through multistakeholder efforts to improve data access and use. This may mean active dialogue between governments, consultations and collaborations with the private sector, and the establishment of Data Protection Authorities (DPA's) competent in the investigation and prosecution of cross border breaches. On top of the inter-governmental dialogue agenda should be the negotiation of mutual assistance agreements that will guarantee similar protection of data in contracting member states and pledges to investigate and prosecute cross-border cybercrimes comprehensively.[74] This will go a long way in moderating the concerns related to the free movement of data. Also, because the majority of African states are still in a developmental state, with some more advanced than others, capacity-building in relation to data protection, cybersecurity and institutional data governance in relevant agencies should be prioritised and realised through policy and asset allocation. In addition, where institutional arrangements and regulations come about as a result of the consultations and dialogue, these arrangements ought to be established through inclusive, consultative and transparent processes. Accountability and transparency are the answer to most of the concerns that follow the shift to data liberalisation and use.

[74] Chapter 7 of the GDPR.

As argued at Sect. 5.2, it is important to highlight that personal and non-personal data should not be treated the same, hence distinct approaches exist in other parts of the world. Whilst the concerns around the protection and regulation of personal data are legitimate, non-personal data which has a lot of value within itself, should not be subject to the same scrutiny. In this regard, lessons can be drawn from the approach that the EU has taken in ensuring that the two are distinct (see key lessons outlined at Sect. 2.1).

The current position, as summarised at Sect. 5.3, confirms that most African countries' attempts at regulating data have overly pre-occupied themselves with personal data, neglecting non-personal data. In the same breath, because personal data is of higher value, it is no surprise that protection laws in this regard may be overbearing. While the current forms of data localisation laws may be thought of as being national governments' attempts to assert sovereignty over data, a borderless medium, the reality is that as more countries enact updated data protection frameworks, it is highly likely that some policymakers will propose more stringent data localisation laws as they believe that the best way to protect data is to store it within a country's borders. However, evidence has shown that the security of data does not depend on where it is stored. Instead, by allowing for the free movement of data across international borders, cybersecurity concerns are less likely to materialise. By allowing cloud service providers to draw from data flows from all over, they will be able to establish best practices in cybersecurity. Similarly, while cloud computing does not guarantee security, it will lead to better security because implementing a robust security program requires resources and expertise, which many organisations and African countries lack. But large-scale cloud computing providers are better positioned to offer this protection. In fact, the security of data depends primarily on the logical and physical controls used to protect it, such as strong encryption on devices and perimeter security for data centres. The nationality of who owns or controls servers or which country these devices are located in, has little to do with how secure they are. Therefore, given the potential benefits that open cross-border flows would bring about, it would be prudent to start aligning policy with the promotion of open cross border data flows. Furthermore, because a comprehensive data regime also makes provision for data sovereignty, data specificity should also be prioritised. Data specificity is used to refer to countries being able to specify what kinds of data can and cannot move freely. Data specificity should be prioritised to avoid unintended restrictions on productive data sharing.

As the AfCFTA and the Digital Transformation Strategy for Africa (2020–2030) seek to increase e-commerce and digital trade in Africa, it is important to consider how supporting the free movement of data across Africa can enhance these efforts. It has been shown that cross border data flows are instrumental and have the potential to greatly influence a new economic resurgence for the continent, as can be drawn from experiences of countries or regional bodies that have adopted a liberal approach to data regulation. Their experience has evidenced that data localisation does not serve the purpose that many think that it does and in actual fact could be thought of as being counterproductive in terms of securing and drawing value from data. Most African countries that have enacted data localisation laws in one way or the other have done so under the justification that the security of data is dependent on where it is stored or collected, which is in fact a fallacy. It has been shown that open policies towards cross-border data flows have generated better security measures and better revenues for the countries that have adopted these systems and the African continent can learn from these experiences now to adequately support the free flow of data. It is also necessary to emphasise that the adoption of open standards for data which will complement the cross-border data flows ensuring that they are the flows are conducted in a safe and transparent manner and to ensure that barriers into accessing the data economy are reduced, thereby encouraging more players to get involved within the data economy. By adopting open data standards and decentralising the power to collect, use and aggregate data, participation in the data economy is encouraged and the chances of illegitimate uses of data are lessened. In the process, Governments are also afforded the opportunity to work on and strengthen their impact in key areas such as policy, technology and development as well as the economy. Such an approach recognises the importance of cybersecurity and supports it within an ecosystem that encourages open data participation.

Ultimately, there is a need to adopt a cohesive legal approach that is unambiguous and offers protection and obligations across the continent while taking cognizance of the value that the liberalisation of data has. Going forward, existing legal instruments should be revisited regularly, where necessary, to eliminate conflicts in law and also to keep abreast with the latest levels of protection and obligations within member states.

References

Regulations, Laws, Policies and Guidelines

Angola – The Data Protection Law (Law 22/11).

EU – The General Data Protection Regulation 2016/679 (EU).

Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, OJ L 303.

Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information.

Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act).

Egypt – Personal Data Protection Law (Law No. 151 of 2020).

Kenya – Data Protection Act No 24 of 2019.

Kenya – Data Protection (Registration of Data Controllers and Data Processors) Regulations. (2021). Available at. https://www.odpc.go.ke/resources/data-protection-compliance-and-enforcement-regulations-2021/

Kenya – Data Protection (General) Regulations. (2021). Available at. https://www.odpc.go.ke/resources/data-protection-general-regulations-2021/

Kenya – Data Protection (Compliance and Enforcement Regulations). (2021). Available at. https://www.odpc.go.ke/resources/data-protection-registration-of-data-controllers-and-data-processors-regulations-2021/

Ivory Coast – Law No. 2013 450 dated June 19, 2013 on the protection of personal data available at. https://ictpolicyafrica.org/fr/document/4wo0y6uby6j

Nigeria – Guidelines for Nigerian Content Development in Information and Communications Technology (ICT) of 2015 and as amended in Aug 2019. Available at https://nitda.gov.ng/wp-content/uploads/2020/11/GNCFinale2211.pdf

Nigeria – The Central Bank of Nigeria's mandatory 2011 Guidelines on Point of Sale (POS) Card Acceptance Services. Available at https://www.cbn.gov.ng/cashless/POS_GUIDELINES_August2011_FINAL_FINAL%20(2).pdf

Rwanda – Ministerial order N°001/MINICT/2012. Available at https://www.rlrc.gov.rw/fileadmin/user_upload/LawsofRwanda/Laws%20of%20Rwanda/7._Administrative/5.9.%20State%20Finance/5.9.3.%20Procurement/5.9.3.3._M._Instructions_Procurement_of_ICT_goods_and_services_by_public_institutions.pdf

South Africa – Protection of Personal Information Act No 4 of 2013.

South Africa – Draft National Policy on Data and Cloud. Available at https://www.gov.za/sites/default/files/gcis_document/202104/44389gon206.pdf

Cases

Patrick Breyer v Bundesrepublik Deutschland Case C-582/12.

Secondary Sources

African Union. (2020). Decision on the African Continental Free Trade Area (AfCFTA). Assembly/AU/Dec.751(XXXIII). Available at https://www.tralac.org/documents/resources/cfta/3176-au-assembly-decision-on-the-afcfta-february-2020/file.html

African Union. (2022). https://au.int/en/documents/20220728/au-data-policy-framework

African Union. *Digital Transformation Strategy for Africa 2020–2030*. Available at https://au.int/en/documents/20200518/digital-transformation-strategy-africa-2020-2030 Accessed on 30/04/2021.

African Union. Agenda 2063 – Background Note. Available at. https://au.int/sites/default/files/documents/33126-doc-01_background_note.pdf Accessed on 30/04/2021.

Analysts, Staffing Industry. The human cloud, the gig economy & the transformation of work. (2017). Available at https://www2.staffingindustry.com/content/download/246507/9128496/HumanCloudSummary2017_170912.pdf Accessed on 30/04/2021.

Aryan, A. (27 July 2020). Explained: What is non-personal data? *The Indian Express*. Available at https://indianexpress.com/article/explained/non-personal-data-explained-6506613/ Accessed 17 January 2022.

Bird and Bird. (2019). *EU data economy: Data-related legal, ethical and social issues*. Available at https://www.twobirds.com/~/media/pdfs/eu-data-economy-legal-ethical%2D%2Dsocial-issues.pdf Accessed on 30/04/2021.

Bowman, C. (6 January 2017). Data localization laws: An emerging global trend. *Jurist*. Available at https://www.jurist.org/commentary/2017/01/courtney-bowman-data-localization/ Accessed on 30/04/2021.

Chaytor, B. (2020). *AfCFTA, an enabler of digital trade and e-commerce*. Available at https://resilient.digital-africa.co/en/blog/tech_voices/afcfta-an-enabler-of-digital-trade-and-e-commerce-1-4/ Accessed on 30/04/2021.

CNBC Africa ‘Rwanda Utilities Regulatory Authority fines MTN US$ 8,5 M‘17 MAY 2017. Available at https://www.cnbcafrica.com/2017/rwanda-utilities-regulatory-authority-fines-mtn-us-85m-non-compliance/ Accessed on 30/04/2021.

Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions: A European Strategy for Data. Available at HTTPS://OP.EUROPA.EU/EN/PUBLICATION-DETAIL/-/PUBLICATION/AC9CD214-53C6-11EA-AECE-01AA75ED71A1/LANGUAGE-EN accessed on 30/04/2021.

Cory, N., & Dascoli, L. (2021). How barriers to cross-border data flows are spreading globally, what they cost, and how to address them. *Information Technology and Innovation Foundation*. Available at https://itif.org/publications/2021/07/19/how-barriers-cross-border-data-flows-are-spreading-globally-what-they-cost

Daniel, J. (15 February 2021). Data protection laws in Africa: What you need to know. *CIO Africa*. Available at https://www.cio.com/article/3607734/data-protection-laws-in-africa-what-you-need-to-know.html?upd=1619734077195 Accessed on 29/04/2021.

Data Centre Planet 'Egypt New Data Law Supports Data Center Localization' (Daily News..., 5 August 2020). Available at https://www.datacenterplanet.com/data-center/egypt-new-data-law-supports-data-center-localization/ Accessed 11 November 2021.

Data.Europa.EU. *The economic benefits of open data*. Available at https://data.europa.eu/en/highlights/economic-benefits-open-data accessed 09/10/2021 Accessed on 27/08/2021.

Data Protection Africa | By ALT Advisory. (2023). https://dataprotection.africa/

Diffen 'Data vs Information'. Available at https://www.diffen.com/difference/Data_vs_Information Accessed on 28/04/2021.

GSMA. *Cross-Border Data Flows: Realising benefits and removing barriers*. Available at https://www.gsma.com/publicpolicy/resources/cross-border-data-flows-realising-benefits-and-removing-barriers Accessed on 30/05/2021.

GTFS: Making Public Transit Data Universally. Available at https://gtfs.org/ Accessed 10/10/2021.

Hutt, R. (2015). *What are your digital rights?*

Interpol. (2021). *African Cyberthreat Assessment Report: Interpol's Key Insight into cybercrime in Africa*. AfriPol.

Le Bihan, J.-F. (16 July 2018). Regional regulatory capacity building. *GSMA Africa Policy Day*. Available at https://www.gsma.com/publicpolicy/wp-content/uploads/2018/07/M360-Africa-Policy-Day-Presentation.pdf Accessed on 11/11/2021.

Manzo, V. (2019). The internet of things and intellectual property rights: The protection of data. *WIPO Academy, University of Turin and ITC-ILO – Master of Laws in IP – Research Papers Collection – 2017–2018* Available at SSRN: https://ssrn.com/abstract=3387417.

Math is fun 'what is data?'. Available at https://www.mathsisfun.com/data/data.html Accessed on 28/04/2021.

McGilvray, D. (2008). *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information (TM)*.

Ncube, C. B. (2016). Recent developments in African regulation of cybercrime: An overview of proposed changes to the South African framework. *Journal of Internet Law, 19*(7), 3–20.

OECD. *Glossary of statistical terms.* Available at https://stats.oecd.org/glossary/detail.asp?ID=532 Accessed on 28/04/2021.

Open Data Institute. (2018). Social Impacts of open standards. *The open standards for data guidebook.* Available at https://standards.theodi.org/creating-impact/social-impacts/ Accessed on 28/08/2021.

Open data Institute. (2018a). Economic Impacts of open standards. *The open standards for data guidebook.* Available at https://standards.theodi.org/creating-impact/economic-impacts/ Accessed 27/08/2021.

Open data Institute. (2018b). Policy Impacts of open standards. *The open standards for data guidebook.* Available at https://standards.theodi.org/creating-impact/policy-impacts/ Accessed on 27/08/2021.

Yang, L., Li, J., Elisa, N., Prickettm, T., & Chao, F. (2019). Towards big data governance in cybersecurity. *Data-Enabled Discovery and Applications, 3*, 10. https://doi.org/10.1007/s41688-019-0034-9

CHAPTER 6

Digitalisation and Financial Data Governance in Africa: Challenges and Opportunities

Bitange Ndemo and Ben Mkalama

6.1 Introduction

The influence of digitalisation for future economic development continues to fascinate scholars and practitioners. Digitalisation often initially involves sizeable costs, and its benefits become apparent after the accompanying data is transferred to an environment that minimises operational expenses. To provide distinction—digitisation refers to the act of converting information from an analogue to a digital format, and digitalisation refers to the subsequent process of leveraging on the digital content to attain an enhanced user experience (Rijswijk et al., 2020). Digitalisation is considered as the process of integrating digital technologies into everyday life. Among other advantages, digitalisation allows users to spend less on computer hardware and software, thereby allowing team members to focus on other tasks. Additionally, digitalisation aids data collection in that

B. Ndemo (✉)
Kenya's Ambassador, Belgium and the EU Mission, Brussels, Belgium
e-mail: bndemo@bitangendemo.me

B. Mkalama
University of Nairobi, Nairobi, Kenya

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8_6

materials are optimised for analysis and investigation, thus allowing firms to prosper. Similarly, with digitalisation, data and resources can be consolidated into a set of tools that would offer a single perspective on their customer journey, operations and other business opportunities. As a result, data-driven consumer insights can facilitate a customer-centric corporate strategy, affording a better overall consumer experience in an era that rewards limitless choices, low prices and quick delivery. Businesses can, therefore, become more agile with digitalisation, increasing their speed to market while embracing innovative approaches (Backbase, 2021; International Finance Corporation, 2017).

Digitalisation increases efficiency and raises productivity through greater automation of manual activities, integration of data across the company and empowering of team members with a collaborative digital culture that offers tools that are adapted to their context. Digitalisation in the finance sphere provides convenience, efficiency and security to clients looking to access their funds; additionally, digitalisation discourages corruption, financing of illicit activities and tax evasion (Jafri, 2021), while promoting greater inclusion which enables access to formal financial services for those lacking other avenues of access (Jafri, 2021; Shipalana, 2019). Put simply, a financial technology company (FinTech) innovations encourage financial inclusion and bolster economic development.

6.1.1 The Spread of Digitalisation

New technologies and their attendant effects on digitalisation are fuelling the emerging Fourth Industrial Revolution (4IR). Technologies like big data analytics, artificial intelligence (AI), blockchain, Internet of Things (IoT) and robotics are being used, independently and jointly, to create new data-driven business models and to generate better opportunities for enterprises to expand rapidly. Technology companies have taken to digitalisation both to understand the customer (through data) and to use this data to design new products, to improve productivity and convenience. As part of its 2030 Agenda, the United Nations General Assembly approved the 17 Sustainable Development Goals (SDGs) as a blueprint for sustainable development in the world. Numerous experts have opined that realising the SDGs requires the effective exploitation of data across numerous sectors (Macmillan, 2020).

Technology firms that were not previously in the financial-services sector are carving out a niche for themselves (Frost et al., 2019); the number of FinTechs has, therefore, multiplied rapidly, rendering financial resources increasingly available and affordable. Increasingly active FinTechs have, however, spawned additional challenges varying across different sectors and jurisdictions, such as e-commerce platforms registering far-reaching cross-border impacts like tax liability for local firms across entire economies, as well as the emergence of regional payment platforms necessitating transnational and interdisciplinary regulatory discourses. There are also governance issues regarding cyber-resilience and data-management concerns; what's more, digital currencies—especially amid economic globalisation—continue to reconstitute financial and monetary policies, with potentially significant economic impact.

Mobile money platforms were built on the back of GSM connectivity and have grown phenomenally; they are now used by over one billion people around the world (Enberg, 2019). Over the past decade, mobile money services have rapidly expanded access to financial services in Africa (Kirui, 2020). The rate of growth of the FinTech industry in Africa has been influenced, too, by existing mobile network operators and their relationships with central banks (International Finance Corporation, 2017). Similarly, in Africa, the payments industry leads the FinTech sector, where historically, more than 90% of the economy has been cash-based (Yermack, 2018). Indeed, across the region, some countries have become global leaders in mobile-money transactions. For instance, while the rest of the world hovers at 5%, as a proportion of GDP, the average for such transactions in sub-Saharan Africa is closer to 25% (International Finance Corporation, 2017). Indeed, by 2017, mobile transactions in Kenya—a country that pioneered the use of mobile for financial transactions—executed through the Mpesa platform, accounted for 44% of national GDP (Rolfe, 2019). Similarly, Zimbabwe's EcoCash platform transacted more than $23 billion in 2017, or 54% of national GDP (Sengere, 2017). These profound rates of use occurred despite remittances in sub-Saharan Africa, registering just under 10% of the global value and yet had the highest transaction costs in the world, thereby creating opportunities for exploitation (Yermack, 2018).

The growth of mobile money in Kenya and Zimbabwe was spurred by an enabling regulatory environment, whereby this innovation was led by the dynamic private sector technological communication players rather than the financial system players (Chitimira & Torerai, 2021; Muthiora, 2015). This situation was due to an appreciation that mobile money combined both a telecommunication service and a financial service. Some of the factors that contributed to financial exclusion included "a trust deficit in banking institutions, their distance from financial institutions and their failure to satisfy documentary requirements such as a proof of identity and a proof of residence" (Chitimira & Torerai, 2021, p. 10). As a result of mobile banking, relaxed requirements to open a mobile money account enabled the poor and unbanked persons to participate in the mainstream financial system, thereby flourishing this sector. Similarly, the mobile money agents' network and the prepaid model played a critical to the aggressive growth and assurance of security for mobile money in Kenya and Zimbabwe (Chitimira & Torerai, 2021; Muthiora, 2015). Yermack (2018) opined that mobile money flourished in markets with weak institutional structures as a result of the following characteristics: low profit margins, asset being light in nature, scalable designs, innovativeness and compliance friendly.

Digitalisation has strong implications for financial access and inclusion in the Global South (Jafri, 2021), and FinTechs are poised to hasten these processes (Lewis et al., 2017). Discourses that favour the dematerialisation of money as "a technological fix to broader problems of poverty and financial exclusion" have seen enhanced digitalisation of the financial sector even as it is unclear whether those at the periphery of the digital money ecosystem genuinely benefit from the influence of digitalised operations (Muralidhar et al., 2019). Lack of identification documents had, previously, been viewed as an impediment to financial access, but this challenge has been addressed recently by providing individuals with a digital identity. Digital identities have become increasingly prominent, and a digital welfare state is developing, depicted by a gradual uptake of digital data and technologies in welfare schemes, partnerships, administrative processes and the provision of services (Jafri, 2021). Kenya, for example, now transacts social service programmes through electronic means like mobile wallets (Republic of Kenya, 2017), notwithstanding hindrances to fulfilment and success stemming from challenges such as literacy (including IT literacy), the socio-emotional perspective of marginalised individuals and the complexity of information and service requirements that users confront (Malladi et al., 2021; Mervyn et al., 2014; Shipalana, 2019).

Innovation, prevalent poverty levels, financial sector stability, financial literacy and existing regulatory frameworks that vary between countries each affect the financial inclusion ecosystem (Ozili, 2021). Malladi et al. (2021) posited that individuals finding themselves within the financial inclusion ecosystem due to digitalisation ended up being marginalised and, owing to socioeconomic circumstances, incapable of sustaining themselves within the framework. Malladi et al. (2021) further observed that whereas some consumers might be savvy with technology—relishing apt gadgets—most users might not be able to afford even basic phones, meaning they would likely struggle to understand and utilise technology efficiently, thus leading to significant variance between users in terms of product and service uptake. Finally, Malladi et al. (2021) opined that a lack of financial knowledge and an understanding of financial cybercrimes led to a general mistrust among the marginalised, thereby diminishing digital penetration.

Digitalisation carries the risk of increasing levels of data inequality at multiple levels. At the household level, those without access to data due to constraints such as infrastructure or cost are excluded from the opportunities that are generated through digital platforms. At the global level, there are massive inequalities in the generation and usage of data, which put developing economies at risk of becoming mere originators of raw data to global platforms, then turned into paying consumers of digital intelligence obtained from their data. Additionally, global companies operating digital platforms enjoy significant financial, market and technological power over smaller competing firms that put them in firm control over data markets.

Global development agendas, such as Millennium Development Goals s (2000–2015) and Sustainable Development Goals (2015–2030) set in motion demand for data on complex set of social, economic and political indicators, hitherto believed to be unmeasurable. Digital technologies, supported by satellite interfaces generate massive data by the second which can only be processed by computation capabilities that exist mainly in advanced economies, creating the space for inequality of opportunities. The UN defines this phenomenon as data revolution which is “an explosion in the volume and production of data matched by growing demand for data from all parts of society”. The increasing use of data as a backbone of the Fourth Industrial Revolution has generated interest on who benefits from access to data, how data production shapes the world and how data mining and algorithmic discrimination techniques shape individual and collective identities and life chances.

Cinnamon (2020) identified inequality in data in three dimensions: “access to data, representation of the world as data, and control over data flows.” Countries, and regions vary considerably in their capability and policy in all these dimensions. Africa lags significantly behind the rest of the world where the “digital” divide morphs into a “data divide” and the inequality of opportunities it engenders, and this "digital capabilities" gap is widening over time. Further, Africa does not only lag in the capacity of the internet bandwidth to access, use and transfer data but also in the levels of subscription to mobile phones as well as high-speed mobile network technology. Therefore, the potential for Africa as a continent to remain disfranchised is real and significant unless concerted effort is made by governments to accelerate investment in the information technology infrastructure as well as skills necessary to utilise complex and large data sets for making decisions and facilitate socioeconomic development.

Digital financial inclusion raises key regulatory issues regarding agency, money laundering, financing of illicit activities, regulation of e-money, consumer protection, payment-system regulation and competition. The increasing influence of technology firms underscores the need for global monitoring and a regulatory framework capable of accommodating the competing priorities of different countries and stakeholders, whilst supporting inter-jurisdictional coordination and minimising the risk of regulatory fragmentation. Global action has, however, thus far lacked coordination, emphasising the need for urgent and concerted remedial action (Lopez, 2020). To help meet this need, this issue paper analyses the impact of digitalisation and data governance in the financial sector and focusses on the challenges and emerging opportunities for private and public organisations across Africa.

6.2 The Current State of Affairs

6.2.1 Emerging Concerns

Digitalisation has transformed global financial markets (International Finance Corporation, 2017; Lehner & Simlinger, 2019), and the COVID-19 pandemic has accelerated this transition (Caldwell & Krishna, 2020). As finance has been democratised, credit has become at once more accessible, available and affordable; several platforms, for example, now deploy big-data analytics and AI to build customer credit scores that offer more convenient services, such as digital loans. Digital transformation has also spurred fundamental changes in the customer experience, affecting decisions and transactional processes across the global world of finance (Backbase, 2021; International Finance Corporation, 2017).

Challenges remain, however, even with the greater inclusion delivered by emerging technology solutions. The digitalisation process has raised questions about data governance, as private entities have generated and deployed massive amounts of customer data without necessarily receiving consent or offering assurances of privacy (Davis, 2021). Indeed, there are numerous instances where consumers have balked at having their data used without their consent. Panian (2010) identified the goals of data governance as enabling secure data collection to complement business needs, managing data as an important competitive advantage whilst optimising costs of data storage. Macmillan (2020) argued that successful data governance meant recognising the economic and social utility of data whilst delivering public benefits across the economy and distribution that should be done securely and in accordance with widely observed norms of consumer protection and privacy. Figure 6.1 illustrates the tension between the need for data protection and production with respect to data governance issues.

Ultimately, governance influences financial stability. We adopt the conceptual framework in Fig. 6.2, and we argue that financial data governance requires a system of self and external regulation.

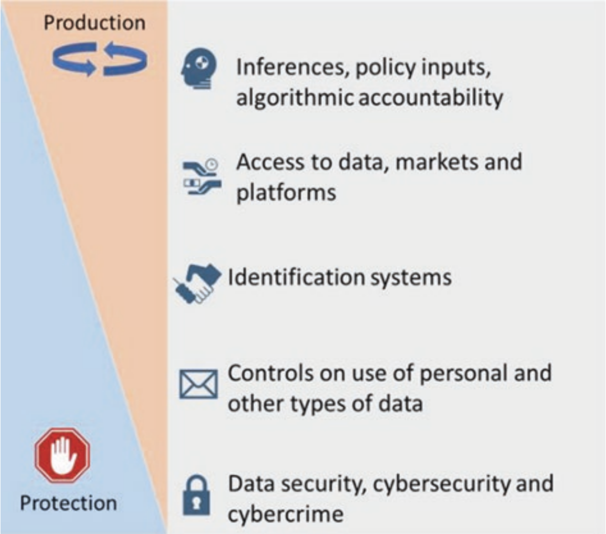

Production
Inferences, policy inputs, algorithmic accountability
Access to data, markets and platforms
Identification systems
Controls on use of personal and other types of data
Data security, cybersecurity and cybercrime
Protection

Fig. 6.1 Data production and data protection. Note. From Data Governance: Towards a Policy Framework, by Macmillan (2020), Industrial Development Think Tank (IDTT)

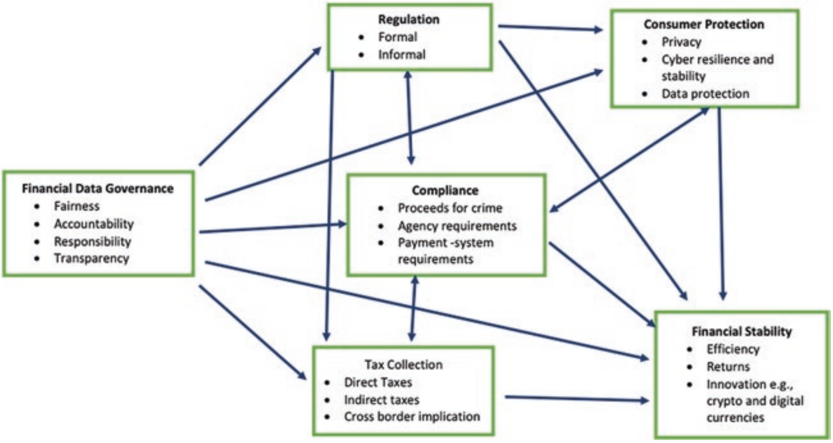


Fig. 6.2 Conceptual framework. Source: Authors

The need for regulation in FinTechs revolves around four principal pillars: compliance (such as preventing money laundering), consumer protection, tax collection and financial stability. There are interacting actions around these pillars, and scholarly work on the same is still fairly nascent.

It is estimated that more than 500 companies across Africa provide financial services enabled by technology (African Union Commission/Organization for Economic Co-Operation and Development, 2021), and the total is even higher when accounting for firms in other sectors that also base their work in technology. Furthermore, the FinTech industry is highly dynamic and is replete with investments, acquisitions, buyouts and partnerships (Wójcik, 2021). Because massive amounts of data are now held by private corporations operating in jurisdictions other than those in which they are based, data governance has emerged a top priority—especially in sub-Saharan Africa (Devermont & Harris, 2021). In response, the European Union of the General Data Protection Regulations (GDPR) developed data-protection laws in 2016, and several African countries have produced similar legislation. Figure 6.3 demonstrates the status of such legislative enactments on the continent.

Laws such as those referenced in Fig. 6.3 address the risk of personal information leaking into the wrong hands while also attending to matters of privacy infringement, concerns about data monopoly, responsibility for

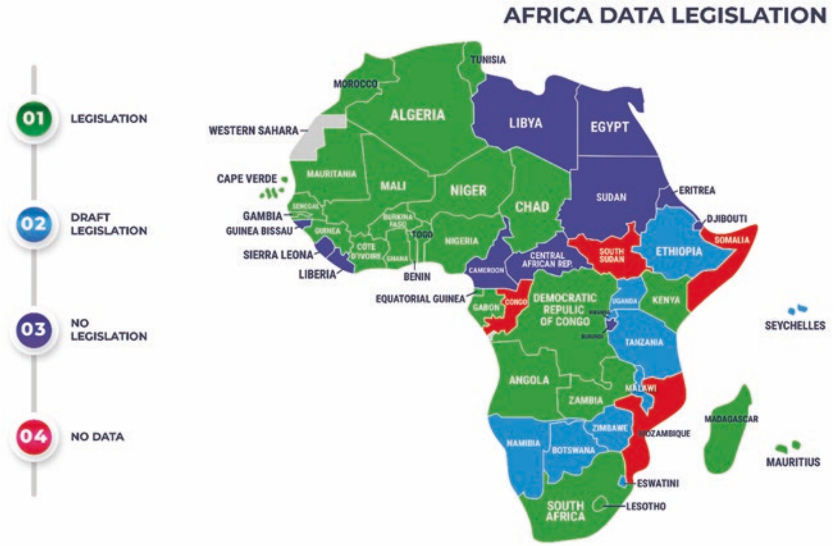


Fig. 6.3 Data protection and privacy legislation in Africa. Note: Adapted from Data Protection and Privacy Legislation Worldwide, United Nations Conference on Trade and Development, 2020, https://unctad.org/page/data-protection-and-privacy-legislation-worldwide

data quality and service and the limits of data governance for Big Data implementation. Additionally, these laws aspire to protect personal information whilst safely accessing and linking open and non-sensitive data; at the same time, they offer proportionality in the use of data and autonomy for individuals and communities deploying data to achieve commercial benefits and efficiency in public services. Most countries, at present, refer to the principles of AI as developed by the OECD.

Yermack (2018) posited that regulation of FinTech companies in developing countries was undeveloped due to immature regional institutions and ever-evolving technology. Malladi et al. (2021) observed that while financial inclusion has broadened, access to credit was still a concern and has caused a spike in predatory lenders charging exorbitant rates of interest, typically due to insufficient penetration by government and credible formal schemes, thereby requiring more outreach to enhance would-be

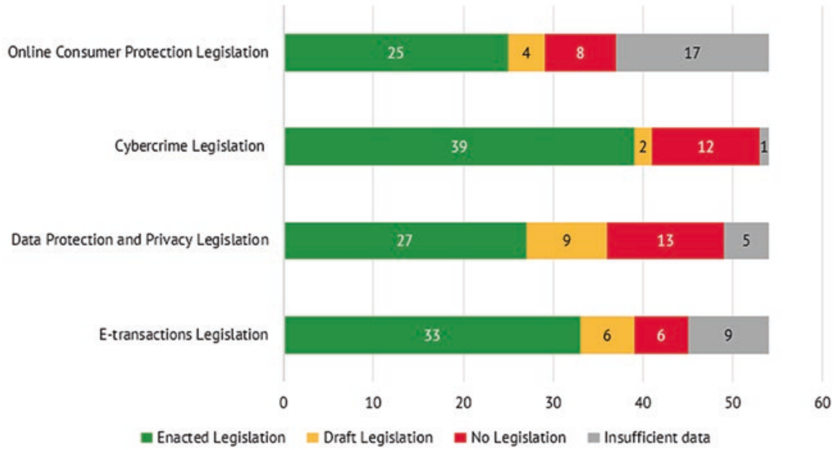


Fig. 6.4 Status of digital protection-related legislation across African countries. Note: From United Nations Conference on Trade and Development E-Commerce Legislation Index, 2021

users' access to credit. Moreover, avenues for digital lending and online loans from credible financial institutions are still missing.

Despite clarity at the continent level, numerous studies have shown that not all countries have progressed towards operationalising their data-protection laws and governance mechanisms (Ademuyiwa & Adeniran, 2020; Deloitte, 2017). According to the United Nations Congress on Trade and Development (2021), African countries are at different stages of legislating various aspects of safeguards on digital transactions and data use.

As indicated by Fig. 6.4, by 2021, only 25 of the 54 countries in Africa had fully enacted legislation on consumer protection. Similar numbers exist regarding cybercrime (39), privacy and data protection (27) and electronic transactions (33). These rates of enactment have significant implications for trade potential and subsequent financial transactions in the region.

6.2.2 Digitalisation Has a Heterogeneous Impact Across Africa

Digitalisation is a multidimensional and rapidly evolving concept (International Finance Corporation, 2017), one that has emerged as a practical and feasible approach to improving governance (World Bank, 2021). Information asymmetries have been reduced by increasing transparency and access to information, leading different organisations to move towards digitalisation. On account of individual-level data, different organisations can offer advanced search outcomes, tailored product, service recommendations, useful ratings, timely traffic data and targeted advertisements (Devermont & Harris, 2021). Furthermore, as the World Bank (2021) has observed, digitalisation can improve linkages between citizens and governments, a connection that is increasingly important as the public and private spheres have become increasing intertwined and reliant on digital technology amid the COVID-19 pandemic. Virtually, every organisation realises that customer data is like gold and a major corporate asset, hence the need for careful protection and active management to preserve the integrity and realise the value of this data.

As many organisations have realised lately, notwithstanding its value, data can be a liability in the event of data breaches (Gregory, 2011; Gressin, 2018; Wang & Johnson, 2018), especially in the financial sector, where such episodes are a common occurrence. Wang and Johnson (2018) estimated that between 2005 and 2018 there were over 8000 documented breaches affecting over 10.3 billion records. Of these cases, more than 2300 breaches implicating 9.8 billion records affected e-commerce transactions in finance and insurance services. For example, in July 2020, a data breach was reported involving as many as 7.5 million banking users who had been exposed on a forum used by hackers to sell and swap ill-gotten data (PYMNTS, 2020).

The African Cyber Immersion Centre (2020) noted a marked increase in cyber-attacks transcending key sectors of the economy and found that these attacks were mostly coordinated across different countries. Cyber resilience, therefore, remains a critical component in the relationship between digitalisation and economic development; the African Union (AU) Heads of States thus adopted and approved the AU Convention on Cybersecurity and Personal Data Protection at the 2014 Malabo Convention (AU, 2014). The convention obliged member states to

> establish in each state party, mechanisms capable of combating violations of privacy that may be generated by personal data collection, processing, transmission, storage, and use; that by proposing a type of institutional basis, the convention guarantees that whatever form of processing is used shall respect

> the basic freedoms and rights of local communities and the interest of business; and take on board internationally recognized best practices. (p. 2)

Verification, vigilance and keen detection efforts are more important than ever in this rapidly changing environment. Without ratification and subsequent operationalisation by the different states of the Malabo Convention, the good intentions will remain just that.

Africa is a large heterogeneous collection of countries with distinct societal objectives and institutional frameworks. There are significant intra-country disparities within the region and at the most basic levels of digitalisation; higher-income countries, for example, enjoy superior connectivity (International Finance Corporation, 2017). Yermack (2018) analysed different patterns of legal systems in Africa, as well as their influence on the adoption of technology and FinTech platforms, and deemed that common law countries (as compared to civil law countries) provided more infrastructure incentives that, in turn, encouraged robust growth for FinTech platforms. Ademuyiwa and Adeniran (2020) further established that most existing laws required substantial amendments to qualify them as appropriate to the dynamics of digitalisation and highlighted emerging concerns regarding policies directed at competition and taxation in the digital space. Table 6.1 offers a summary of progress regarding different aspects of data governance in light of existing trade agreements and protocols.

6.2.3 What Does This Portend for Africa?

Africa has unique demographic advantages that include a young and increasingly literate population, a burgeoning middle class, initiative-taking mobile network operators and escalating internet penetration (Gyori, 2018). Accordingly, digitalising Africa today paves the way for more resilient economies in the future (International Finance Corporation, 2017). Innovative digital financial services, including mobile money, can increase financial opportunity in Africa and transform the landscape of financial inclusion for the unbanked or underserved (Gyori, 2018). By enhancing unconstrained access to financial services, capital, goods and services, these innovations can integrate the poor and minorities into the mainstream economy whilst addressing inequalities. Such innovations in economic development include SME development and support systems, infrastructural development, circular economic activities and increases in

Table 6.1 Missing clauses—digital provisions in Africa trade agreements

	Intellectual Property (IP) Protection *Legal rights that creators have over their works, inventions, technological developments, etc. Examples of IP protection include patents, trademarks, copyrights and trade secrets.*	*Data Protection* *The process of protecting individuals' personal information. Includes measures to secure data, data encryption, masking erasure and backup.*	*Cyber Security* *Protection against criminal use or manipulation of cyberspace, including breaches known as "cyberattacks".*
The Cotonou Agreement (between the EU and the African, Caribbean, and Pacific Group of States) (2000)	Present but does not contain unique clauses governing IP. Relies on adherence to the Trade-Related Aspects of Intellectual Property Rights (TRIPS) and the WTO's agreement regarding the Convention on Biological Diversity	Absent	Absent
EU–EAC Economic Partnership (2014)	Absent; contains a clause dictating that negotiations on IP are to be completed within five years of UK–Kenya EPA start date. However, subsequent negotiations have not resulted in an IP protection agreement	Absent	Absent
AfCFTA Agreement (2018)	Absent; to be included in Phase II of the AfCFTA Negotiations	Absent; to be included in Phase II	Absent; to be included in Phase II
UK–Kenya Economic Partnership Agreement (2020)	Present; contains brief language noting that the clause on IP is replicated from the EU-ACP EPA	Absent	Absent
US–Kenya FTA Negotiation Principles (2020)	Present; identifies IP as a key issue and calls on US to provide capacity building and technical assistance implementing IP provisions once in place.	Absent	Absent
UK-Ghana Interim Trade Partnership (2021)	Absent	Absent	Absent

Note: Adapted from Digital Africa: Levelling up Through Governance and Trade, by Devermont and Harris (2021), Centre for Strategic and International Studies employment. As examples of what is possible through access to technology, some of these innovations enable remote micropayments to and from SMEs, while other proposed FinTech propositions include involving central banks in embracing blockchain technology and transitioning to sovereign virtual currencies. Transformations such as these have vast implications for macroeconomic stability and would likely improve tax collection, compliance and other administrative functions of the government (Yermack, 2018).

At the same time, the International Monetary Fund (IMF) has noted that as countries become increasingly digitalised, they must invest in four broad policy pillars: infrastructure, policy frameworks, skill-building and appropriate risk management policies. At present, numerous African countries lack these data governance pillars (as well as the enforcement capacity) needed to support a well-functioning data-based economy (Adeniran & Osakwe, 2021). The discourse around the consequences of FinTechs remains controversial and should be viewed from a geographical perspective (Wójcik, 2021); in as much as Mpesa was nurtured by, and successful in, Kenya, very few African countries have, within their current legislative framework, taken a stand on the legitimacy of digital currency and initial coin offerings (ICO). Data governance provides invaluable support for these developments.

6.3 Formulating a Research and Policy Agenda

6.3.1 Opportunities Associated with Financial Data Governance

Technology is evolving at a fast pace, and there is a constant need to both evaluate these developments and engage in research to nurture the capacity for the management and governance of data. Financial data governance should be viewed from the perspective that it can create value; if appropriately managed, data governance enables organisations to be competitive and agile, managing costs to more effectively address the needs of their customers (Panian, 2010). The desirable attributes that facilitate value creation include the availability of superior, relevant and consistent data, and compatible technology, combined with clear audibility and easy accessibility (Panian, 2010; Petzold et al., 2020).

What is the expected level of data quality and service attendant to new financial innovations? Inadequate data quality and poor availability yield poor productivity, with significant amounts of time wasted on non-value-added activities, including data sourcing, data aggregation, data reconciliation, data cleansing and manual reporting (Petzold et al., 2020). Hence, effective data governance may require rethinking organisational design to accommodate a balance between the setting of standards, strategic direction and execution—a balance necessarily affected by data complexity (which increases with the scope of business operations), the speed with which core data evolves and the maturity of the underlying and predominant technology.

What are some of the emerging models of organisational design that would best utilise available opportunities? With its focus on systems that appreciate autonomy for individuals and communities, and its emphasis on using open data to achieve commercial benefit and efficiency in public services, digitalisation balances public and private interests. In its Policy Brief No. 89, the United Nations outlined some areas in which the use of open data and big data analytics could be considered useful (United Nations, 2020), such as in attaining real-time reactions to changes in economic phenomena, facilitating coordination and collaboration across different stakeholders in the financial system and enhancing public trust and countering misinformation when explaining complicated economic cycles. The COVID-19 pandemic propelled digitalisation to the forefront of discourse on the nature of the world economy of the future (Wójcik, 2021), but digitalisation can also identify and address vulnerable groups to provide effective resource management. Such value-added actions, however, necessitate enhanced data governance to complement their efforts. The question that follows then is: What have been the linkages between different levels of data governance and resource management?

6.3.2 Challenges and Risks Associated with Big Data in the Financial Sector

As FinTechs expand the geographical limits of the financial sphere, initiating new products and offerings based on copious financial data, they complicate the predicament of regulators (Wójcik, 2021). Firms that have underinvested in governance make their organisations vulnerable to real and often expensive breaches. The concerns around the handling of personal information, as well as privacy infringement by FinTechs, have not been fully assessed but certainly include issues of customer privacy and protection. Data-handling risks arise at three nodes: data collection, processing and even archiving. These risks crystallise when data is inadequately handled at the three nodes and in a manner that compromises the firm. In a January 2020 Consultative Group to Assist the Poor (CGAP) report, Medine and Murthy noted that

> As the commercial use of personal data grows exponentially, so do concerns over whether that data will be used in consumers' best interests. This is particularly true for financial services in emerging economies, where data expand the potential for reaching poor and underserved communities with suitable products but where customer protection risks are great. In many markets ranging from Indonesia to India and Kenya, it is unfair to impose the burden of consent on individuals to protect their data when such a large proportion of the population are opening accounts or coming online for the first time, literacy rates are low, and individuals face potential language and technological barrier. (p. 1)

To address these concerns, Janssen et al. (2020) argued that data governance principles should include evaluation of data quality and bias (before data is used), post-processing validity checks and data minimisation/need-to-know protocols (e.g., only necessary information should be shared as opposed to complete data sets). Other policies advocated by Janssen et al. (2020) include bug bounty schemes (i.e., rewarding those who detect errors and issues), informing when sharing, data separation (e.g., sensitive versus insensitive data) and citizens' control of data (i.e., citizens and organisations should be sufficiently empowered to validate the accuracy of their data). Finally, Janssen et al. (2020) stressed the importance of collecting data at the source, as well as authorisation to access data (including separation of concerns so no single person could misuse or abuse data), distributed storage of data (because distributed systems are less vulnerable and circumvent easy data combining) and the appointment of data stewards for accountability.

In short, the data risks posed by FinTechs can be better managed by an improved regime of data management, adoption of advanced analytics techniques and reliance on cognitive technologies. The overriding concern is the extent to which these principles, policies and procedures have been implemented, as well as the degree to which they have influenced data governance. Malladi et al. (2021) observed that stakeholders exercised disjointed efforts to achieve sustainable last-mile delivery models; furthermore, open access to information like healthcare-schemes data, social-inclusion data, COVID-19 data and vaccination data had not been fully leveraged, demonstrating incidences of incoherence. In a dynamic environment, last-mile technological systems and artefacts are especially vulnerable to exposure and exploitation.

Cybersecurity remains the biggest threat facing digitalisation, with few available experts to mitigate attacks in Africa (International Development Research Centre, 2019). Evidence indicates that risks increase as organisations digitalise and automate their operations (Kaplan et al., 2019). Owing to an emergent worldwide commitment to diminish cybersecurity threats, the International Telecommunication Union (ITU, 2021) developed an index, the Global Cybersecurity Agenda (GCA), to assess countries according to five strategic pillars (Legal Measures, Technical Measures, Organisational Measures, Capacity Building and International Cooperation) and to then aggregate an overall score. Based on this calculus, the top 10 African countries most committed to cyber stability are (with their overall scores in parentheses) as follows: Mauritius (96.89), Tanzania (90.58), Ghana (86.69), Nigeria (84.76), Kenya (81.70), Benin (80.06), Rwanda (79.95), South Africa (78.46), Uganda (69.98) and Zambia (68.88) (ITU, 2018). Each of these countries has, subsequently, addressed challenges in a collaborative manner; because of this collaboration, 18 African countries currently have an institutional framework for reporting cybersecurity incidents.

6.3.3 Social and Ethical Issues Affecting Digitalisation of Financial Data

Beyond the technological challenges of digitalisation are the social and ethical issues it presents. The predominant view of regulatory structures is that they stifle innovation; data privacy continues to be a major concern as vast amounts of captured data is easily available to unauthorised stakeholders because personal information and privacy norms are not honoured (Malladi et al., 2021). There are also technology concerns around some of the technologies like AI, whose algorithms may be discriminatory. Zook and Grote (2020) imagined digitalisation as a decentralised techno-utopian vision of society that would enshrine individual liberty and resist the centralised and surveillance nature of regulations. Royakkers et al. (2018) identified privacy, autonomy, security, human dignity, justice and the balance of power as impacted by digitalisation and, consequently, in need of protection; moreover, they opined that, whereas regulation had been developed around privacy and security, official scrutiny was not as well-articulated in the other four areas (which are inherently fundamental in many modern-day constitutional architectures). The question then becomes: What policies and systems can hold governments and private citizens accountable for the use of financial digital data in a manner that does not run afoul of their basic rights?

6.3.4 Enhancing Institutional Frameworks

Institutional frameworks must be addressed because they affect data collection, transmission, processing, storage, access and interoperability. The preferred approach has, thus far, been regulatory sandboxes that offer co-development of regulation by stakeholders as well as private self-regulation (Yermack, 2018). However, as policymakers chart their way forward and consider various factors, they must question what structures exist to incentivise FinTechs to adopt and adapt data governance. Davis (2021) proposed a series of actions that could strengthen transparency, accountability and participation in data protection; one such recommendation would require proactive verification of the compliance activities by each of the players in this space.

The impact of FinTechs on the fragility and profitability of financial institutions is still uncertain (Fung et al., 2020). The impact of digitalisation through the coordination of decentralised data systems across institutions must still be assessed. What, for example, are the risks to financial data associated with data monopoly? What is the scope for antitrust regulation in so far as data concentration is concerned? Davis (2021) opined that applicable regulatory sanctions and monetary fines should deter breaches of existing governance requirements, but others aren't so sure.

From a regulatory perspective, the limits of data governance for Big Data implementation are unclear, as are the regional and in-country impacts of recently enacted frameworks and legislation. Davis (2021) recommended that regulatory authorities enjoy multiple mandates affording them multi-stakeholder engagements and, where possible, regional collaborations. Is there a scope for regulatory sandboxes that would allow experiments with direct feedback between citizens and the government? In Africa, this terrain is uncharted; more areas call for in-depth analysis regarding the possible challenges and opportunities of digitalisation and financial governance.

6.4 Conclusions and Implications

The literature on digitalisation reveals several emerging themes. Scholars should consider how the abovementioned issues will guide any future discourse on digitalisation and financial governance in Africa. In this issue paper, we have outlined the state of financial-data governance and practices in Africa and have discussed areas that necessitate more investigation to narrow existing gaps in knowledge and policy. Finally, we have identified questions and dilemmas facing scholars, practitioners and policymakers as we all proceed on this journey into a digitalised future.

References

ACIC. (2020). *Africa cyber security report -2019/2020*. Africa Cyber Immersion Centre.

Ademuyiwa, I., & Adeniran, A. (2020). *Assessing digitalization and data governance issues in Africa*. Centre for International Governance Innovation.

Adeniran, A., & Osakwe, S. (2021). *Why digitalization and digital governance are key to regional Integration in Africa*. Centre for Global Development. https://www.cgdev.org/blog/why-digitalization-and-digital-governance-are-key-regional-integration-africa

African Union. (2014). *African Union*. African Union Convention on CyberSecurity and Personal Data Protection. https://au.int/sites/default/fles/treaties/29560-treaty-0048-_african_union_convention_on_cyber_security_and_personal_data_protections_e.pdf

AUC/OECD. (2021). *Africa's development dynamics 2021: Digital transformation for quality jobs*. African Union Commission/Organization for Economic Cooperation Development. https://doi.org/10.1787/0a5c9314-en

Backbase. (2021). *Banking 2025: Four pillars of the digital First Bank*. Backbase. https://go.backbase.com

Caldwell, J., & Krishna, D. (2020, July 30). *The acceleration of digitization as a result of COVID-19*. Deloitte. https://www2.deloitte.com/global/en/blog/responsible-business-blog/2020/acceleration-of-digitization-as-result-of-covid-19.html

Chitimira, H., & Torerai, E. (2021). The nexus between mobile money regulation, innovative technology and the promotion of financial inclusion in Zimbabwe. *Potchefstroom Electronic Law Journal*, *24*, 1–33. https://doi.org/10.17159/1727-3781/2021/v24i0a10739

Davis, T. (2021). *Data protection in Africa: A look at OGP member progress*. Open Government Partnership.

Deloitte. (2017). *Privacy is paramount: Personal data protection in Africa*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/za/Documents/risk/za_Provacy_is_Paramount-Personal_Data_Protection_in_Africa.pdf.

Devermont, J., & Harris, M. (2021). *Digital Africa: Levelling up through governance and trade*. Centre for Strategic and International Studies. https://www.csis.org/analysis/digital-africa-leveling-through-governance-and-trade

Enberg, J. (2019, October 24). *Global mobile payment users 2019*. eMarketer. https://www.emarketer.com/content/global-mobile-payment-users-2019

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., & Zbinden, P. (2019). *BigTech and the changing structure of financial intermediation* (Working Paper No. 779). Bank for International Settlements. https://www.bis.org/publ/work779.pdf

Fung, D., Lee, W., Yeh, J., & Yuen, F. (2020). Friend or foe: The divergent effects of FinTech on financial stability. *Emerging Markets Review, 45*(C), 100727. https://doi.org/10.1016/j.ememar.2020.100727

Gregory, A. (2011). Data Governance—Protecting and unleashing the value of your customer data assets. *Journal of Direct, Data and Digital Marketing Practice, 12*(3), 230–248. https://doi.org/10.1057/dddmp.2010.41

Gressin, S. (2018). Federal Trade Commission. *The Marriott data breach*. https://www.consumer.ftc.gov/bog/2018/12/marriot-data-breach

Gyori, D. (2018, February 23). *Africa's ten key advantages in digital transition*. The Asian Banker. http://www.theasianbanker.com/updates-and-articles/africas-ten-key-advantages-in-digital-transition

IDRC. (2019). *Shaping an Internet for women's empowerment*. International Development Research Centre. https://www.idrc.ca/en/research-in-action/internet5-shaping-internet-womens-empowerment. International Telecommunications Union (2021). Global Cybersecurity Ind.

International Finance Corporation. (2017). *How fintech is reaching the poor in Africa and Asia: A start-up perspective*. World Bank Group. https://www.ifc.org/wps/wcm/connect/d3217a44-60d5-45f0-a590-caa24482e22f/EmCompass+Note+34+DFS+and+FinTech+Mar+28+FINAL.pdf?MOD=AJPERES&CVID=lIlOTMQ

International Telecommunications Union. (2018). *IG_workshop_August2018*. https://www.itu.int/en/ITU-D/Capacity-Building/Documents/IG_workshop_August2018/Presentations/Session5_SergeZongorev.pdf

International Telecommunications Union. (2021). *Global cybersecurity index (GCI)*. https://www.itu.int/pub/D-STR-GCI.01

Jafri, J. (2021). *Fintech, philanthropy and development: Emerging issues with digital inclusion*. Financial Geography Working Paper Series: ISSN 2515-0111, 2.

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy artificial intelligence. *Government Information Quarterly, 37*(3), 1–8. https://doi.org/10.1016/j.giq.2020.101493

Cinnamon, J. (2020). Data inequalities and why they matter for development. *Information Technology for Development, 26*(2), 214–233. https://doi.org/10.1080/02681102.2019.1650244

Kaplan, J., Richter, W., & Ware, D. (2019). *Cybersecurity: Linchpin of the digital enterprise*. McKinsey & Company. https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/Risk/Our%20Insights/Cybersecurity%20Linchpin%20of%20the%20digital%20enterpri

Kirui, B. (2020). *The role of mobile money in international remittances: Evidence from Sub Saharan Africa*. AERC working paper.

Lehner, O. M., & Simlinger, R. (2019). When function meets emotion, change can happen: Societal value propositions and disruptive potential in fintechs. *International Journal of Entrepreneurship and Innovation, 20*(4), 277–288. https://doi.org/10.1177/1465750319857974

Lewis, R., Villasenor, J., & West, D. (2017). *Building a secure and inclusive global financial ecosystem*. Brookings Financial and Digital Inclusion Project report p. 68.

Lopez, C. (2020). *Principals of financial regulation for Big Tech*. Observer Research Foundation. Digital frontiers. https://www.orfonline.org/expert-speak/principles-financial-regulation-big-tech/.

Macmillan, R. (2020). *Data governance: Towards a policy framework*. Industrial development think tank (IDTT).

Malladi, C., Soni, R., & Srinivasan, S. (2021). Digital financial inclusion: Next frontiers—Challenges and opportunities. *CSIT, 9*, 127–134.

Mervyn, K., Simon, A., & Allen, D. K. (2014). Digital inclusion and social inclusion: A tale of two cities. *Information, Communication and Society, 17*(9), 1086–1104. https://doi.org/10.1080/1369118X.2013.877952

Muralidhar, S. H., Bossen, C., & O'Neill, J. (2019). Rethinking financial inclusion: From access to autonomy. *Computer Supported Cooperative Work, 28*(3–4), 511–547. https://doi.org/10.1007/s10606-019-09356-x

Muthiora, B. (2015). Enabling mobile money policies in Kenya: Fostering a digital financial revolution. *GDMA mobile money for the unbanked*.

Ozili, P. K. (2021). Financial inclusion research around the world: A review. *Forum for Social Economics, 50*(4), 457–479. https://doi.org/10.1080/07360932.2020.1715238

Panian, Z. (2010). Some practical experiences in data governance. *World Academy of Science, Engineering and Technology, 62*, 939–946.

Payments. (2020). *Security and risk*. FinTech Dave reports data breach involving 7.5M users. https://www.pymnts.com/news/security-and-risk/2020/fintech-dave-data-breach-hackers/.

Petzold, B., Roggendorf, M., Rowshankish, K., & Sporleder, C. (2020, June). *Designing data governance that delivers value*. McKinsey Technology.

Republic of Kenya. (2017). *Kenya Social Protection Sector review report 2017*. Ministry of Labour and Social Protection.

Rijswijk, K., Bulten, W., Klerkx, L., den Dulk, L., Dessein, J., Debruyne, L., & en Nematoden, O. (2020). *Digital Transformation: Ongoing digitisation and digitalisation processes.* Desira.

Rolfe, K. (2019). Payments industry intelligence. *Mobile money transactions equivalent of half of Kenya's GDP.*

Royakkers, L., Timmer, J., Kool, L., & Est, R. (2018). Societal and ethical issues in digitization. *Ethics and Information Technology, 20*, 127–142.

Sengere, L. (2017). *EcoCash has processed over $23 billion since launch and that's not the only impressive figure.* TechZim.

Shipalana, P. (2019). *Digitising financial services: A tool for financial inclusion in South Africa?* South African Institute of International Affairs.

UNCTAD. (2020). *Data protection and privacy legislation worldwide.* United Nations Conference on Trade and Development. https://unctad.org/page/data-protection-and-privacy-legislation-worldwide

UNCTAD. (2021). *Summary of adoption of e-commerce legislation worldwide.* United Nations Conference on Trade and Development. https://unctad.org/topic/ecommerce-and-digital-economy/ecommerce-law-reform/summary-adoption-e-commerce-legislation-worldwide

United Nations. (2020). *Strengthening data governance for effective use of Open data and big data analytics for combating COVID-19.* United Nations Policy. Department of Economic and Social Affairs, United Nations.

Wang, P., & Johnson, C. (2018). Cybersecurity incident handling: A case study of the Equifax data breach. *Issues in Information Systems, 19*(3), 150–159.

Wójcik, D. (2021). Financial Geography II: The impacts of Fintech—Financial sector and centres, regulation and stability, inclusion and governance. *Progress in Human Geography, 45*(4), 878–889. https://doi.org/10.1177/0309132520959825

World Bank. (2021). *Data, digitalization and governance.* The World Bank Group.

Yermack, D. (2018). *Fin Tech in Sub Saharan: What has worked well and what hasn't* p. 25007. National Bureau of Economic Research Working Paner No.

Zook, M., & Grote, M. H. (2020). Initial coin offerings: Linking technology and financialization. *Environment and Planning A, 52*(8), 1560–1582. https://doi.org/10.1177/0308518X20954440

CHAPTER 7

More Than Just a Policy: Day-to-Day Effects of Data Governance on the Data Scientist

Vukosi Marivate

7.1 Introduction

The continued rise of the information economy meant an increase in the use of data to build and deploy many data-driven products. These data-driven products are used to extract meaningful insights from raw information, which is then used to address challenges across many different fields. This has coincided with the emergence and development of Data Science as a unique field of expertise, building data-driven products. Data Science is unique from Computer Science (the study of theory and practice of how computers work), and it encompasses many fields. From the perspective of users, the data-driven products have brought many new services and conveniences.

In health, for example, there were rapid deployment of data tools to inform the public on the COVID-19 pandemic (Alamo et al., 2020; Shuja et al., 2021), pandemic prediction models (Ray et al., 2020) and estimations of impact of COVID-19 (Bradshaw et al., 2021). At the same time,

V. Marivate (✉)
Department of Computer Science, University of Pretoria, Pretoria, South Africa
e-mail: vukosi.marivate@cs.up.ac.za

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8_7

some of the tools developed to deal with diagnostics/treatments were not as successful. An example of such data-driven products are the many tools/algorithms that were developed or deployed to improve radiology scans (Roberts et al., 2021; Wynants et al., 2020). On the one hand, one may be tempted to say such deployments were a complete failure. However, on the other hand these challenges highlight some of the shortcomings of data tools and areas of improvement. More importantly, these challenges outline the need to manage data (and its products) so that we take into account the human factors and impacts data may have across all domains. Keeping with the COVID-19 topic, the pandemic also put a spotlight on the lack of basic data infrastructure (Mbow et al., 2020), lack of data skills and/or lack of political will in many countries to focus on the improvement of data-driven products. These data-driven products and tools ultimately impact on the quality of responses to the pandemic. The aforementioned examples, highlight the need for Data Governance that takes a refined view of data in.

I look at the Data Scientist (or Data Science Team) as the ones who make most of the decisions on the data tools they develop or create. This simplified view does not encapsulate all the challenges associated with what is currently taking place. It would be better to look at data-driven products through the lens of socio-technical systems. Socio-technical systems are systems which have interactions between humans, machines and the environment (Baxter & Sommerville, 2011). Even within the organization, the Data Science Team or Data Scientist cannot make decisions without a variety of different stakeholders, especially decisions that have an impact on humans and other environmental factors. As such, the Data Scientist should be able to understand the other inter-dependencies of organizations and society to better understand where they fit and that governance structures should exist to guide the development of systems with such inter dependencies.

In this work, I aim to provide a better understanding of the governance/human factors that Data Scientist and organizations should be aware of. To address this challenge, I will answer fundamental research questions for the domain.

Research Question: What are the salient points that Data Scientists should be aware of when it comes to Data Governance within organizations?

Research Sub-Questions:

- Do the current policies or mechanisms on the African continent provide a coherent view that can be used by Data Scientists to navigate and respond appropriately to the needs of the organization.
- Can we learn from the ICT4D community to better understand how interventions should take care of more than just deploying a tool.

It is important to contextualize why we need to answer these questions. We are at a time where policy is lagging deployment of data tools (this is discussed in this paper). This means there are gaps and blind spots that both Data Science practitioners and policy makers (both in public and private sectors) have. These blind spots have consequences. There has been much written about the data protection policy making and much written about Data Science practice and limitations. In this work I want to link the two in order to have a joint understanding that decision making has to be done together. The rest of the document is organized as follows. First, I look at the field of Data Science and how Data Governance fits into practice. The next step is to look at Data Governance on the African continent. I will set the scene and identify gaps that then intersect both areas of Data Science and Data Governance. In the proceeding section, I discuss how ICT4D may have already blazed a path that allows us to learn from in understanding the interactions of Data Science and Data Governance. The latter sections deal with the different stages of the Data Science process and proposals on how best Data Scientists can navigate human factors such as privacy, bias and security. Lastly, I conclude and summarize the viewpoints and evidence elaborated on in this paper.

7.2 Data Science and Practice

I first look at the practice of Data Science and its connections to Data Governance. As such I provide an overview of what Data Science is. An important definition that is still evolving but is important for joint understanding between the reader and the author.

7.2.1 What Is Data Science?

Data Science is a discipline that has arisen due to a number of factors. Data Science itself is a field that uses scientific modelling techniques (typically from a diverse set of scientific disciplines) to extract patterns/information/knowledge from a wide variety of data (Dhar, 2013). The rise in this discipline has been swift for many reasons. Organizations (public and private) have been working to explore the data that they have amassed over time and mine information for patterns and trends that may give them a competitive advantage. There has been an explosion in the number of large internet-based organizations and internet-generated content. Simply, with more users on the internet, and more content on the internet, the information economy needs better data and data tools to monetize these users (Mandl & Kohane, 2016; Zhang, 2017) (e.g. for advertising) or for services that motivate users staying within a company's products (a walled garden) (Best, 2014; McCown & Nelson, 2009; Skorup & Thierer, 2013).

On the side of public organizations, Data Science has meant the work to analyse or collect data that improves on services provided by governments or new forms of ways to understand citizens (sometimes resulting in mass/hyper surveillance. It is very important to understand these factors, especially as they are connected to "value creation in the information age". Consideration of the political economy of data, whereby incentives for the monetization of data may be at odds with the interests of private citizens is critical. Issues of concern include the ability of data scientists to shape and influence data governance around private incentives, as well as their ability to collect and utilize information for purposes beyond the intentions of the individual providing data (Nyamwena & Mondliwa, 2020). The factors necessitate that we understand the foundational data infrastructures (physical, virtual, human and otherwise) through the lens of governance, specifically Data Governance. Let us first break down the process of Data Science.

7.2.2 The Data Science Process

To provide the reader with better understanding of Data Science, I use the data analysis cycles to provide an insight into the typical Data Science Process. One can use the CRoss Industry Standard Process for Data Mining (CRISP-DM) as a representation of the process (Wirth & Hipp, 2000). The steps are typically: (a) understand a business problem, (b) understand the data required, (c) collect data, (d) prepare data, (e) perform modelling, (f) evaluate the solution to the problem and (g) adjust understanding and/or deploy (see Fig. 7.1).

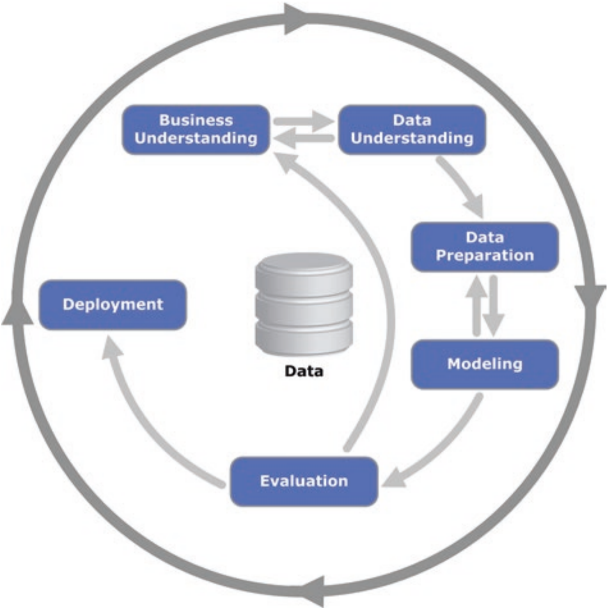


Fig. 7.1 CRISP-DM flow model. Source: Wikimedia Jensen (2012)

One notes that all of this focuses on solving a business challenge. We can easily extend this to solving any societal/organization/scientific challenge, it does not need to be business. This process is similar to the Epicycles of Analysis (Peng & Matsui, 2015) that splits the processes of the problem and the analysis for a solution to the problem. The former tries to separate the problem formulation from the modelling. Problem formulation takes understanding the correct data to gather or get access to. Ultimately, with all of these, we need to understand the human factors and dimensions that arise in all parts of the cycles. The inter-dependencies are discussed later in the document.

The rise of Data Science has also coincided with the rise of Machine Learning and Artificial Intelligence (West & Allen, 2018), and typically it is expected that Data Scientists have an understanding of, and can use, concepts from these fields (Tang & Sae-Lim, 2016). Machine Learning is a field of study concerned with creating tools that learn analytical models from data (Alpaydin, 2020) and is a subset of Artificial Intelligence. Artificial Intelligence is a field of study concerned with creating machines which mimic the intelligence of humans, typically defined as creating an agent that can perceive its environment, and perform actions to maximize some utility or achieve some goal(s) (Russell & Norvig, 1995).

Many Data Science researchers/practitioners are also Artificial Intelligence and/or Machine Learning practitioners/researchers. As such, from here on I will refer to Data Science researchers/practitioners even if I am talking about Artificial Intelligence and/or Machine Learning. Many Data Science researchers or practitioners are comfortable with the above models of understanding data and the subsequent analysis. For this to be successful, society and organizations have an over growing need to understand what actually happens during developing and deploying a system or model in the real world. Governance, in more ways than one, comes into play. The data collection needs considerations of humans and the human dynamic (Bender & Friedman, 2018; Gebru et al., 2018; Jo & Gebru, 2020). The choice of modelling requires consideration of people and their needs (Mitchell et al., 2019), the deployment further requires the consideration of the human dimension in all its guises (Raji, Gebru, et al., 2020; Raji, Smart, et al., 2020). As such Data Governance can be a useful tool for the Data Scientist to be aware of these human factors and the challenges when humans and data [collection, modelling or products] interact (Buolamwini & Gebru, 2018; Hooker, 2021; Ledford, 2019; Mehrabi et al., 2021; Sujan et al., 2019).

7.2.3 Why Do We Need Data Governance?

From the perspective of governments, as part of economic development and growth, they want to embrace "value creation in the information age" (Nyamwena & Mondliwa, 2020). To do so, the collection, use and flow of data has to be governed in order to be able to have oversight over this value creation. In short, Data Governance has to touch every part of the Data Science life cycle as discussed earlier. Data Governance also rises to prominence as a result of historical pushes for digitization of countries especially that of African countries. Governments are concerned that if they do not capitalize the data opportunity, they will be left behind on another economic development. The challenge arises when we look at ways Data Governance has to be shaped for different countries. Without adequate Data Governance in countries, the opportunities for both public and private sectors are at risk of not realizing the full potential of the information economy. This is a big risk as products that may fall short of the values of the countries citizens may be deployed and ultimately cause harm. Such examples of falling short are inadequate privacy protections (Osakwe & Adeniran, 2021), limitations on what data can be used for, regulation of data-driven products that could be harmful (Metcalf & Crawford, 2016), guidelines on data sovereignty (Hummel et al., 2021), and how specific sets of data should be treated as public goods to be shared within or outside a country (Borgesius et al., 2015). Good Data Governance is not only about the data creation stage, but about how governance permeates the full Data Science cycle (Metcalf & Crawford, 2016). Furthermore, good Data Governance requires the contextual knowledge of and from decision makers (in both public and private sectors) to understand the Data Science cycle (data, modelling, algorithms, etc.) (Kearns & Roth, n.d.). It is harder for the gatekeepers to regulate industry if they themselves do not have a foundational understanding of what typically happens within the Data Science cycle. This is an important point to highlight because industries such as finance, for example, have well defined regulators in most countries. These financial regulators regulate the industry to mitigate corruption and harm. Regulatory boards are made up of experts in the field who then work to set best practice, limitations and also penalties for breaches of the regulations. The challenges with many of the data-driven products we see nowadays is that many of the decision makers in the process of deploying these tools have little experience with the field itself and see most of what is going on as a black box that takes in data, and "magically" produces answers. This highlights the needs for basic foundational regulation that asks the right questions when developing data-driven products but also sets the path for a joint understanding of the field which should be understood by all people (not just experts). In the proceeding section I look at important parts of the Data Science cycle and highlight the human factors and questions that should be asked by Data Scientists and also be understood by decision makers.

7.3 Human Factors and the Data Science Cycle

In order to champion the joint understanding of Data Science and Data Governance, in this section I discuss the human factors in the Data Acquisition, Modelling and Presentation phases of the Data Science cycle.

7.3.1 Data Acquisition

One of the steps that is fraught with tension in the Data Science process is the data acquisition process. This can be a blind spot (Mitchell et al., 2018; Zhang et al., 2018) that can make or break many projects. Imagine using a dataset collected in the 1950s on financial lending by banks. Now building a predictive tool to assist in lending decisions with such a dataset will be full of gender and racial biases in many countries (Bond & Tait, 1997; Rice, 1996). Put simply, the model would learn to discriminate. This is still a challenge today (Runshan et al., 2021). Even if the data is taken as representative of the population being studied, it may encode societal bias and discrimination. Most times when talking and interacting with decision makers or clients, those without much experience tend to overlook the challenges in the acquisition of data. These challenges are connected with governance issues (Veale & Binns, 2017).

7.3.2 Processes and Procedures

In acquiring data, as part of the Data Science process, one connects the problem being approached with the data that will be needed to solve the problem. At some point, there may be data before the questions are clear, while at other times there is a question to be answered but the data has not been mapped out. In all instances, data has to move from where it rests and staged for processing by the Data Science team. This requires identification of the relevant data source, identification of which subset of the information is important and how the transmission will occur. In doing these identification steps we have to look at the human factors.

7.3.3 Human Factors

For each of the proceeding steps of the Data Science process, I focus on these three human factors. For the Data acquisition I focus on: Where does the data come from? Why is/was it being collected? Who is the data about? There are many more factors, but for conciseness and to communicate our message, the message will remain with three factors per step of the Data Science cycle. Where does the data come from? When identifying the source of data, it quickly becomes clear that one has to understand the structures of the organizations internally or externally that control access and use of the data. In an ideal case, there is a clear Data Governance structure that also provides information on how a data scientist can request data, how the data should be handled and any sensitive and salient information that the scientist should be aware of (Abraham et al., 2019). There would be questions that are related to the sensitiveness of the data. Was the data collected in an ethical manner? Is the data part of an open data repository? What licensing is the data under and expectations of use? Is the data from a governmental entity, what are the national expectations on Open Government data? For example, in a municipality, one may expect that aggregated water use data by municipal ward should be open and available (especially as many areas in some countries face water shortages), but there may be some resistance by some officials in making this data available.

It may be that there is not enough human resource to create and keep the data available, the data may normally be available for a fee that adds to revenue, there may be issues of transparency etc. Why is/was it being collected? This is an important factor as it establishes prior expectations on what the data that was collected or is being collected was used for. If we imagine that we have data about the transaction habits of bus riders in a city, the original use of the data and expectation was to manage the transportation system. If now the data will be used to understand behaviour to deliver advertising to bus riders, this new use may not be covered by original terms of reference. More importantly, bus riders may not agree with the change of the use of their data and there is a responsibility the organization has with them to treat their information with care and thought.

Who is the data about? In carrying through the process to build up the data one has to think if it is representative of the population it is serving. Again, focusing on when the data is about people, we need to understand who the data represents and if this distribution is equitable, fair (Mitchell et al., 2018; Zhang et al., 2018)? Further does this distribution of people actually match those we expect to make decisions about in the end data-driven product? If not, this may be a problem that introduces biased decision making. For example, in the recent decade, much has been highlighted about the bias in facial recognition systems (Raji, Gebru, et al., 2020). Some of this bias comes from the original data that was used to train them (Mitchell et al., 2018; Zhang et al., 2018). Some of this bias comes from the designs of the systems and also how success is measured. I will discuss more on this later in the modelling and the presentation subsections). One can see just from looking at the above, that there are important human factors that cannot just be left to the Data Scientist or organization to make decisions about. There needs to be foundational expectations on data handling, data storage, security, ethics and regulatory tests on what the data would be used for.

7.4 Data Analysis and Modelling

In the Data Analysis and Modelling step, the Data Scientist focuses their energy on using the correct approaches to extract meaningful information from the data. These choices will influence the final result as well as be the foundation on which many will choose to believe the results or not. Even though these may be established computational, statistical or mathematical approaches, we still need to understand how choices impact the end product and people.

7.4.1 Processes and Procedures

The Data Scientist takes the data that has been acquired in the prior step. They then work to clean it, transforming it into a form that can be used by downstream modelling tasks and then loading it into their modelling systems. The Data Scientist will make choices on metrics to be measured or optimized. Ultimately, these metrics are used to decide on success and then are used to know if new data should be sourced, the question should be re-framed or can one move to the next step of the Data Science cycle.

7.4.2 Human Factors

For the data analysis and modelling stages I focus on these factors: How are the modelling choices made? Who has the skills to model? What are the models for the use-case being used? How are the modelling choices made? For a period, there was a popular retort that people are biased and machines are unbiased. When it has highlighted that machines cannot be unbiased because the data that they use to learn may be biased, the needle moved to that algorithms cannot be biased, only the data (Birhane & Cummins, 2019). But, this still ignores many factors that modelling choices also impact the results of the final models (Jiang et al., 2020). In Machine Learning, we pride ourselves in working to build better and better generalizable, accurate and efficient algorithms, but this does not absolve us about thinking about our modelling choices (Birhane et al., 2021). Work by Hooker et al. (2020) highlighted the biases in compressed models. Further, more and more ML models use transfer learning (building on prior models or datasets), this then carries forward biases. This is one of the reasons Data Scientists should work to document their modelling choices (Mitchell et al., 2019). Modelling may seem insignificant at the time of decision making, but may lead to big consequences later. A recent example (Birhane et al., 2021) is how models influences the collection of massive (in order to fight against bias) dataset that, when looked at under a microscope, to not be as representative as the dataset authors claimed. This highlights the lack of participation and inclusive design choices that also call in to question, who has the modelling skills?

Who has the skills to model? ML/AI/Data Science is a field that typically is skewed in terms of demographics and who ends up building the underlying technologies. One may argue that this does not apply on the African Continent when it comes to racial makeup. But that is not a true reflection of the field. For a long period, in major technology companies on the continent, the senior technical roles were skewed Male and White (mirroring the challenges that have been criticized about Silicon Valley). Further making this worse is the lack of Data Science skills on the continent. Without these skills, we further have less connection between decision makers and those who design models. How many of the decision makers have a data/computational background? Another factor is that the major tech companies that do drive most of the internet economy tend to only have business offices on the continent (Birhane, 2020). Their aim, to sell their services (Birhane, 2020), extract data (Coleman, 2018) and handle regulatory issues (if there is regulation (Birhane, 2020; Coleman, 2018)). The offices do not build or shape the core technologies at these companies. As such, if we connect this question to the prior one, we see how modelling choices can become a life changing decision for those on the downstream tasks. Imagine how in organizations, automated hiring systems, were deployed to assist in the hiring process by using AI to screen or monitor candidates. These systems have been shown to be discriminatory (Sánchez-Monedero et al., 2020), but what are the odds that the decision makers and internal Data Science teams had the skills to be able to evaluate their facial recognition systems or text screening services against bias?

What are the models for the use-case being used? Recent work in the ML/AI field has brought about focus explainable models in the fight against harm and pursuit for better fairness. These choices of such models are in every use-case. Let's take, for example, the increase in surveillance systems and facial recognition systems internationally [ref]. How the models are chosen for such use-cases and evaluated impact the ultimate impact these systems will have on society. Much work has highlighted how biased facial recognition systems (Raji, Smart, et al., 2020) can lead to discriminatory behaviour by law enforcement. This may end up being a life of death situation for someone at the end of these automated systems. A Data Scientist and decision maker needs to ask themselves, what is the cost of an error of our model? These should then impact how the deployment is done. Further, depending on the societal expectations, there may be regulatory restrictions in making one choice or another.

7.5 Presentation and Deployment of Data-Driven Products

The final step in many Data Science projects is presenting results to decision makers and/or the deployment of the data driven products.

7.5.1 Processes and Procedures

In this step, the Data Scientist would work to present a report on findings of the modelling in order to answer the original questions. From here, decisions may be made on these reports. Reports may be visualization, simulations or data-driven products with metrics that show their efficacy. Decisions on what to show and who the data-driven products will be aimed at will be made. These have human factors.

7.5.2 Human Factors

For the Presentation and Deployment of data-driven products stages, I focus on these factors: What decisions are being made with the models? What choices are being made in what to be shown? How will the models be kept updated? What decisions are being made with the models? The ultimate test for the usefulness of a model for the decision maker is when it is deployed for used or presented for decision making. This is a spot in the Data Science life cycle that requires careful understanding of the prior parts of the cycle or wrong decisions could be made. When looking at the data product or predictions of a model, the user must understand how the model works, how it was built and what limitations it has. The sub-question here could be, how do people interpret the results/predictions from the data product? This requires more than just displaying a result but also working with human computer interaction practitioners to design in such a way that is fair, transparent and mitigates bias or discrimination (Holstein et al., 2019; Lee & Singh, 2021).

What choices are being made in what to be shown? As in the statistical domain, we can also lie with data-driven products. The COVID 19 pandemic had many examples where decision makers worked to distort data, distort model predictions and even censor data researchers and practitioners in order to fit with a view that the decision maker held (A hostile environment, 2021; Vigjilenca, 2020; Zhang & Barr, 2021). This may be taken as an extreme public example, but this does happen in many ways. One may be testing for harm at run-time. How will the models be kept updated? When deploying data-driven products, the internal models have to be kept updated. The world did not stop changing when the model was trained and deployed. As such, the models will start exhibiting drift. This drift may also come from the how users respond to what the model does itself. Does the organization of Data Science team have procedures on the maintenance of the models in the data-driven product and how to test for drift before the system has high error in its results (predictive, prescriptive, diagnostic etc.?).

In this section I have discussed how Data Science and Data Governance intersect. In the latter part of the section, I chose three sections of the Data Science cycles to be able to analyse for human factors. Through identifying these human factors, we can better understand how Data Governance is an integral part of the full cycle as decisions being made by the scientist will impact users and humans in general. In the next section I then discuss Data Governance on the African continent.

7.6 Data Governance and the African Continent

With calls for African countries to jump on to the current advances of data driven economies, there has been some movements towards strategies and governance policies by governments that cover data. The African Union released the "The Digital Transformation Strategy for Africa 2020–2030" (African Union, 2020). This strategy should be understood in the context of the wider and more localized Data Governance and digitization challenges in different African countries. When it comes to privacy, the European general data protection regulation (GDPR) (European

Commission, n.d.) has had wide ranging effect and impact on the internet economy as many companies who processed European citizen data had to abide by the rules set out by the EU. Around the African continent, as shown by the research in (Davis, 2021), there are efforts to strengthen data protection policies, even with only about 52% of African countries having such legislation.

The African Union Convention on Cyber Security and Personal Data Protection (known as the Malabo Convention) (African Union, 2014) was adopted by AU member states in 2014. It sets out to provide protections for cyber infrastructure, protection of personal information, cyber security and the necessary foundations to enable an information economy across the African continent. Even though ratified in 2014, only eight countries had ratified the convention by 18/06/2020.[1] The convention touches on many aspects that can form a unified foundation for African countries to benefit from the information economy. Without ratification, we have the reality that organizations and practitioners do not have a unified view on how to deploy data tools and for some countries the reality is much worse with very lax or non-existent protections (Davis, 2021).

In South Africa the Protection of Personal Information Act (POPIA) (Government of South Africa, n.d.), which has taken many years to get enacted, has also begun a discussion in the public on data acquisition, protection of personal information and the use of the data for downstream tasks (especially when it is not for the original purpose of data collection). Even so, Data Governance is not only the protection of personal information, but there are also many more human and organizational factors that data interacts with. I hope the preceding section has made it clear that Data Governance should cover more than just the data being used. But, as earlier discussed, there are many human factors that should be taken into consideration in all the stages of the Data Science cycle. To effectively govern the full process, countries have to have a clear understanding of the stages as well as the responsibilities of governments towards the Data Scientists and the responsibilities of the Data Scientists towards the public.

The African continent has made big strides in the ICT sector and building local skills and also championing local companies. Even so, there is still

[1] https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection biases and lead to discrimination. This illustrates another governance gap (whether planned or unplanned) as decision makers have to be able to evaluate the risks and harms such systems may pose to the population (Mudongo, 2021).

a dominance of the Big Tech Giants (Microsoft, IBM, Google, Facebook etc.) on the continent physically or with services that cross borders. Even though we do not have an agreed definition of the data skill gap, the work by (Sey & Mudongo, 2021) highlights how there is lack of understanding of the need of AI skills and that we need to have efforts to build these skills on the continent and this must connect public and private sectors. These insights are important as they place in context how few of the Big Tech firms have few or any research and development that is done on the continent. AI governance skills are recommended as part of the development of AI skills on the continent (Sey & Mudongo, 2021), echoing the message in this paper on the broader Data Science and Data Governance nexus.

The continent risks being just a source of data (Birhane, 2020) to build services that then are used by citizens without any local development of these services. This has been recently brought to bear with how Facebook only has 13% of its abuse team (which fights abuse on their online platforms) working on non US content, even though 90% of Facebook users are outside the US (Purnell et al., 2021). This is important as misinformation on Facebook outside the US has effect on many countries, but cannot be battled by Facebook itself. Further on, governments have to be able to govern the digital space and insure that the citizens get to benefit from the digital public goods (Gillwald & van der Spuy, 2019). Another challenge is the use of some of the data-driven products for surveillance by both governments and private sector on the continent (Mudongo, 2021). As already highlighted, the systems are less likely to be developed locally and may encode.

7.7 Case Study: Learning from Our Recent Past, Enter ICT4D

Data Science and Artificial Intelligence hailed as a silver bullet to many problems, data itself referred to as the new oil to be exploited by nations and organizations (Hirsch, 2013). But a challenge that organizations and nations should be able to spot rears its head again. With the rise of ICT and digitization efforts, many problems were pointed to where ICT could be the solution (Curtis, 2019). Throw in development practices, ICT4D has been a force for the last two or more decades (Walsham, 2017).

I argue that we now have had enough time that some of the shortcomings of seeing many problems as requiring ICT as the solution, especially from practitioners who would come from outside, drop in, deploy and then leave is very much akin to what is happening in the Data Science world currently and needs change (Shilton et al., 2021). There may be differences, chief among them, familiarity with what ICT is and less familiar with what Data Science, Artificial Intelligence or Machine Learning are (Osoba & Welser, 2017). Basically, Data Science researchers and practitioners are just seen as magicians you throw a problem and data at, and a solution arrives on the other side. We see this with the advent of touting of 4IR strategies for African nations that are driven by public institutions that do not have the skills or knowledge to really engage with the subject they are touting as a solution to many of the problems they face (McBride et al., 2018; Moorosi et al., 2017). In ICT4D, a historical debate was on the efficacy of having researchers and practitioners who were not locals come in with "solutions" using ICT to many development issues (Andrade & Urquhart, 2012; Toyama, 2015). Over time this has become an area of study within the field itself. It became very apparent on how the development and design of systems should be participatory (Andrade & Urquhart, 2012; Tongia & Subrahmanian, 2006; Toyama, 2015) and take into account more than just the technical challenge. This tough challenge took time and many failures. In contrast within Data Science and Artificial Intelligence field, a lot of work has been put into understanding fairness, ethics and the longer term effects of the technical interventions. This is a welcome change to the ICT4D history, but we still are lagging in the understanding of the need of participatory design as well as governance that guides the field (Singh & Flyverbom, 2016). We have large international bodies like the International Telecommunications Union, that many states belong to, that has shaped ICT policies across regions.

In Artificial Intelligence, one can say the debate on fairness and harm has been very much open due to the threats of wide scale impact on people. But, this does not mean that debates solve the problems. In most of the debates and discussions, it is mostly researchers, and not decision and policy makers who are doing work to document harm and make recommendations to mitigate it (Whittaker et al., 2018). Policy makers need to come to the table to also shape the debate by providing input from government. We need to draw from lessons of other fields while at the same time understanding the uniqueness of the take up of data-driven products before we even had the time to think about their impact.

7.8 Conclusion

In this paper, I used a survey of literature around Data Science and Data Governance to bring to the fore the connections within this nexus. Leaving decisions of design to only the Data Scientist ignores the many human factors that data-driven products have. As such, Data Governance is key to being able to create and deploy products that do add to the developing economies on the continent while mitigating harm. This requires that African countries have an appreciation of the needs of governance and skills to enable effective policy. The case study presented on ICT4D allows us to learn from a related discipline that has been active for two decades and has had similar challenges in deploying interventions in the Global South.

Recommendations:

There is a need for African governments to work together to practically implement Data Governance policy. The glaring reality that only 8 countries (as of this writing) have ratified the African Union Convention on Cyber Security and Personal Data leaves much to be desired.

Both public and private industries must engage with data scientists to better get an understanding of the areas of concern highlighted in this paper beyond data privacy. Most policy on the continent focuses on privacy protections and some automated decision making, but there are many other decisions made in the process of developing data tools that impact the final outcome.

For the data scientist, it must be a reality that policy and development of data tools go hand in hand. Even if national, regional or continental policies have not caught up, there is growing movement within our practice that works to develop best practice and also highlight challenges in ethics, fairness and mitigating abuse.

References

A hostile environment. (2021). *Brazilian scientists face rising attacks from Bolsonaro's regime*. ScienceMag.

Abraham, R., Schneider, J., & Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management, 49*(2019), 424–438.

African Union. (2014). African Union convention on cyber security and personal data protection. *African Union: Addis Ababa, Ethiopia*.

African Union. (2020). The digital transformation strategy for Africa (2020–2030). *Addis Ababa*.

Alamo, T., Reina, D. G., Mammarella, M., & Abella, A. (2020). Covid-19: Open-data resources for monitoring, modeling, and forecasting the epidemic. *Electronics, 9*(5), 827.

Alpaydin, E. (2020). *Introduction to machine learning*. MIT Press.

Andrade, A. D., & Urquhart, C. (2012). Unveiling the modernity bias: A critical examination of the politics of ICT4D. *Information Technology for Development, 18*(4), 281–292.

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers, 23*(1), 4–17.

Bender, E. M., & Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling betterscience. *Transactions of the Association for Computational Linguistics, 6*(2018), 587–604.

Best, M. L. (2014). The internet that Facebook built. *Communications of the ACM, 57*(12), 21–23.

Birhane, A. (2020). Algorithmic colonization of Africa. *SCRIPTed, 17*, 389.

Birhane, A., & Cummins, F. (2019). *Algorithmic injustices: Towards a relational ethics*. arXiv preprint arXiv:1912.07376.

Birhane, A., Kalluri, P., Card, D., Agnew, W., Dotan, R., & Bao, M. (2021). *The values encoded in machine learning research*. arXiv preprint arXiv:2106.15590.

Birhane, A., Uday Prabhu, V., & Kahembwe, E. (2021). *Multimodal datasets: misogyny, pornography, and malignant stereotypes*. arXiv preprint arXiv:2110.01963.

Bond, P., & Tait, A. (1997). The failure of housing policy in post-apartheid South Africa. In *Urban forum* (Vol. 8, pp. 19–41). Springer.

Borgesius, F. Z., Gray, J., & van Eechoud, M. (2015). Open data, privacy, and fair information principles: Towards a balancing framework. *Berkeley Technology Law Journal, 30*(3), 2073–2131.

Bradshaw, D., Dorrington, R. E., Laubscher, R., Moultrie, T. A., & Groenewald, P. (2021). Tracking mortality in near to real time provides essential information about the impact of the COVID-19 pandemic in South Africa in 2020. *South African Medical Journal, 111*(8), 732–740.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77–91). PMLR.

Coleman, D. (2018). Digital colonialism: The 21st century scramble for Africa through the extraction and control of user data and the limitations of data protection laws. *Michigan Journal of Race and Law, 24*, 417.

Curtis, S. (2019). Digital transformation—the silver bullet to public service improvement? *Public Money & Management, 39*(5), 322–324.

Davis, T. (2021). *Data protection in Africa: A look at OGP member progress (August 2021)*. Technical Report. Alt Advisory.

Dhar, V. (2013). Data science and prediction. *Communications of the ACM, 56*(12), 64–73.

European Commission. (n.d.). *2018 reform of EU data protection rules*. European Commission. https://ec.europa.eu/commission/sites/betapolitical/files/data-protection-factsheet-changes_en.pdf

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H. III, & Crawford, K. (2018). *Datasheets for datasets*. arXiv preprint arXiv:1803.09010.

Gillwald, A., & van der Spuy, A. (2019). *The governance of global digital public goods: Not just a crisis for Africa*. GigaNet.

Government of South Africa. (n.d.). *Protection of personal information Act 4 of 2013*. Government of South Africa. https://www.gov.za/documents/protection-personal-information-act

Hirsch, D. D. (2013). The glass house effect: Big Data, the new oil, and the power of analogy. *Maine Law Review, 66*, 373.

Holstein, K., Vaughan, J. W., Daumé, H. III, Dudik, M., & Wallach, H. (2019). Improving fairness in machine learning systems: What do industry practitioners need?. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–16).

Hooker, S. (2021). Moving beyond "algorithmic bias is a data problem". *Patterns, 2*(4), 100241.

Hooker, S., Moorosi, N., Clark, G., Bengio, S., & Denton, E. (2020). *Characterising bias in compressed models*. arXiv preprint arXiv:2010.03058.

Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. *Big Data & Society, 8*(1), 2053951720982012.

Jensen, K. (2012). *CRISP-DM process diagram*. https://commons.wikimedia.org/wiki/File:CRISP-DM_Process_Diagram.png

Jiang, Z., Zhang, C., Talwar, K., & Mozer, M. C. (2020). *Characterizing structural regularities of labeled data in overparameterized models*. arXiv preprint arXiv:2002.03206.

Jo, E. S., & Gebru, T. (2020). Lessons from archives: Strategies for collecting sociocultural data in machine learning. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 306–316).

Kearns, M., & Roth, A. (n.d.). Ethical algorithm design should guide technology regulation. *The Brookings Institution*. https://www.brookings.edu/research/ethical-algorithm-design-should-guide-technology-regulation/

Ledford, H. (2019). Millions of black people affected by racial bias in health-care algorithms. *Nature, 574*(7780), 608–610.

Lee, M. S. A., & Singh, J. (2021). Risk identification questionnaire for detecting unintended bias in the machine learning development lifecycle. In *Proceedings of the 2021 AAAI/ACM conference on AI, ethics, and society* (pp. 704–714).

Mandl, K. D., & Kohane, I. S. (2016). Time for a patient-driven health information economy? *New England Journal of Medicine, 374*(3), 205–208.

Mbow, M., Lell, B., Jochems, S. P., Cisse, B., Mboup, S., Dewals, B. G., Jaye, A., Dieye, A., & Yazdanbakhsh, M. (2020). COVID-19 in Africa: Dampening the storm? *Science, 369*(6504), 624–626.

McBride, V., Venugopal, R., Hoosain, M., Chingozha, T., & Govender, K. (2018). The potential of astronomy for socioeconomic development in Africa. *Nature Astronomy, 2*(7), 511–514.

McCown, F., & Nelson, M. L. (2009). What happens when Facebook is gone?. In *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries* (pp. 251–254).

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR), 54*(6), 1–35.

Metcalf, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society, 3*(1), 2053951716650211.

Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2018). *Prediction-based decisions and fairness: A catalogue of choices, assumptions, and definitions.* arXiv preprint arXiv:1811.07867.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency* (pp. 220–229).

Moorosi, N., Thinyane, M., & Marivate, V. (2017). A critical and systemic consideration of data for sustainable development in Africa. In *International conference on social implications of computers in developing countries* (pp. 232–241). Springer.

Mudongo, O. (2021). *Africa's expansion of AI surveillance-regional gaps and key trends.*

Nyamwena, J., & Mondliwa, P. (2020). *Policy brief 3: Data governance matter lessons for South Africa.* https://www.competition.org.za/ccred-blog-digital-industrial-policy/2020/7/28/data-governance-matters-lessons-for-south-africa

Osakwe, S., & Adeniran, A. P. (2021). *Strengthening data governance in Africa.*

Osoba, O. A., & Welser, W., IV. (2017). *An intelligence in our image: The risks of bias and errors in artificial intelligence.* Rand Corporation.

Peng, R. D., & Matsui, E. (2015). The art of data science. *A guide for anyone who works with data. Skybrude Consulting, LLC.*

Ponelis, S. R., & Holmner, M. A. (2015). *ICT in Africa: Building a better life for all*.

Purnell, N., Scheck, J., & Horwitz, J. (2021). *Facebook employees flag drug cartels and human traffickers*. The Company's Response Is Weak, Documents Show. https://www.wsj.com/articles/facebook-drug-cartels-human-traffickers-response-is-weak-documents-11631812953.

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2020). Saving face: Investigating the ethical concerns of facial recognition auditing. In *Proceedings of the AAAI/ACM conference on AI, ethics, and society* (pp. 145–151).

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 33–44).

Ray, E. L., Wattanachit, N., Niemi, J., Kanji, A. H., House, K., Cramer, E. Y., Bracher, J., Zheng, A., Yamana, T. K., & Xiong, X. et al. (2020). *Ensemble forecasts of coronavirus disease 2019 (COVID-19) in the US*. MedRXiv.

Rice, W. E. (1996). Race, gender, redlining, and the discriminatory access to loans, credit, and insurance: An historical and empirical analysis of consumers who sued lenders and insurers in federal and state courts, 1950–1995. *San Diego Law Review, 33*, 583.

Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A. I., Etmann, C., McCague, C., Beer, L., et al. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence, 3*(3), 199–217.

Runshan, F., Huang, Y., & Singh, P. V. (2021). Crowds, lending, machine, and bias. *Information Systems Research, 32*(1), 72–92.

Russell, S. J., & Norvig, P. (1995). *Artificial intelligence: A modern approach*.

Sánchez-Monedero, J., Dencik, L., & Edwards, L. (2020). What does it mean to 'solve' the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 458–468).

Sey, A., & Mudongo, O. (2021). *Case studies on AI skills capacity building and AI in workforce development in Africa*.

Shilton, K., Finn, M., & DuPont, Q. (2021). Shaping ethical computing cultures. *Communications of the ACM, 64*(11), 26–29.

Shuja, J., Alanazi, E., Alasmary, W., & Alashaikh, A. (2021). COVID-19 open source data sets: A comprehensive survey. *Applied Intelligence, 51*(3), 1296–1325.

Singh, J. P., & Flyverbom, M. (2016). Representing participation in ICT4D projects. *Telecommunications Policy, 40*(7), 692–703.

Skorup, B., & Thierer, A. (2013). Uncreative destruction: The misguided war on vertical integration in the information economy. *Federal Communications Law Journal, 65*(2), 157.

Sujan, M., Furniss, D., Grundy, K., Grundy, H., Nelson, D., Elliott, M., White, S., Habli, I., & Reynolds, N. (2019). Human factors challenges for the safe use of artificial intelligence in patient care. *BMJ Health & Care Informatics, 26*, 1.

Tang, R., & Sae-Lim, W. (2016). Data science programs in US higher education: An exploratory content analysis of program description, curriculum structure, and course focus. *Education for Information, 32*(3), 269–290.

Tongia, R., & Subrahmanian, E. (2006). Information and Communications Technology for Development (ICT4D) – A design challenge?. In *2006 International conference on information and communication technologies and development*. IEEE (pp. 243–255).

Toyama, K. (2015). *Geek heresy: Rescuing social change from the cult of technology*. Public Affairs.

Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society, 4*(2), 2053951717743530.

Vigjilenca, A. B. A. Z. I. (2020). Truth distancing? Whistleblowing as remedy to censorship during COVID-19. *European Journal of Risk Regulation, 11*(2), 375–381.

Walsham, G. (2017). ICT4D research: Reflections on history and future agenda. *Information Technology for Development, 23*(1), 18–41.

West, D., & Allen, J. (2018). *How artificial intelligence is transforming the world*. Technical Report. Brookings Institute.

Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., West, S. M., Richardson, R., Schultz, J., & Schwartz, O. (2018). *AI now report 2018*. AI Now Institute at New York University New York.

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (Vol. 1). Springer.

Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., Bonten, M. M. J., Dahly, D. L., Damen, J. A., Debray, T. P. A., et al. (2020). Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ, 369*, m1328.

Zhang, Y.-C. (2017). The information economy. In *Non-equilibrium social science and policy* (pp. 149–158). Springer.

Zhang, J., & Barr, M. (2021). Harmoniously denied: COVID-19 and the latent effects of censorship. *Surveillance & Society, 19*(3), 389–402.

Zhang, B. H., Lemoine, B., & Mitchell, M. (2018). Mitigating unwanted biases with adversarial learning. In *Proceedings of the 2018 AAAI/ACM conference on AI, ethics, and society* (pp. 335–340).

CHAPTER 8

The Economics of Blockchain Within Africa

Aaron Thegeya

8.1 Introduction

The advent of blockchain technology offers massive potential for revolutionary innovations that address fundamental constraints and market failures across a wide span of sectors in Africa. Blockchain applications have the potential to overcome information asymmetry problems as well as property rights and governance barriers, and applications within the continent have already enjoyed a measure of success in doing so. In turn, these innovations have the ability to boost levels of productivity and unlock capital flows to underserved sectors, in addition to leveraging the increasing returns of information as an input to production in order to spur economic growth.

Blockchain adoption in Africa is supported by high rates of mobile phone adoption, but hindered by low rates of data access, low quality of Internet connections, and high data access costs. Access to data is therefore inequitable, with higher levels of access amongst better-off households (World Bank, 2021). The COVID-19 pandemic has contributed to an acceleration of digitization across the continent with the movement of more social interactions and commercial transactions online. This in turn

A. Thegeya (✉)
Aliquot Limited, Nairobi, Kenya

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8_8

has increased the need for solutions that allow secure and transparent economic activity. The closure of critical institutions, particularly education facilities, in addition to the need for continued economic activity, has exposed gaps in the quality of Internet infrastructure.

Blockchains enable transparency within decentralized networks by eliminating the need for a centralized intermediary. Within an African context, the use of blockchains can potentially be very impactful within informal labor markets by providing a mechanism to increase transparency and secure property rights, thereby increasing allocational efficiency and overcoming credit rationing constraints. The utilization of smart contacts can further underpin the scope of impact. However, the use of blockchains is not without cost. The implementation of decentralized networks introduces the need for artificial costs of verification, which create new constraints.

Policy implications of blockchains, particularly within an African context, touch on some critical dimensions. An enabling and conducive policy and regulatory framework is critical to realize the untapped potential of blockchains. These frameworks must be incentive-compatible and therefore internalize the incentives of agents within the networks. This is particularly important within the public sector, where a failure to implement applications with potentially massive positive externalities may otherwise occur. Additionally, regulatory frameworks should account for the individual data privacy preferences of agents, where the assignment of rights and obligations over data is a key determinant of growth and equity. A Pareto-efficient outcome may be possible through the assignment of rights to consumers, complemented with the development of markets for the sale of data.

The economics of blockchain is still very much a nascent field. This paper gives a background of recent advances in blockchain within Africa, as well as the key underlying economic principles within blockchain networks, both universally and within an African context. It also discusses policies that would be impactful in increasing the use of blockchain to improve governance and transparency. The document is structured as follows: Section 8.2 gives a background of blockchain innovations in Africa, and Sect. 8.3 gives a review of the existing literature on the economics of blockchain. Section 8.4 discusses the economic theory of blockchains, and Sect. 8.5 discusses economic aspects particularly relevant within an African context. Section 8.6 discusses policy implications of blockchain networks, and Sect. 8.7 concludes.

8.2 Blockchain Innovations in Africa

Blockchain is a relatively new technology that has seen rapid global growth since its first commercial implementation in 2008. A blockchain is a digital ledger of transactions that is duplicated and distributed across an entire network of computer nodes. Each entry in the ledger is known as a block, and the addition of new entries is done by appending a new block to the existing set of blocks, resulting in a chain of blocks. Each block in the chain contains a number of transactions, and every time a new transaction occurs, a record of that transaction is added to every participant's ledger. The nodes within a blockchain network may be anonymous and therefore untrusted across the participants. Therefore, mechanisms such as proof of work or proof of stake are implemented to establish trust amongst the network's participants so that there is no need for a central trusted authority or clearinghouse within the network. These mechanisms work by introducing costs that discourage malicious behavior amongst network nodes.

Blockchain innovations in Africa have been most successful in the financial sector. The most widespread application of blockchain within the financial sector is in cryptocurrency trading, as well as in payments and cross-border transactions. Blockchain offers alternative payment solutions to existing payments systems and also allows cross-border remittances. With 63 percent of Africa's population unbanked,[1] there is enormous untapped potential for the further adoption of blockchain-based solutions as an alternative to traditional payment options. In addition, the adoption of mobile money payments has lowered the costs and cultural barriers to entry for digital payments solutions. Countries with high currency depreciation risk and capital controls have tended to be the first adopters of cryptocurrencies in order to protect against inflation.[2] This combination of factors has contributed to rapid blockchain technology adoption within the financial sector.

In particular, there has been a rapid rise of cryptocurrency trade on the continent, which is now well developed in Africa. Peer-to-peer payments with digital currencies have started to become an alternative to local

[1] The banked population is defined as individuals, aged 15 years and above, that have an account at a financial institution or with a mobile money service provider.

[2] AllAfrica, *How Hyperinflation is Driving Cryptocurrency Adoption in Africa*; https://allafrica.com/stories/202109100733.html.

currencies, with a number of growing blockchain African-run startups.[3] Global cryptocurrency market capitalization reached nearly US$2 trillion in the first quarter of 2021. Despite a large number of currency exchanges, the cryptocurrency market is highly segmented, with the largest ten cryptocurrencies globally accounting for roughly 82 percent of the market by volume. Bitcoin is the largest exchange by market capitalization. Due to market volatility in cryptocurrency markets, market capitalization can change significantly. As of June 1, 2013, Bitcoin's market capitalization was $522bn, relative to $224bn for Ethereum (www.coinmarketcap.com).[4]

Additionally, central banks in Africa and around the world have begun to explore the possibility of introducing central bank digital currencies (CBDCs). A CBDC is a form of virtual currency that is issued and backed by the Central Bank. Globally, 76 countries are exploring the introduction of a CBDC with 5 of the countries—Grenada, Saint Kitts and Nevis, Antigua and Barbuda, Saint Lucia, and the Bahamas—having already launched digital currencies. Fourteen other countries are in the pilot stage while three countries within Africa (Nigeria, South Africa, and Mauritius) have their CBDCs in the development stage.[5]

Other sectors that have seen a significant number of blockchain innovations within the continent are insurance, telecommunication, and agriculture, as well as in the securing of identity and property rights. The traceability aspect of blockchain has eased the complexity and opacity of processes within agricultural supply and value chains and has facilitated proof of asset ownership as a means to enabling access to resources and unblocking credit rationing, as well as allowing better exercising of property rights. The capacity to track the origins of consumer goods has also increased the visibility of local producers along supply chains and improved logistical efficiencies.

Blockchain traceability and digital identity has gained momentum in Africa as seen from a number of large blockchain innovations within Africa. For example, a scheme launched in 2018 by the Democratic Republic of Congo uses blockchain technology to monitor cobalt mining in the country. This process involves organizations throughout the supply chain, from

[3] African blockchain startups include Kenya's BitPesa and Bitsoko, Ghana's bitcoin exchange BTCGhana, and South Africa's Luno, Ice3X, and GeoPay.

[4] See https://coinmarketcap.com.

[5] Atlantic Council Research: Central Bank Digital Currency (CBDC) Tracker https://www.atlanticcouncil.org/cbdctracker/

on-the-ground monitors checking that sites are not using child labor, through the refining process to end users. An additional innovation utilizing artificial intelligence and blockchain to trace coffee from Uganda to Colorado and from Ethiopia to Amsterdam while cutting intermediaries out of the coffee supply chain. The blockchain platform also ensures instantaneous payments to producers. Additionally, De Beers of South Africa imprints a digital fingerprint into its diamonds that is then tracked by blockchain as gems are sold, giving a forgery-proof record of the movement from the mine to cutter and polisher, then through to a jeweler.

Leading innovators in blockchain technology are Nigeria, South Africa, and Kenya. These countries account for over 80 percent of blockchain innovations. Within these countries, the major innovations are in finance and insurance, the Internet and telecommunications, and the health sector. The key innovations by sector in each country are given in Fig. 8.1.

Africa's blockchain footprint is small and accounts for a small portion of the global blockchain network. For example, if the Bitcoin network is considered, there are 12,867 nodes globally, of which only 16 nodes are in Africa, representing less than 1 percent of the total. The distribution of

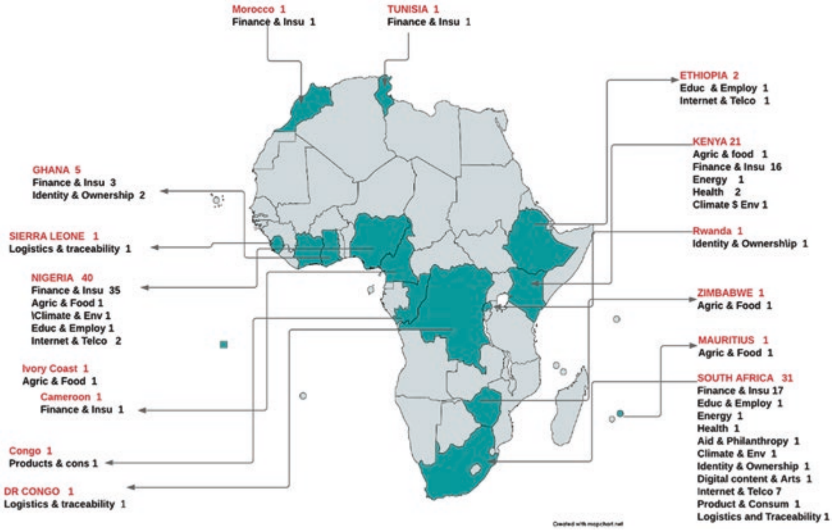


Fig. 8.1 Blockchain innovations in Africa by sector. Source: Positive blockchain.io and Internet searches

nodes is across South Africa, which accounts for 14 nodes, and Nigeria and Egypt, which each account for one node.[6] The largest share of Bitcoin nodes is found in Europe, which has a total of 4258 nodes.

While there has been significant growth in blockchain adoption, there is great latent potential in a number of areas related to governance and transparency. While the benefits of deploying blockchain technology are clear, the implementation of solutions has been slow and fragmented. In the area of public spending and governance, blockchain-based workflow tools allow for efficient project implementation by enabling expenditure tracking in a collaborative and transparent way. The deployment of blockchain technology has also increased the transparency of property rights and transfers of asset ownership, particularly related to land registration. This addresses challenges in many developing countries, where a large share of land owners lack accurate documentation of property ownership. Additionally, in order to increase trust in educational certificates, blockchain-based systems are being deployed for the verification of digital documents. This helps to establish trust in labor markets.

Digital Identity is a key innovation on which the success of some critical blockchain innovations rests. Of digital technologies, biometrics technology is the most common means of identity authentication. In the public sector, governments in almost 50 African nations have issued e-passports.[7] However, some countries are looking into deploying the technology to other sectors including health and security. The African biometrics industry is estimated at US$1.6[8] while globally it's estimated at US$24.1 billion with Africa and Middle East biometrics market forecasted to grow at an annual rate of 21 percent.[9] Digital identity innovations have been implemented to provide trusted identities for economically excluded individuals and small businesses, as well as to enhance the security of transactions through digital authentication.

The World Bank estimates that one billion people worldwide lack the means to prove their identity, and approximately 51 percent of this population live in Africa (GSMA).[10] Within the African continent, 33 percent of the population lack legal proof of identity with more than half of the

[6] Total reachable Bitcoin nodes that were active as of August 6, 2021; https://bitnodes.io/

[7] Un.org:African countries embracing biometrics, digital IDs https://www.un.org/africarenewal/magazine/february-2021/african-countries-embracing-biometrics-digital-ids

[8] As estimated by Acuity Market Intelligence.

[9] Biometrics—Global Market Trajectory & Analytics 2020.

[10] GSMA, The commonwealth Digital Identity Initiative.

population in nine African countries being unregistered.[11] Therefore, leveraging mobile technology, especially in developing countries, is seen as an enabler of digital identity and associated services, given that a large percentage of countries around the world require mandatory prepaid SIM card registration.

Enabling infrastructure is also critical for the implementation and utilization of blockchain innovations. At the end of 2019, 651 million people were connected to mobile services in Africa, comprised of 477 million people in sub-Saharan Africa (45 percent of the population) and 176 million people in North Africa (70 percent of the population).[12] Half of the total connections are through smartphones, as cheaper devices have become available, with the number of smartphone connections projected to reach 67 percent of Africa's population by the end of 2025 (65 percent of the population in sub-Saharan Africa and 75 percent of the population in North Africa respectively).

Despite high levels of mobile phone ownership, levels of data accessibility across the continent are low. Additionally, there are major differences in Internet connectivity both across countries in Africa, as well as within countries. On average, only 24 percent of individuals are connected to the Internet in Africa, which is 60 percentage points lower than the proportion of individuals connected in Europe which has the highest level of Internet connectivity. Morocco has the highest connectivity rates in Africa, with a connectivity rate of 62 percent. Conversely, a total of 11 countries have connectivity rates below 10 percent. The lowest connectivity rates across the continent are in Burundi, Somalia, and Eritrea, where less than 3 percent of the population are connected to the Internet.

Additionally, there are major differences in the costs of data access across the continent. A standard metric to gauge the relative cost of access to data is the average price of one gigabyte of data. Africa has the second-highest cost of data access globally, second only to North America, with a cost of US$5.80 per gigabyte of mobile data. Comparatively, the price of 1 gigabyte of data is US$8.21 in North America, and US$1.79 in Asia, which has the lowest average cost. Costs of broadband access are even higher. In Africa, the average price of a broadband Internet connection is US$77, excluding Mauritania, which is an outlier with an average cost of

[11] The Equatorial Guinea, Chad, South Sudan, Zambia, Angola, Ethiopia, Eritrea, Nigeria, and Somalia.

[12] GSMA.

US$695.[13] In comparison, the average cost of broadband access in the United States is US$60, while the cost of access in Europe, the lowest globally, is US$30.

8.3 Economics of Blockchain: A Literature Review

A nascent but growing literature on the economics of blockchain is beginning to emerge. Blockchain networks have brought attention to some areas of economics that have a long history of research. These include the economics of information, and more recently, the economics of data, which has been spurred by advances in processing power, machine learning, and storage. Additionally, the economics of privacy is a relevant topic within the context of blockchains.

Catalini and Gans (2016) discuss how blockchain technology can shape innovation and competition in digital platforms. In their discussion, two key costs affected by the technology are considered: the cost of verification and the cost of networking. By reducing the costs of running decentralized networks of exchange, blockchain technology allows for the creation of ecosystems where the benefits from network effects and shared digital infrastructure do not come at the cost of increased market power and restricted data access by platform operators. Reduction in the cost of networking allows open-source projects and startups to directly compete with entrenched incumbents through the design of platforms.

Abadi and Brunnermeier (2018) note that although blockchains keep track of ownership transfers, a centralized record-keeping best compliments the enforcement of possession rights. They note that the ideal qualities of any record keeping would be correctness, decentralization, and cost efficiency. However, blockchain fails to satisfy all three properties simultaneously. Unlike centralized record-keepers that extract rents due to monopoly power based on restricted access to their ledgers, blockchains allow for free entry of record-keepers and thus drive down rents. Blockchains provide static incentives for correctness through expensive proof-of-work algorithms and giving permission to record-keepers to undo fraudulent reports by rolling back history.

Jones (2020) discusses the non-rival property of data and argues that because of its infinite usability, there are large social gains to allocations in

[13] How do U.S. Internet costs compare to the rest of the world? See: https://www.broadbandsearch.net/blog/internet-costs-compared-worldwide

which the same data is used by multiple firms simultaneously. However, firms are incentivized to hoard data in order to avoid competition, leading to inefficiencies and lower productivity. The paper indicates that giving data property rights to consumers can lead to allocations that are close to optimal and is more efficient than government-imposed restrictions on selling data.

Gans and Gandal (2019) discuss the economic limits of blockchain and find that an economically sustainable network will involve the same cost of running a network regardless of whether it is proof of work or proof of stake. Additionally, the authors find that the regulation of the number of nodes permitted within a network does not lead to additional cost savings relative to networks in which free entry is allowed.

Chen et al. (2021) examine recent research on the economics of blockchains. They opine that a game-theoretical approach to understanding consensus challenges is the most promising. Further, the authors find that solutions to blockchain challenges that involve local consensus, local centralization, or local scalability are the most promising. The authors also find that blockchain innovations are likely to emanate from mechanism design approaches to consensus protocols that have clear objectives for specific applications. The core of blockchain economics, according to the paper, are agent and incentive issues, and key problems are thus likely to arise from information asymmetry.

There is an expansive literature on cryptocurrency as a subset of blockchain applications. Some key papers include: Chiu and Koeppl (2017) examine the optimal design of cryptocurrencies to assess whether they can support bilateral trade, and find that adopting an optimal design based on money growth instead of transaction fees to finance mining rewards, could lower the welfare loss caused by cryptocurrency mining. Zimmerman (2020) analyzes the model of cryptocurrency price formation and concludes that speculation could crowd out monetary use, limiting the ability of cryptocurrencies to act as a medium of payment. Halaburda et al. (2020) conduct a literature review of studies spanning different disciplines, focusing on the emergence of cryptocurrency, its demand, supply, trading price, and competition.

The economics of blockchain builds upon the extensive literature on the economics of information. Within this literature, agents interact in economic environments characterized by perfect or imperfect information. In their famous paper, Rothschild and Stiglitz (1978) showed that imperfect information caused market disequilibrium in the insurance market. In addition, Akerlof's seminal article (1978) on quality uncertainty and market mechanisms explored adverse selection in markets where sellers are better informed than buyers about the quality of goods.

The dimension of privacy is a key element of consideration within blockchain networks. Acquisti et al. (2016) highlight three themes that connect diverse insights on the economics of privacy: first, they highlight that there are situations where the protection of privacy can both enhance and detract from individual and societal welfare. Second, consumers' ability to make informed decisions about their privacy is hindered because of imperfect or asymmetric information, especially in digital economies. And third, they note that because privacy issues of economic relevance arise in widely diverse contexts, it is hard to characterize a single unifying economic theory on privacy.

Cecere et al. (2017) reviewed literature on the importance of personal data in markets. The authors highlight the puzzle that individuals face in sharing data in order to enable access to customized products and information, while at the same time protecting their personal information from misuse. Further, Cecere et al. (2017) elaborate that personal data can spur growth in new industries, but highlight that a tension exists within regulatory frameworks that aim to protect personal data without hindering firms' ability to innovate.

Chellappa and Sin (2005) highlight that a consumer's intent to use personalization services is positively influenced by their trust in the vendor. However, they note that although online vendors offer useful web-based personalization products, these products increase switching costs and are an important means of acquiring valuable customer information. They note that investments in online personalization may be severely undermined by privacy concerns.

8.4 Fundamentals of Blockchain Economics

Blockchains are immutable records of information that encode some form of economic activity, such as the value of an asset or evidence of a transaction. Hence, the economics of blockchains builds upon the foundations of the economics of information. Data can be defined as a collection of symbols that encode the properties of observables or the representation of facts, while information can be defined as data within a given context.[14]

[14] See Russell Ackoff, Data, Information, Knowledge and Wisdom (DIKW) pyramid.

Ideas in turn can be defined as novel processes that generate economic value from data.

While there is a rich economic literature on the economics of information, and more recently an expanding literature on the economics of data spurred by advances in data storage that generate big data, as well as machine learning that allows leveraging of big data in new productive ways, a distinction is typically rarely made between data and information. For our purposes, we define information as the sum-total of data and ideas and focus on the dynamics of information within blockchains.

Information is an input to production in addition to capital and labor that typically leads to an increase in productivity. Returns to information are recorded as gains in total factor productivity and in practice are recorded as a residual in a growth decomposition. Given that information builds upon existing information, it is likely that returns to information are increasing in scale (Romer, 1990). For example, binary representation of information leads to the creation of a programming language, which in turn leads to the creation of programming libraries that significantly increase the efficiency of coding and therefore production. It is thus possible that productivity gains to information are not linear but exponential. Measurement of the contribution of information to growth is challenging due to the difficulty in quantifying information in terms of level and increase, unlike other factors of production. For example, unlike information, labor input can be measured in terms of hours worked and capital can be measured by asset value.

Productivity gains from information are a function of the ability to generate and utilize ideas, which is determined by the quality of human capital. The maximization of gains from information is therefore cross-sectoral and must be integrated across other key social functions. For example, this stresses the need to protect human capacity to learn, through access to healthcare systems that protect learning capacity all the way from conception, with proper maternal nutrition, to early stage childhood development that prevents stunting of growth and brain development, to access to quality educational systems that develop the capacity of human capital to create and use ideas. Network infrastructure is then critical to ensure the dissemination of information.

Information is non-rival and partially excludable. Hence, there are massive positive externalities to the sharing of information, as information is infinitely usable at the same moment in time and without depletion over time. The excludability of information is a function of the security of storage and transmission, as well as the ability of others to generate the same information independently. This makes information only partially excludable based on the likelihood of circumventing network security systems, keeping data secure on stand-alone password-protected systems, or the ability of other agents to generate the same information from first principles. The costs of information vary between data and ideas. While the costs of data relate only to its storage, the costs of generating ideas include the necessary investment to learn.

The excludability of information generates incentives for those who collect big data to overinvest in data collection and to hoard their data, while at the same time incentivizing under-investment in data privacy, thereby generating negative externalities. By virtue of the productivity gains that accrue from the utilization of information, agents with access to information enjoy a competitive advantage over those without access. Additionally, the information has resale value in itself. These features allow those in control of big datasets to generate economic rents and therefore create incentives for them not only to overinvest in collecting big datasets, but also to hoard their data. Conversely, these dynamics also create incentives to underinvest in data privacy, as those collecting data do not internalize the costs of privacy that data subjects may place a high value on.

The excludability of information within blockchain networks is dependent on the type of network. Public blockchain networks are fully transparent by design, and hence data within these networks is not excludable. Transparency is enabled by imposing an artificial cost of proving the validity of data within a distributed network, through methods such as proof of work, and replaces the mechanism of trust in a centralized information repository. Within a centralized system, an agent in a privileged position is trusted to verify and maintain the validity of data, and the agent's status is dependent of ensuring that trust is maintained. This privileged status also allows the agent to extract rents.

Private blockchains are an intermediate case where information is shared amongst a limited number of pre-verified nodes, where participants only join by invitation or permission. Therefore, a measure of trust is inherent in these networks. Information is not excludable for participants within the network, but it is excludable from those not part of the network.

By virtue of their transparency, public blockchains lower the costs of verification without need for a central intermediary and improve allocational efficiency. Further, these networks lower the ability of those with privileged status in centralized networks to extract rents, or those with incentives to hoard data to extract rents, by excluding others from access to data. However, data collection and hoarding incentives also serve as disincentives for big data collectors to utilize blockchains.

Blockchains have significant implications within the context of the economics of information. A rich literature discusses markets under perfect information as well as those under imperfect information. Some key papers are included in the literature review. Imperfect information results from adverse selection, due to the inability to tell the quality of a given agent, and due to moral hazard, due to the inability to ensure that agents apply a desired level of effort to complete projects. Within the context of contracting, investors typically require higher returns as a consequence of these informational problems, or agents incur costs to signal their quality. In instances where information constraints are not overcome, a suboptimal level of investment occurs, leading to allocational inefficiencies within the economy.

Blockchains drive the ability to overcome imperfect information constraints and therefore increase allocational efficiency. Data-driven machine learning algorithms have improved the capacity to gauge quality by leveraging an agent's past transaction history. Through use of this historical information, an agent's preferences, action tendencies, or quality may be uncovered. For example, transaction history in an online store will give an indication of a person's preferences, while loan history records will indicate a firm's likelihood of defaulting on a future loan. However, electronic data are not immutable, and malicious agents can either change a transaction history or create a false history. Blockchains, by tracking transaction history, implementing hashing and requiring verification, lower the likelihood of data manipulation.

Fundamentally, blockchains have the ability to form the basis of smart contracts whereby the fulfillment of obligations by contracting parties is embedded and executed automatically, thus ensuring contract enforceability. Smart contracts are enabled within blockchains by virtue of the ability to encode contractual obligations through code. Contractual preconditions, actions, and the timing of execution can all be embedded a priori into a contract. This in turn precludes the ability of contracting agents to renege on their promises at a future date, as their ability to do so is eliminated at the very beginning of the contract. The design of the blockchain, through the implementation in a distributed network, reinforces the integrity of contractual obligations.

However, smart contracts have limited scope in applicability. Smart contracts are enforceable when the embedded instructions can be executed electronically and are not dependent on other exogenous factors. A key challenge in the adoption of smart contracts is that parties need to rely on a trusted, technical third party to verify the accuracy of the contractual code between two contracting parties. Further, the immutability of a smart contract, once executed, complicates subsequent amendments to the contract. Additionally, smart contracts may contain unintended programming errors that could result in erroneous execution or may contain weaknesses that could be exploited by hackers.

Within a contracting perspective, blockchains raise important questions on information ownership, custody, and property rights. Data within blockchains are stored on multiple servers that may be distributed across multiple borders. Under this distributed framework, it is imperative to define clear ownership rights, custody rights, and authority over regulation. Additionally, information is frequently generated as a byproduct of economic activity. For example, the repeated purchase of a particular commodity indicates a person's preference profile and is visible to the retailer of the commodity. Clear delineation of the rights to such data is necessary. Further, in many instances, information is collected not on an individual but on a collective level, such as on a household or location basis. For example, satellite data is informative of a particular location, and is reflective of the individuals living in that location. In this case, the clear delineation of rights to collect and own such collective data is also important.

Although blockchains provide a decentralized solution to overcome challenges related to asymmetric information, they do not overcome the costs of verification. The implementation of blockchains is costly and is dependent on the use of an energy-intensive verification mechanism within a proof-of-work context. Proof-of-work requires network participants to solve a complex encryption problem in order to gain the rights to amend a given blockchain. The complexity of the encryption problem can be predetermined by the network and accordingly amended as a function of computing capacity. Proof-of-work solutions are found through brute-force algorithms that iterate through a list of possible solutions until the correct solution is found.

Proof-of-work algorithms impose new barriers to entry in place of those eliminated through the distributed nature of blockchains, through investment requirements in processing power, memory, and technical know-how. As the popularity of blockchains has grown, network nodes have invested in dedicating processing power and memory to blockchains, and as a result, have massively increased energy requirements. For example, the energy consumed by the Bitcoin network is greater than total energy consumption in Argentina.[15] The distributed nature of blockchain networks also demands high memory bandwidth within those networks that require each node to store copies of the chain. Additionally, proof-of-work demands limit the number of transactions that can be executed within a block during any given period. Within the Bitcoin network, on average seven new block transactions are executed within an hour.

Alternatives to proof-of-work have been developed in order to overcome processing constraints, including proof-of-stake, proof-of-burn, and proof-of-capacity. However, none of these alternatives provide a comprehensive solution, and this is still an area of active research. Proof-of-stake is a type of consensus mechanism that works by requiring users to stake their assets in order to become validators. The value of assets at stake determines the validation power of a given node. Validators are responsible for checking and confirming blocks they do not create and stand to lose their stake for bad behavior on the network.

Proof-of-burn is an alternate consensus algorithm, which aims to reduce energy consumption relative to proof of work. Proof of burn functions by providing a different incentive for miners to validate transactions. A network node utilizes existing assets rather than wearing out electricity and hardware, therefore encouraging miners to be good actors. For example, within a cryptocurrency context, network nodes utilize or ‘burn’ existing coins. The more coins miners send to burn, the higher their likelihood of mining a block. Moreover, miners receive a reward each time they correctly validate and mine a block. Proof of activity is a combination of proof of work and proof of stake algorithms, which works by making it more difficult to mine as time elapses. However, increasing the difficulty of mining causes a spiraling in energy consumption and hardware costs.

These methods are not without drawbacks. For example, proof-of-stake encourages hoarding due to the link between returns and the amount held in escrow, rather than spent on transactions. It may also defeat the purpose of decentralization by leading to a concentration of validation markets.

Further, while blockchain can improve transparency, it is not possible for blockchain to enforce the transfer of physical assets. Blockchain can give a transparent account of the transaction history relating to a given

[15] Digiconomist.net: Cambridge University’s Bitcoin Energy Consumption Index.

asset, and smart contracts can enforce the electronic transfer of ownership rights for a given asset, based on a set of a priori agreed to conditions. However, the scope of a blockchain contract remains electronic, and the physical transfer of an asset is dependent on the transferor honoring the promise to turn the asset over to the transferee. A well-functioning justice system is critical to support the successful implementation of blockchains and build trust in their deployment. Additionally, blockchains cannot verify the correctness of information about inputs, which require external verification before introduction into a blockchain.

Blockchains do not resolve public sector incentive compatibility constraints. While blockchain is an impactful solution, it does not address private incentives that agents face in deciding contractual obligations. Additionally, blockchains remove the ability for agents to generate rents on the basis of control of the custody and dissemination of information. The implementation is therefore likely to face resistance, which is particularly important in the case of public sector contracts, where failure to internalize the positive benefits of blockchain adoption is particularly costly, and thus the negative impact may be substantial and spread over a large number of agents. Where a blockchain is likely to lower the rents receivable to an agent that is in control of the decision of whether or not to use a blockchain, there are likely to be upfront costs in incentives or policy to push implementation through.

8.5 An African Perspective on Blockchain Economics

While blockchains have been gaining traction across a large number of economies, a number of distinguishing features are particularly relevant in the African context. These include fundamental issues ranging from the digital inclusiveness of much of the continent, to governance issues related to supporting the implementation of blockchains in an equitable manner.

Africa’s relatively lower levels of Internet access, lower quality of Internet connections, and higher cost of access mean that a smaller share of individuals participate in the data economy relative to the rest of the world. Thus, the impact of blockchain innovations is likely to be segmented at the household level, with higher levels of adoption amongst better-off households. This also means that the greatest impact of blockchain innovations is likely to be felt in those sectors that have a public good element or that have the potential for positive externalities. These sectors would have a direct benefit on those that are digitally included, while also indirectly benefitting those that are not. For example, a blockchain innovation that digitally tracks land registries is a public service that increases the transparency of transactions, and those public records are beneficial to the entire population.

Low levels of digital access across large segments of the African population also mean that the majority of Africans are unable to leverage data to increase their levels of productivity or directly benefit from the utilization of blockchains. By leaving those without access behind, blockchains may contribute to a widening gap in living standards. The development of blockchain therefore needs to be complemented with a concurrent increase in access to data.

Africa's markets are characterized by large informal labor markets. On average, 83 percent of jobs in Africa are accounted for in the informal sector, with the greatest share amongst the youth.[16] The pivotal role played by the informal sector is expected to continue for the foreseeable future, as the formal sector is not creating jobs quickly enough to absorb Africa's growing labor force. The visibility of informal sector labor is limited, and therefore asymmetric information problems are particularly acute within this sector, as it is costly to overcome adverse selection and moral hazard problems. In addition, individuals within this sector are much more likely to have a low asset base, lower levels of income, and higher income volatility than formal sector workers, lowering their ability to pledge collateral. Therefore, the flow of capital to this sector is severely restricted. The integration of blockchain to improve the visibility of the informal sector, through transparent tracing of transactions or securing of property and asset rights, would potentially have a transformative impact on the sector by reducing levels of credit rationing.

Digital currencies, especially within the cryptocurrency market, thrive highly on speculation, which in turn results in high volatility. Volatility is influenced by the fact that markets for cryptocurrencies remain very small in comparison to traditional currencies, and therefore there is a concentration of individuals with large holdings of crypto coins. For instance, there are approximately 18.9 million bitcoins in circulation in comparison to

[16] Kenya National Bureau of Statistics, *Economic Survey,* 2020, p. 44.

KES 288.2 billion and US$2.2 trillion that is circulation.[17] Large transactions by these individuals therefore have the potential to cause swings in the market.

Informal contracts and tacit agreements are common in regions where there is uncertainty about the enforceability of contracts and are typically premised on a reputational foundation or a collective governance framework at the local level. Within this context, the use of blockchain can create a framework to enhance the ability to contract. First, it can substitute the critical role of reputation through transparency, whereby visibility of past contractual obligations amongst any contracting parties is visible to third parties. Consequently, the costs of reneging on contractual obligations are significantly increased through the potential for reputational repercussions, and this creates a strong incentive for parties to adhere to contracts.

Second, the utilization of smart contracts can strongly enhance the ability to contract by automatically executing actions that have been agreed upon ex ante, once certain preconditions have been fulfilled. This in turn means that control of contract execution is taken away from the contracting parties and therefore the contract is secured. These innovations are efficiency-enhancing and are particularly relevant in countries where there are weak governance frameworks or weak rule of law.

In addition, blockchain innovations can help improve institutional governance itself, by enhancing the transparency of legal and judicial processes, thereby improving the accountability of public officials. This can help speed up the improvement of governance across the continent, where indicators highlight the need to improve the quality of institutions and enforce the rule of law.

However, private incentive compatibility constraints indicate that regulatory intervention may be necessary to introduce the use of blockchains to contribute to the solution of governance problems. Individuals in privileged positions of information or authority are frequently able to extract rents as a consequence of their status. Where the decisions to implement blockchains are also within the domain of their control, they face

[17] Total Kenya Shillings in circulation are given by the Central Bank of Kenya, depository corporation survey, https://www.centralbank.go.ke/statistics/monetary-finance-statistics/; total US currency in circulation is tracked at https://fred.stlouisfed.org/series/CURRCIR; total number of bitcoins in circulation are tracked at https://www.blockchain.com/charts/total-bitcoins.

disincentives to carry through implementation, due to the significant loss of private gain, despite the welfare-enhancing properties of blockchains. Incentive compatibility problems are found in both the private and public sectors, but the welfare implications are particularly acute in the public sector, where externalities tend to be larger. As a consequence, external intervention is necessary to ensure that implementation occurs.

Economies of scale, such as minimum account requirements and account maintenance fees, as well as the location of mainstream retail banking services in high-income neighborhoods, work against disadvantaged populations. According to Kshetri (2017), blockchain-based solutions can be used to develop offerings that are appropriate to meet the needs of disadvantaged groups such as by enabling small transactions at low cost. Kshetri also argues that the combination of decentralized access and immutability, which makes it difficult to engage in opaque transactions that take place between companies, individuals, and institutions, is likely to reduce or even eliminate fraudulent lending practices such as insider lending.

8.6 Policy Implications

While blockchains carry transformative potential for Africa, the realization of this potential is unlikely to happen without an enabling and conducive policy and regulatory framework. A number of important policy considerations must be taken into account, as detailed in this section. The large number of individuals that do not yet have a digital presence, coupled with the large variations in the quality of data access, create an opportunity to put an appropriate data infrastructure in place as new users come online.

The regulatory environment for blockchain in Africa is uneven. Currently, no regulatory authority within Africa has issued any regulations on the use of blockchain technology.[18] However, Mauritius and Kenya have created regulatory sandbox environments to provide innovators in financial institutions with licenses to practice. With regard to data privacy, 28 countries in Africa have enacted personal data protection legislation.[19]

The assignment of rights and obligations over data to various stakeholders will determine growth and equity. Assignment of rights to

[18] Baker McKenzie, *Blockchain and Cryptocurrency in Africa: A comparative summary of the reception and regulation of Blockchain and Cryptocurrency in Africa*, 2018.

[19] UNCTAD, Data Protection and Privacy Legislation Database.

consumers may provide the most optimal equilibrium if complemented with the development of markets for the sale of data. This is because under these conditions, households will appropriately reveal their individual preferences in terms of benefiting by sharing data relative to the costs of doing so and price their relative value accordingly. Additionally, the regulatory landscape must allow for sufficient transparency in order to determine the rate of data collection and data distribution. This will allow a clear mapping of the landscape in order to understand privacy concerns in addition to data utilization and productivity.

Blockchain policy and regulatory frameworks must be incentive-compatible. This means that frameworks must internalize private incentives of participants within blockchain networks. For example, the excludability of data creates incentives for agents to hoard data in order to stifle competition, which allows incumbents to benefit, but also leads to data underutilization. An appropriate framework will account for these incentives, and implement policies that promote data sharing in order to result in more optimal outcomes.

A successful blockchain policy for Africa must be an integrated policy and must also promote the development of an enabling environment, by closing access gaps in social and physical infrastructure. A clear policy to close gaps in education across and within countries is a critical foundation necessary to maximize the benefits of blockchain. Additionally, improvement of infrastructure and lowering the costs of access are necessary preconditions to increase productivity.

Data privacy is a critical component of the policy environment. Privacy has some characteristics of a final good, valued for its own sake, and an intermediate good valued for instrumental purposes. These are crucial considerations for a blockchain governance framework. As a final good, individuals gain different levels of utility from maintaining privacy, with a spectrum ranging from those who are highly sensitive to privacy to those who are very comfortable with an open digital presence.

The regulation of privacy can have some unintended effects. Individuals have incentives to disclose favorable information and hide unfavorable or negative information. This results in inefficient outcomes. Hence, regulatory interference that prohibits the flow of personal information can remove signals of quality from the marketplace and introduce inefficiencies. However, reputational effects can also lead to suboptimal private decisions if information disclosure reveals negative information about an individual. For example, in a transparent environment, the stigma associated with checking into a rehabilitation center may deter an individual from seeking treatment.

The optimal level of information disclosure should take into account accountability versus privacy concerns. The Coase Theorem (1960) states that whether or not a person's private information will be disclosed is dependent on the relative valuations of the parties interested in the information. If trade in an externality is possible and there are sufficiently low transactions costs, bargaining will lead to a Pareto-efficient outcome regardless of the initial allocation of property rights. Thus, assigning property rights to information and allowing trade in information is likely to lead to an ex-post efficient outcome.

While private blockchains provide a measure of privacy, they are not truly decentralized. This is because private blockchains delegate specific actors to verify blocks and transactions. Although this provides efficiency and security, it raises concerns that private blockchains are not truly decentralized because the verification of transactions and control are put back into the hands of a central entity. In contrast, public blockchains lower the likelihood of a malicious attack as more people become part of the network.

From a legal perspective, the enforceability of blockchain contracts may face jurisdictional issues, when it is unclear where the agents sit. Given that the nodes of a decentralized ledger can span multiple locations around the world, there is a risk that transactions performed by an organization could fall under every jurisdiction in which a node in the blockchain network is situated, resulting in an overwhelming number of laws and regulations that might apply to a certain transaction. For example, there have been difficulties in the application of existing regulatory regimes on crypto assets, whereby in some countries, regulators have banned cryptocurrencies, while others have issued varying levels of regulations and investor warnings.

Agent domicile is also important particularly from the perspective of recognizing income for taxation purposes, as well as building appropriate structures for taxation. Poor coordination in establishing taxation frameworks can result in multiple taxation, competition, or conflict across authorities in the recognition of revenues or asset gains within blockchain networks, for the purposes of establishing tax bases.

The lack of a centralized coordinator increases the vulnerability of blockchains to volatile fluctuations and limits the ability of a central authority to intervene as a consequence of the volatility. This property is particularly prevalent in cryptocurrency markets. It raises the concern that without appropriate regulation, there could potentially be welfare-reducing swings in volatility and makes the case for a regulatory authority. However, cryptocurrencies by nature have no specific legal and regulatory jurisdiction due to their global portability, making it difficult for policymakers to develop cryptocurrency regulations. Additionally, cryptocurrencies have underlying tokens that are not subject to existing regulations.

Appropriate infrastructure must implement checks and balances to avoid data-based discrimination. While data-driven innovations are often viewed as a positive development, discriminatory biases embedded in these technologies have the potential to create racial and social inequalities. Specifically, in running algorithms, data can be poorly selected, incorrect, incomplete, or outdated and can even incorporate historical biases, for example, in the case of employment selection based on previous job recruitment data. Also, given that the programming decisions are essentially human judgments, concerns regarding the design of the algorithm that is using the data inputs often arise.

8.7 Conclusion

Blockchains have grown rapidly since their first application with the invention of cryptocurrency markets a little over a decade ago, and their potential for additional applications appears endless. In this background paper, we have considered the current state of blockchains within Africa and explored the underpinning economic elements of blockchain networks. Additionally, we have explored the economic aspects of blockchains that are particularly relevant within an African context.

Policy implications of blockchains touch on some important dimensions. An enabling and conducive policy and regulatory framework is critical to realize the untapped potential of blockchains. These frameworks must be incentive-compatible, and therefore internalize the incentives of agents within the networks, as failure to implement applications with potentially massive positive externalities may otherwise occur, particularly within a public blockchain context. Additionally, the frameworks should account for the individual data privacy preferences of agents within blockchain networks, where the assignment of rights and obligations over data will be a key determinant of growth and equity. A Pareto-efficient outcome may be possible through the assignment of rights to consumers, complemented with the development of markets for the sale of data.

Blockchain research is still nascent, and the economic theory of blockchain networks is far from well established. Additional research on blockchain economics is necessary to contribute to the growing literature. These papers will incorporate an element of computer network dynamics from computer science as a matter of necessity, in terms of the efficiency, technical capacity, and deficiencies inherent in networks of anonymous users. In addition, the literature will continue to explore new economic incentives that arise out of interactions within distributed networks. For example, while blockchain networks solve the problem of establishing trust within distributed networks and maintaining accurate data within those networks, it does so at the expense of generating costly verification. While alternatives to costly verification are under experimentation, there is no clear solution, and this remains an active area of research.

References

Abadi, J., & Brunnermeier, M. (2018). *Blockchain Economics.* NBER Working Paper No. 25407.

Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature, 54*(2), 442–492.

Akerlof, G. A. (1978). The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235–251). Academic Press.

Catalini, C., & Gans, J. (2016). *Some Simple Economics of the Blockchain*. NBER Working Paper No. 22952.

Cecere, G., Le Guel, F., Manant, M., & Soulié, N. (2017). The economics of privacy. In *The New Palgrave dictionary of economics.* Palgrave Macmillan.

Chellappa, R. K., & Sin, G. R. (2005). Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information Technology and Management, 6*(2), 181–202.

Chen, L., Cong, L. W., & Xiao, Y. (2021). A brief introduction to blockchain economics. In *Information for efficient decision making: Big data, blockchain and relevance* (pp. 1–40).

Chiu, J., & Koeppl, T. V. (2017). *The economics of cryptocurrencies–bitcoin and beyond*. Working Paper 3,048,124, SSRN.

Gans, J. S., & Gandal, N. (2019). *More (or less) economic limits of the blockchain* (No. w26534). National Bureau of Economic Research.

Halaburda, H., Haeringer, G., Gans, J. S., & Gandal, N. (2020). *The microeconomics of cryptocurrencies* (No. w27477). National Bureau of Economic Research.

Jones, C. I., & Tonetti, C. (2020). *Nonrivalry and the Economics of Data American Economic Review 110*(9) 2819–2858.

Kshetri, N. (2017). Potential roles of blockchain in fighting poverty and reducing financial exclusion in the global south. *Journal of Global Information Technology Management, 20*, 201–204.

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy, 98*(5), S71–S102.

Rothschild, M., & Stiglitz, J. (1978). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. In *Uncertainty in economics* (pp. 257–280). Academic Press.

World Bank. (2021). *Minimum data consumption: How much is needed to support online activities, and is it affordable? Analytical insights*. World Bank.

Zimmerman, P. (2020). *Blockchain structure and cryptocurrency prices*. Bank of England Staff Working Paper No. 855.

CHAPTER 9

Conclusion

Bitange Ndemo, Njuguna Ndung'u,
Scholastica Odhiambo, Miriam Omolo, and Abebe Shimeles

9.1 Introduction

The book has presented the views of various authors' findings on data governance frameworks in Africa. It is a primer in creating a platform for raising awareness on basic principles/tenets of international norms on data

B. Ndemo
Kenya's Ambassador, Belgium and the EU Mission, Brussels, Belgium
e-mail: bndemo@bitangendemo.me

N. Ndung'u
National Treasury and Economic Planning, Nairobi, Kenya
e-mail: njuguna.ndung'u@aercafrica.org

S. Odhiambo (✉)
African Economic Research Consortium, Nairobi, Kenya
e-mail: scholastica.odhiambo@aercafrica.org

M. Omolo
The African Policy Research Institute, Nairobi, Kenya
e-mail: miriamomolo@yahoo.com

A. Shimeles
Department of Economics, University of Capetown, Capetown, South Africa
e-mail: abebe.shimeles@uct.ac.za

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8_9

governance. The experiences and practices shared from across the globe and Africa in particular, will effectively promote and advocate for data governance protocols in the era of digital revolution. The assessment of the potentials of improving the digital market to enhance benefits to African consumers, governments, and businesses, is well articulated in each of the chapters. This chapter draws conclusions building from the findings of the various authors and further provides recommendations on data production, consumption, and utilization.

9.2 Data Governance Frameworks

Data governance frameworks require a delicate balance between data sovereignty, increased productivity, and accountability. While data sharing is expected to improve living standards through improvements in productivity growth due to non-rivalry and infinite reusability of data, leading to increase in scale, there is need to minimize data misuse and exploitation to ensure data safety. Policy makers can play a role in maintaining the balance between data safety and improving lives by putting in place a clearly defined data governance framework integrated with a data strategy that improves data sovereignty, strengthens Africa's competitiveness, and cross-country collaboration in this digital age. An effective data governance framework should incorporate infrastructure, technical protocols, laws and regulations, and institutions that promote secure use of data and enhancing the rights to privacy.

To promote regional connectivity, a pan-African data governance infrastructure that brings about a single market for data that allows for creation, use, and reuse of data by individuals across the continent and spurring economic growth and development should be out in place. The framework must adhere to the principles of accountability, data accuracy and quality, and interoperability and standardization. The costs of compliance to such a framework should be low to promote equity in access and competition.

9.3 Value Chain Approach to Data Production, Use, and Governance

A value chain approach to data production, use, and governance in Africa was found to be of low quality with respect to a wide range of issues including relevance, impartiality, and equal access; professionalism; accountability; prevention of misuse; cost-effectiveness; confidentiality; legislation;

national co-coordination; international coordination and cooperation and documentation. This has been driven by a myriad of challenges but key among them, the inability of African National statistical offices (NSO) to work effectively due to lack of political and institutional independence, which has weakened the managerial and technical ability to generate quality of official statistics. Evidence shows that legal autonomy enhances the trustworthiness of NSOs thus, enabling them to attract and sustain funding from both domestic sources and development partners. Other challenges include weak basic statistical data systems, low use of non-traditional sources of data, lack of a comprehensive data user's engagement strategy, weak coordination of the NSS, inadequate data dissemination systems, lack of a strategy to enhance uptake of official statistics, inadequate funding, and inadequate human resources.

Improving production and use of official statistics as well as governance of official statistics processes, requires that the legal framework governing production of official statistics should ensure that produced statistics are accurate, relevant, timely, comparable, consistent, and impartial in line with the fundamental principles of official statistics as well as African Charter on Statistics. Secondly, enhancing the capacity of NSOs to coordinate the NSS and formation of partnerships with other data producers outside NSS to supplement official statistics, will result in efficient and effective administrative of data systems and will reduce data gaps. African NSOs should develop and enhance the capacity of their staff through hiring addition staff and providing technical training to enhance their technical skills for national statistics data collection and management. Institutionalize knowledge management systems for transfer of managerial and technical skills from experienced staff to newly employed employees and ensure smooth running of operations and strengthen institutional memory.

9.4 Legal and Regulatory Frameworks for Data Protection

Data protection through legal regimes and regulations to ensure free flow of data, open data regimes, and cyber security found that management of big data would always involve the handling of personal data, hence the activation of data protection principles. Very few African countries have binding instruments with provisions on various principles that persuasively impart data governance. Out of 55 African countries, only 30 have fully dedicated data protection laws, and 19 of them have established DPAs to enforce compliance with the laws, this means that there is a weak data legislative and enforcement framework.

There is a lack of trust and confidence amongst African states in developing a unified regulatory system, because of the high value personal data held within each territory. It is imperative that a trusted data environment grounded in the rule of law; comprehensive institutional arrangements and regulations; and competent institutions responsible for overseeing the use of public and private data is established as soon as possible. Active dialogue through multi stakeholders' consultations that involves collaboration with governments, private sector, data protection authorities with the objective of negotiating of mutual assistance agreements that will guarantee similar protection of data in contracting member states and pledges to investigate and prosecute cross-border cybercrimes comprehensively, will moderate governance of movement of cross-border data.

A comprehensive legal and regulatory framework for data governance will require that African countries ratify the Malabo Convention and strengthen their respective municipal data protection legal frameworks in a manner that compliments public and private organizational management of personal data, within and outside Africa. The transborder cooperation of national DPAs envisaged by the regional treaties ought to be encouraged and strengthened, to boost enforcement of regional and municipal data protection laws, with the aim of enhancing trans-border flow of data and international data governance within the confines of uniform cross-border data protection rules.

A legal and regulatory entry point for data protection is though the African Continental Free trade Areas Agreement (AfCFTA) and the Digital Transformation Strategy for Africa (2020–2030) that seek to increase e-commerce and digital trade in Africa, it is important to consider how supporting the free movement of data across Africa can enhance these efforts. It has been shown that cross-border data flows are instrumental and have the potential to greatly influence a new economic resurgence for the continent, as can be drawn from experiences of countries or regional bodies that have adopted a liberal approach to data regulation.

9.5 Digitalization and Financial Data Governance

Financial digitalization has been led by financial technology companies (FinTechs), this has resulted in increased productivity and more financial resources being made available and affordable. Digitalization in the financial sector is quickly evolving and this requires a financial data governance framework to deal with the challenges affecting collection, processing, quality, and security of collected data. Given the emerging trends, policy makers should pre-emptively address these challenges by continuously reviewing and enhancing the institutional frameworks that enhance the operation of FinTechs that develop financial transactions platforms and applications. A dialogue on financial-data policy and governance in Africa is essential and can be facilitated by AERC. Such a forum is the first step towards bringing together practitioners, scholars, policymakers, and institutions to lay the groundwork for the governance of financial data. A key advantage will be that it will set the pace and agenda for financial data governance research and dissemination of high-impact findings and recommendations for application across the continent.

9.6 The Economics of Blockchains in Africa

The emerging blockchain technology has the potential to boost levels of productivity and unlock capital flows to underserved sectors. This is influenced by increasing returns of information as an input to production to spur economic growth. Even though blockchain technology is gaining traction at the international level, this might not quickly permeate the African continents due to the low lower levels of Internet access, lower quality of Internet connections, and higher cost of access. This implies that blockchain innovations is likely to be segmented at the household level, with higher levels of adoption amongst better-off households. The large informal Africa Market, which is estimated to generate 83% of employment particularly for the youth, with a low asset base will result in most African countries being unable to leverage on these technologies to increase their productivity. The integration of blockchain to improve the visibility of the informal sector, through transparent tracing of transactions or securing of property and asset rights, would potentially have a transformative impact on the sector by reducing levels of credit rationing. Informal contracts and tacit agreements, which tend to thrive in informal sectors will create uncertainty given that most players in this sector rely on social capital.

These challenges can be overcome by creating an enabling and conducive policy and regulatory framework that can be used to tap the existing potential for blockchains. These frameworks must be incentive-compatible, so that agents within the networks realize the value of adherence to the frameworks. Data protection and privacy must be enhanced within such frameworks, with an option for the development of markets for the responsible sale of data. The utilization of smart contracts can strongly enhance the ability to contract by automatically executing actions that have been agreed upon ex-ante, once certain preconditions have been fulfilled, within and established legal framework.

9.7 Further Research

Several areas require further research:

1. Identification and prioritization of key strategic interests across the continent that will benefit the most from the implementation of a data governance framework, as well as quantifying the value of the framework.
2. Determining strategies to increase the pace of digitization of offline public data sources, while also increasing access to already digitized public data.
3. Mapping infrastructure gaps and cost-of-access disparities at the subnational and cross-country level.
4. The legal, policy and regulatory framework for optimal implementation of effective data governance.
5. The effects of blockchain data on businesses—micro, small, medium, and large enterprises.
6. The efficiency, technical capacity, and deficiencies inherent in networks of anonymous users, and its application within the African context.

Index[1]

[1] Note: Page numbers followed by 'n' refer to notes.

B. Ndemo et al. (eds.), *Data Governance and Policy in Africa*,
Information Technology and Global Governance,
https://doi.org/10.1007/978-3-031-24498-8