OPEN SCIENCE IN RESEARCH AND INNOVATION FOR DEVELOPMENT IN AFRICA – RESEARCH PAPER

Boulton Geoffrey ;Loucoubar Cheikh ;Mwelwa Joseph;Wafula Muliaro;Ozor Nicholas;Bolo Maurice;

;

IDRC Grant: 108460-001-Networking Africa's science granting councils

Open Science in Research and Innovation for Development in Africa

Geoffrey Boulton
University of Edinburgh, UK

Cheikh Loucoubar
Pasteur Institute, Dakar, Senegal

Joseph Mwelwa
Joint Minds Consult Institute, Gabarone, Botswana

Muliaro Wafula
Jomo Kenyatta University of Agricultural Technology, Nairobi, Kenya

Nicholas Ozor
African Technology Policy Studies Network (ATPS)

Maurice Bolo
The Scinnovent Centre

AFRICAN TECHNOLOGY POLICY STUDIES NETWORK

TheScinnovent Centre
SCIENCE, INNOVATION AND ENTERPRISE

The African Technology Policy Studies Network (ATPS) is a transdisciplinary network of researchers, policymakers, private sector actors and the civil society promoting the generation, dissemination, use and mastery of Science, Technology and Innovations (STI) for African development, environmental sustainability and global inclusion. In collaboration with like-minded institutions, ATPS provides platforms for regional and international research and knowledge sharing in order to build Africa's capabilities in STI policy research, policymaking and implementation for sustainable development.

Published by the African Technology Policy Studies Network (ATPS)

P O Box 10081, 00100 GPO Nairobi, Kenya

ISBN: 978-9966-124-93-7

Table of Contents

List of Boxes

List of Figures

List of Tables

About Science Granting Councils Initiative

The Science Granting Councils Initiative in Sub-Saharan Africa (SGCI) seeks to strengthen capacities of Science Granting Councils (SGCs) in Eastern, Southern, Central and West Africa in order to support research and evidence-based policies that will contribute to economic and social development. It is jointly funded by the United Kingdom's Department for International Development (DFID), Canada's International Development Research Centre (IDRC), South Africa's National Research Foundation (NRF) and the Swedish International Development Cooperation Agency (Sida).

The objectives of SGCI are to strengthen the ability of participating SGCs to 1) manage research; 2) design and monitor research programmes, and to formulate and implement policies based on the use of robust science, technology and innovation (STI) indicators; 3) support knowledge transfer to the private sector; and; 4) establish partnerships with one another, and with other science system actors. The implementation of these objectives is achieved through regional training courses, individualised on-site training sessions, on-line training, webinars and, collaborative research. The SGCI works with 15 councils in Kenya, Rwanda, Uganda, Tanzania, Ethiopia, Cote d'Ivoire, Burkina Faso, Senegal, Ghana, Zambia, Mozambique, Botswana, Malawi, Namibia and Zimbabwe.

The SGCIs principle output include 1) more effective research management practices among Councils, 2) strengthened ability of Councils to design and monitor research programmes, and to formulate and implement policies based on the use of robust science technology and innovation indicators, 3) increased knowledge transfer to the private sector and 4) increasingly coordinated and networked Councils. More effective Councils are expected to strengthen national science systems, and ultimately lead to nationally-led research that contributes to development in participating African countries.

About the African Technology Policy Studies Network (ATPS)

The African Technology Policy Studies Network (ATPS) is a trans-disciplinary network of researchers, policymakers, private sector actors and the civil society promoting the generation, dissemination, use and mastery of Science, Technology and Innovations (STI) for African development, environmental sustainability and global inclusion. ATPS has over 1,300 members and 3000 stakeholders in over 51 countries in 5 continents with institutional partnerships worldwide. We implement our programs through members in national chapters established in 30 countries (27 in Africa and 3 Diaspora chapters in the Australia, United States of America, and United Kingdom). In collaboration with like-minded institutions, ATPS provides platforms for regional and international research and knowledge sharing in order to build Africa's capabilities in STI policy research, policymaking and implementation for sustainable development.

About Scinnovent Centre

The Scinnovent Centre is a science, technology and innovation (STI) policy think tank registered in Kenya as a not-for-profit company. Their preliminary concern is that despite advancements in science, technology and innovation (STI), poverty levels in Africa are increasing; environment degradation is worsening; the ecosystem has become more fragile; sustainability has been compromised and livelihoods threatened.

So they ask three big questions: Why have the developments in science, technology and innovation not made any significant difference in African development? Why have STI policies not translated into practical change on the ground? How come pockets of success piloted across countries have not scaled?

Acknowledgement

This research paper was published by the African Technology Policy Studies Network (ATPS) in partnership with the Scinnovent Centre under the Science Granting Councils Initiative in sub-Saharan Africa. The ATPS and Scinnovent Centre acknowledge the funding support received from United Kingdom's Department for International Development (DFID), Canada's International Development Research Centre (IDRC), South Africa's National Research Foundation (NRF) and the Swedish International Development Cooperation Agency (Sida).

Executive Summary

Chapter 1 The first target for this paper was the 2019 Annual Forum of sub-Saharan Africa's Science Granting Councils as basis for discussion of the new Open Science paradigm, its significance for Africa, and possible roles for the Councils in promoting its development. It has developed further as a consequence of those discussions. Its substantives arguments are as follows.

Chapter 2 summarises the digital revolution of the last three decades that has seen unprecedented developments in the means whereby vast fluxes of data and information are acquired, stored, communicated. It is a world historic event with profound, global implications. Its scientific and technological importance lies in enabling a new scientific mode, of data-driven science, that contrasts with the classical mode of hypothesis-driven science. It permits the recognition of deep patterns in complex phenomena, an apparently simple process, but one that is at the heart of the digital revolution's benefit to science, to society and to development. It is also the fundamental driver of the 4^{th} industrial revolution. A strong national science and educational base is crucial in exploiting these opportunities, with the new "open science" paradigm as a means of addressing the opportunities and amplifying the impact of the collective scientific effort. The digital skill base is also a vital means of protecting key national assets and formulating relevant regulations and legislation.

Chapter 3 discusses the pervasive and efficient technologies unleashed by the digital revolution and argues that they cannot be sidestepped. Their innovations create new capabilities and reduce costs in ways that undermine and disrupt many established ways of working, in both public and private sectors, and create new opportunities for innovative application. National science systems, both in Africa and beyond, must adapt to the new demands and opportunities these technologies create, with open science as a potentially cost efficient way of doing so. The recent finalisation of agreements about the African Continental Free Trade Area (AfCFTA) and the mobility and dynamism that it is designed to enable, would be an opportunity for powerful synergy if it were matched with timely creation of an African open science area.

Chapter 4 describes the array of essential tools and processes required for the new paradigm of open science, and the rationale for sharing scientific data to permit their re-use by others as open data. It requires efficient management of research data, common standards that enable data usability and data citation, open licensing, and access to cutting edge machine learning capacity. It requires the outputs of scientific research, and the authoring of scientific papers, to be openly accessible at affordable cost; a major problematic issue for global science, and one with which Africa must engage. It must be open to society, which requires the engagement of science as never before in joint creation of actionable knowledge that has greater potential for application and greater socio-political legitimacy. Realising the potential benefits of the open science mode, and exploiting the capacities of data-driven science require access to powerful computational and cloud systems and communication networks.

In **Chapter 5** we argue that the demands on researchers, research groups, or even institutions to satisfy these varied requirements of open science are potentially overwhelming if they are dealt with in a piecemeal fashion. We posit that these functions are inter-related, and are parts of a system of functions that need to be integrated, rather than being stand-alone processes. We describe open science platforms or commons that provide more or less seamless support to the research process, from information technology infrastructure to high-level analytic and artificial intelligence (AI) procedures. Such platforms generate major economies of scale, and enhance impact and voice through coordination of effort. They have lessons for any African initiative: that major programmes of pan-African relevance are the best way to enthuse scientists to create virtual critical masses and intra-African collaboration; and that hub and node networks ensure both national commitment and effective coordination.

In **Chapter 6** we highlight the magnitude of the task of building a strong open science capacity on the contemporary framework of African science and its open science activities. African science suffers from the lowest rate of investment in science per head of population of any continent, such that the largest part of investment in science comes from outside the continent. It has few research centres of a critical mass, a low level of intra-African collaboration, and many of its universities are deeply underfunded by international standards. There are few centres of high performance computing, effective Cloud systems are rare, networks are under-funded, and open science policies and standards are not coordinated across the continent. Strengths that have potential for development and impact include the Square Kilometre Array (SKA) that is developing powerful computational capacity between several member states. There are a number, though too few, excellent databases and several platform projects that have high potential, and major World Bank investments, particularly in digital education. There are excellent circum-Africa internet connections, whilst the National Research and Education Networks (NRENs) have the potential to develop as an effective intra-African network provided that they are better funded and federated. Commitment and sustained support from international agencies will be important for future development.

Chapter 7 identifies crucial enablers of open science that need to be put in place and inhibitors that need to be minimised or removed. Common policies are required for intellectual property, data standards, open access publishing and shared and interoperable infrastructure. Inter-institutional work is required to develop performance metrics for scientists that incentivise rather than punish open science, and the skills required to support open science processes need to be evaluated and planned. Africa, as elsewhere, needs to develop its scientific cultural norms in favour of open science. Institutions should be encouraged to endorse the Science International Accord on Open Data. A concordat should be developed with international funders that protects the IP and career development prospects of African scientists, and there needs to be a conversation between stakeholders about how to contextualise open science in an African setting. A questionnaire circulated to 15 SGCI members elicited a unanimous view that engaging with the digital revolution was a key priority for Africa, and a strong view that a collaborative open science initiative in which the Science Granting Councils played a role should be a priority.

In **Chapter 8** we bring together the strands of our enquiry in a series of recommendations. The strength of the Science Granting Councils lies in their intermediary position between governments and the science community, influencing and being influenced by both. Acting as a collective, they could achieve efficiencies of scale, stimulate virtual critical masses, intra-African collaboration and enhanced impact. They should consider the timely creation of an African open science area. They should explore the potential for convergence of relevant national policies, for radical changes in the modes of scientific communication and the use of science evaluation metrics. They should explore means of federating IT systems. They should engage with stakeholders in plotting a way forward, including governments, policymakers and science academies; researchers and their institutions, particularly the universities; and international supporters in seeking greater strategic convergence between their respective priorities.

1. Introduction

This paper has been commissioned by the African Science Granting Councils' Initiative (SCGI) as a motivating contribution for the theme of the Science Granting Council's 2019 Annual Forum in Tanzania: *Open Science in Research and Innovation for Development.*

1.1 Context for the paper

Science systems worldwide are grappling to adapt to the consequences of the digital revolution, to the opportunities of the 4th industrial revolution that has been enabled by it, and to the challenges of global sustainability and Agenda2030. The new paradigm of open science has been widely seen as a powerful vehicle for responding to these challenges, and potentially as the future for science in the 21st century. Given the inevitable uncertainty surrounding the hypothesis that open science is indeed the future, the dilemma for Africa is whether national systems should be left to respond in their own ways, or whether the issue is so important that coordinated, collective action is required to generate the energy and impact needed to avoid Africa being left on the wrong side of a major knowledge divide. This report is partly designed to help the Science Granting Councils assess the risks associated with these choices.

The global Open Science movement has accelerated in development and up-take over the last decade, and in a variety of exploratory forms. The emerging paradigm is fashioned from converging, mutually reinforcing trends: universal access to knowledge via the world-wide-web, open access to digital publishing that has displaced the restrictions of paper text, data-driven science that adds a new dimension to the classical hypothesis-driven mode of scientific enquiry, and data sharing that enhances the efficiency of discovery and opens novel data-enabled potential in understanding the complexity at the heart of most major contemporary societal challenges, of development and of global sustainability. This confluence is creating new approaches to the generation, diffusion and governance of the scientific process, using new tools, technologies and frameworks that are a consequence of the digital revolution of the last 3 decades, which is also the driver of the "4th industrial revolution", that is fundamentally dependent on these scientific, technological and sociological innovations. This Open Science (OS) is therefore premised on the need for enhanced collaborations in research and innovation, increased knowledge exchange and greater uptake and utilization of knowledge for socio-economic development [1].

At the same time science confronts an information-rich world that although it needs scientific understanding more than ever, does not do so as a passive recipient of scientific wisdom. To be effective in its societal contribution, science must also be open to society in a two-way process of dialogue in which science engages more deeply with business, policymakers, governments, communities and citizens as knowledge partners in ways that are action-oriented and increase both effectiveness and socio-political legitimacy [2].

.

1.2 Methodology

The study was undertaken by four scientists (appendix 1), experienced in the domain of science policy, with a range of complementary experiences of open science at both the global and African levels. The study method comprised three parts:

a) An initial analysis of the new paradigm of open science, its evolution, its tools and its potential for Africa, was submitted as a basis for the proposed paper's chapter sequence and content. It was the core of our bid for the contract, and we have largely followed that sequence in the resulting paper.
b) Analysis of the peer-reviewed and grey literature with the purpose broadening and deepening the issues identified

in a) and in particular exploring arguments about workable conceptual and operational frameworks to enable efficient and effective open science in Africa. Knowledge derived from this process informed our recommendations on potential roles of SGCs in enabling open science research and innovation for development. (The involvement of three of us - Boulton, Mwelwa, Wafula - in the landscape study of open science in Africa conducted by the South African Academy of Science on behalf of the African Open Science Platform was a valuable source of information about the African landscape of open science - chapter 6).

c) A questionnaire survey of SGCs was undertaken to elicit their experiences of and approach to open science, and its relevance to development and to the 4th industrial revolution. The questionnaire is shown in appendix 3. Responses are analysed and discussed in chapter 7 (Sections 7.7). The relevance of these analyses to our conclusions is presented in chapter 8.

At the end of the each of the chapters in the following text, we have set out shaded text boxes containing key messages that are relevant to the potential roles of the SGCs. These then contribute to the synthesising discussion and recommendations in the final chapter.

1.3 Remit

The remit of this paper is to review the issues surrounding the evolving open science movement, the challenges and opportunities it presents for Africa, and the ways in which the Councils could beneficially intervene. It was framed by the following questions:

- What roles could Science Granting Councils play in fostering Open Science in research and innovation for Africa's development and how can they effectively play this role within the OS ecosystem?
- What tools, interventions, policies, incentives, infrastructure and frameworks are required to foster OS in research and innovation for development? Which of these are of immediate

relevance and importance for Africa's Science Granting Councils?

- What are the key enablers and inhibiters for mainstreaming and implementing OS policies, initiatives and activities in Africa; and how can they be sustained and resolved respectively?
- How is OS governed? Who are the key players? How are the rules, roles and responsibilities determined in the co-creation and utilization of open knowledge? What are the experiences across the 15 SGCI countries?
- What are the pros and cons of OS? Is OS increasing marginalization or bridging the divides? How can OS benefit excluded/vulnerable groups?

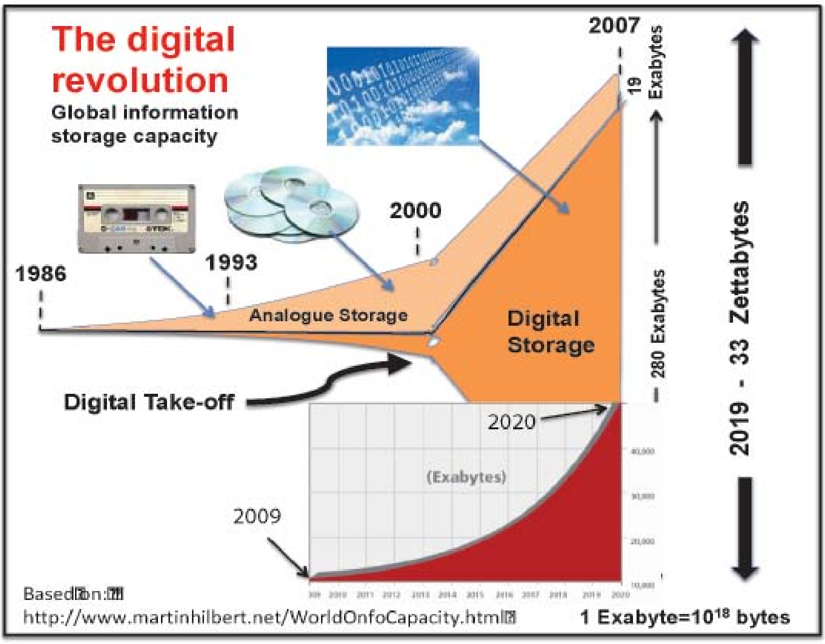
2. The digital revolution: complexity, innovation and open science

Open science is not new. The first published scientific journals in the 17th century ushered in the modern era of scientific openness. They required authors not only to submit their ideas, but also the evidence, the data, on which these were based. This permitted others to scrutinise the logic of the postulated concept/data relationship, and to replicate experiments or observations. It was a process that was well adapted to the discovery of error, a process termed scientific self-correction by historians of science, and one on which the rigour of modern science is based. It is reflected in Albert Einstein's comment, that *"no amount of experimentation can prove me right. A single experiment can prove me wrong"* [3]. It is the reason why science has become the most reliable way to acquire new knowledge and the basis for its benefit to society.

The conditions for such openness have now changed. Since the turn of the millennium, the replacement of analogue by digital technologies for the acquisition, storage, communication and analysis of data have created a digital revolution (Box 1) with powerful and pervasive consequences for science, economies and society, as a consequence of cost savings and flexibility [4]. This digital revolution has had four broad consequences for science and society, which are summarised below (2.1-2.3) and which lie at the heart of the rationale for action by the Science Granting Councils.

BOX 2.1 – THE DIGITAL REVOLUTION

New modes of digital data acquisition have created enormous and growing volumes, now measured in zettabytes, equivalent to a trillion gigabytes, with 1 zettabyte equivalent to approximately 3 million galaxies of stars. Some data acquisition systems acquire "big data", flowing into storage systems at formidable speeds, and which quickly create enormous data volumes. But we also collect a huge diversity of data, much of which is not "big' as described, but potentially of immense value in permitting integration of data about wide sets of attributes that characterise components of complex systems, enabling the recognition of deep patterns that have never previously been seen. The challenge for data science is to integrate data from diverse sources to reveal deep-lying patterns the complexities of nature and society.



2.1 Opportunity and challenges for science: unravelling complexity

The major, pressing global scientific, economic and societal issues of the 21st century (including climate change, sustainable development, disaster risk reduction) are inherently complex. They are embedded in complex systems whose property is to show emergent behaviour, which is behaviour that cannot be predicted simply by considering the inputs separately, but requires the interoperation of all major system elements to be analysed. The outcome is not merely the sum of system parts. Achieving these ends depends on the use of new tools, new tasks and new ways of working.

New tools. Machine learning algorithms mimic human cognitive functions such as trial-and-error learning and pattern recognition that have always been essential components of scientific analysis. Like us, they learn from experience, with data as the experience. A continuing "big data" flux permits progressive learning that has the capacity to reveal deep, hitherto unrecognised patterns in data. They offer a novel route to understanding and action using machine learning to extract knowledge directly from the data deluge. It is the basis of a new scientific paradigm of "data-driven science", which permits us to discern spatial and temporal structures in data that go far beyond pre-existing capacities. It creates a basis for models that learn much more than traditional data assimilation approaches and can form a firmer basis for policy and action in science and other areas of life.

New tasks. Understanding such systems can only be achieved through research that works across disciplines, and which uses a transdisciplinary approach to translate understanding into action. Achieving this depends upon our capacity to extract knowledge from the large and diverse volumes of heterogeneous data that are increasingly available, and which reflect the behaviour of complex systems. However, our ability to combine data from heterogeneous sources and across disciplines remains rudimentary at worst, excessively resource intensive at best. It is a foundational issue for 21st century science that is an increasing focus of international attention (see International Science Council Action Plan: 2019-2021- 3). A further task lies in managing unprecedented data fluxes so they are open to scrutiny at the time of publication of concepts based on them in ways that uphold the vital principle of scientific self-correction. Not only does all relevant data need to be made available, but also the metadata (the data about data), relevant computer codes, and in many instances the details of machines used in computational analysis. These requirements are summed up in the FAIR principles (Findable-Accessible-Interoperable-Reusable). Failure to ensure that data management and statistical procedures keep pace with the digital explosion in ways that are compatible with the requirements of the principle of self-correction are in part responsible for the epidemic of non-reproducibility that has occurred in many fields of science [5].

New ways of working. Addressing the challenge of complexity requires scientists to have access to broad ranges of interdisciplinary data. Maintaining the pre-existing mode, whereby scientists only have access to data that they have created, subverts this potential and the potential of much data-driven science. It is a major driver for the promotion of the new open science paradigm. Such data sharing and marketing between commercial companies provides the feedstock for the technologies that companies use to enhance their efficiency and market impact. The same is true in science.

2.2 The digital revolution, driver of the 4th industrial revolution

The digital revolution is an event of world historic significance. Its technologies together constitute a “general purpose” technology that is driving what has become known as the fourth industrial revolution (Figure 2.1), though penetrating far beyond the confines of industry. These are technologies that continually transform themselves, progressively penetrating new domains, boosting productivity across all sectors and industries because of their cost effectiveness. They are globally pervasive, with profound economic and social implications that fundamentally disrupt pre-existing norms. They have unleashed an unprecedented new era of innovation, with profound implications, not only for science, industry and economies, but also for society and all levels of governance.

Revolutionising, economies, societies, lives

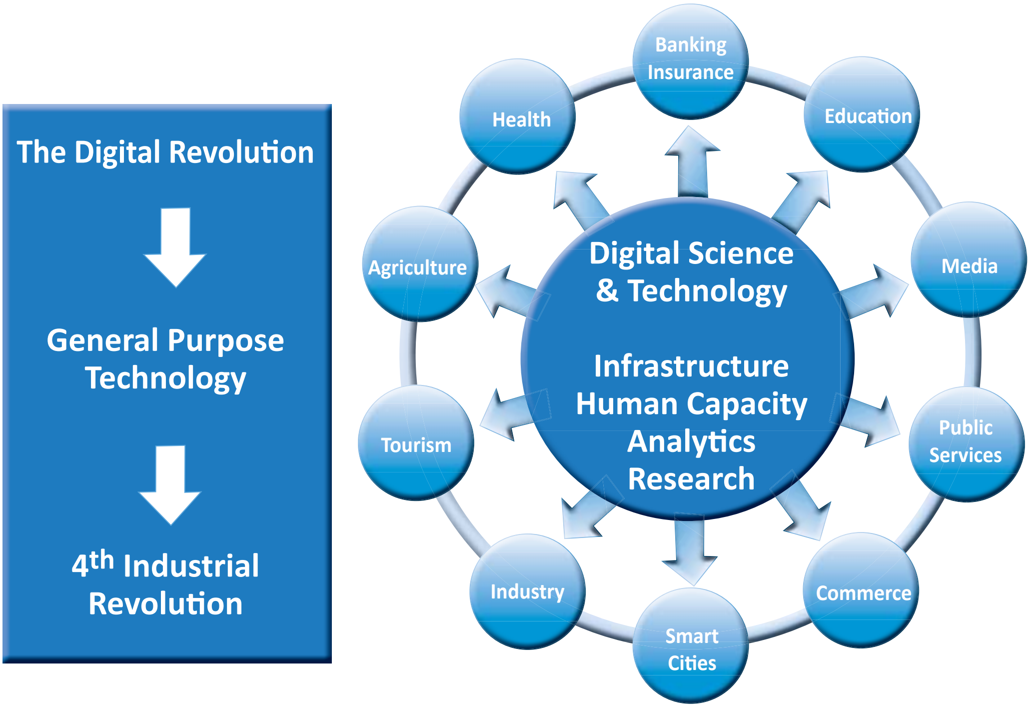


Figure 2.1. How the digital revolution drives and is at the core of the "4th industrial revolution" and its ubiquitous applications.

2.3 Empowering diverse voices and "post truth"

A major consequence of the digital revolution has been the creation and adoption of digital communication devices, with a global penetration of "smartphones" at 37% of the global population, and with Africa at 26%, but increasing at the fast annual compound rate of 6.7%. They are now the preferred means of web access, and have fundamentally changed social and commercial interactions and retail activities. Whilst a decade ago it was assumed that these technologies would democratise communication and action, the unanticipated dynamic has been their use as means of dividing societies into political and social siloes. It is increasingly described as a means of broadcasting misleading or blatantly untruthful statements, as an aspect of a so-called "post-truth" world, where a partial opinion, no matter how outlandish, can be given the same credibility as a fact.

A recent study of Twitter [6] found that falsehood diffused faster, deeper, and more broadly than truth in all categories of information. They were 70% more likely to be re-tweeted than the truth, and reached more people, due in part to peer-to-peer sharing. Within the typology of false news, political news travelled more deeply and more broadly. For scientists this appears most damaging where well-established relationships or rigorously tested concepts, such as the smoking-cancer link, the health benefits of vaccination or the evidence of human-induced climate change, are denied without credible evidence. It has been supposed that this reflects a lessening in public trust for science, but if anything, trust in science is rising [7]. However, although the proportional trust for science may be rising, the powerful tools of ubiquitous communication and the world-wide web have given dissenting voices a broadcasting power that they have previously lacked, making them "the most powerful machine for the spreading of lies that the world has ever known" [8].

2.4 Emergence of the new "Open Science" paradigm

The suite of powerful digital technologies that have emerged in recent decades have naturally led to new ideas about the opportunities that they offer for science and how science systems and norms might need to be re-configured if these opportunities are to be seized. This new paradigm of open science is based on open data and open access to the results of scientific inquiry, as means of enhancing efficiency, the rate of discovery, understanding of complex systems and, in collaboration with other societal actors, of innovation. At the same time, it has become clear, in the face of the data deluge, that greater discipline in data use in particular will be needed if science is to retain statistical rigour and uphold the principle of reproducibility in the face of the data deluge. As Jim Gray [9] commented "we scientists do terrible things with our data". The recognition of these opportunities and challenges has led to the definition by computer scientists of what has been called a fourth paradigm for science, an "E-science" that adds computer simulation linked to data-intensive science (with its three basic activities of: capture-curation-analysis), to the classical scientific paradigms of observation, experiment and theory. It adds an approach of "data-led science" to the classical approach of hypothesis-led science.

This is an important technical root for the new open science paradigm, but by no means the only one. Analysis by social scientists of the assumptions, drivers and habits of scientific enquiry suggested that the established paradigm of scientific discovery (mode 1), characterised by the hegemony of disciplinary science, with its strong sense of an internal hierarchy between the disciplines and driven by the autonomy of scientists and their host institutions, the universities, was being superseded, although not replaced, by a new paradigm of knowledge production ('Mode 2') which was socially distributed, application-oriented, trans-disciplinary and subject to multiple accountabilities [10] [11].

These two perceptions are the forebears of the modern open science paradigm. The first based on the use of powerful digital technologies to enhance the capacity of science to discover new knowledge: the second as a response to the need for a broader disciplinary input in understanding the complexities of nature and society, and a responsibility for broader societal engagement in translating this understanding into actionable knowledge.

These are perspectives that have arisen from and been driven by the science community and are the background to and conceptual drivers of open science. They create an evolving setting within which the necessary inter- and trans-disciplinary collaborations needed to understand complex systems can be forged. They imply that without science becoming a more publicly engaged enterprise, the application of its understanding to the problems of the modern world will have a diminished potential. As a consequence, the initial focus of open science on open data and open access publishing has extended recently to include open engagement with society.

2.5 The “dark side”

Most technologies are ethically neutral but have the potential to be used for harm as well as benefit. It is incumbent on governments to determine whether and how regulations or legal restraints are required to prevent harmful use, but also to ensure that national technical skills are developed that are able to identify and mitigate risks. The dark side of the digital revolution lies in cyber-fraud, a massively growing industry, cyber-warfare, including attacks on national infrastructure, cyber-espionage, including attacks on the integrity of databases, and cyber-lies that undermine civic consensus and electoral integrity. Access to the skills and capacities necessary to identify and deflect threats or mitigate their consequences are vital capabilities for a modern state.

2.2 Relevance to the Science Granting Councils’ initiative

Points of advocacy to government

- the digital revolution is a world-historic event
- its technologies are the bedrock of the 4th industrial revolution
- a strong science base is essential to exploit its opportunities and address its challenges
- a new paradigm of “open science” is developing as an efficient way of doing so
- high level data science and IT are vital in protecting key national assets and formulating relevant regulations and legislation

3. The imperative for Africa

3.1 Why must a modern state respond energetically to the digital revolution?

The technologies of the digital revolution have already shown enormous capacity to create long-term benefit precisely because they are so flexible and pervasive, with many benefits coming not simply from adopting the technology, but from adapting to it. But by their very nature, they are highly disruptive in the short-term, rapidly redefining relationships between customers, workers and employers, and permeating almost everything we do, progressively overhauling all industries whilst creating new ones. African governments cannot avoid these forces that technology has unleashed, which may have short-term disruptive consequences as well as long-term benefits. Just as their western and Asian counterparts are doing, they must promote creative thinking and acting, in and beyond government, about structural adaptation, widespread re-skilling and educational innovation, to minimise short-term disruption and maximise long-term benefit.

The digital revolution, which has largely replaced the printing technologies invented in the fifteenth century, offers immediate, democratised access and has destroyed distance as a barrier to the spread of information. It has also reduced costs and has done away with the space limitations of print pages and books. It has serious implications for the conduct of science and technology in Africa and throughout society. We argue that African governments need to take a cue from counterparts elsewhere by developing systematic adoptive and adaptive responses that are aligned with the aspirations of the *African Union Science, Technology and Innovation 2024 report* [12], in order to capitalise on the digital revolution. However, time is short. Whereas the full impacts of the printing revolution were centennial, those of the digital revolution are decadal. Although the latter has not yet run its course and its ultimate destination remains hidden from us[G], playing catch-up holds much less promise than being near the head of the pack.

3.2 What are the crucial questions to which African states must respond?

The imperative to respond to the digital revolution is global. Its impacts cannot be avoided, in Africa or elsewhere. Governments, industry, commerce and national science systems worldwide are struggling to understand their long-term significance whilst adapting to what are seen as immediate imperatives. For science, these are currently supposed to be in three dimensions:

a) How should priorities, incentives, infrastructure and funding in national science systems be adapted to exploit the new digital world to best effect across the whole spectrum of science and its application?
b) How should capacities and capabilities in informatics (computer science-data science-artificial intelligence) and in data engineering be developed and prioritized, not only for the benefit of the science system but also in their provision of skills for public and private sectors?
c) A national science system does not operate in a social vacuum. It is an essential element of national intellectual infrastructure with a value to society largely determined by the way in which it interacts with society to simulate innovation. How does a science system need to adapt to a digitally-aware society with its social media, instantaneous communications and global information and disinformation webs?

[G] *When Zhou Enlai, Chinese premier from 1949 to 1976, was asked what he thought were the benefits of the French Revolution (1789), he is reported to have replied: "it is too early to say".*

These questions are as urgent and insistent for African states as they are for all others. They cannot be deflected or ignored, for the alternative would be to risk stagnating in a scientific backwater, isolated from creative streams of social, cultural and economic opportunity. A country that fails to develop its own capacities will inevitably be dependent upon skills bought in from elsewhere as a passive and ill-informed consumer of expensive data services, lacking the creativity to thrive in a fast-changing world [13]. A clear danger is that Africa's relatively weak contribution to global knowledge creation (see 6.1) [14] could deteriorate, with potentially profound consequences for the continent's vitality.

Whilst there is a risk that even by strenuous, adaptive digital policies, Africa's economic performance could lose "market share" because of the creativity of other better-favoured economies, failure to adapt at all would certainly lead to serious economic deterioration. Studies of the impacts of digitized information flows show that they have only slightly decreased inequalities, with Africa lagging behind the rest of the world [15]. On this basis, for a state to do other than equip itself to the best of its abilities with the skills, the support mechanisms and the opportunities for translation of cutting-edge digital technologies would be unwise in the extreme.

3.3 Could Open Science be the vehicle for a cost–effective response from Africa?

Open science is, in part, developing as a means of maximising the scientific and socio-economic impact of the digital revolution at the national and supra-national levels and at the levels of disciplines, and with the intention of mainstreaming its processes within national or disciplinary science systems.

The cost effectiveness of the open science project has been a major issue [16a, 16b], and whether, in the African case, the net economic effect would be positive or negative. To this extent, our search of the African landscape for data on the potential economic effect of open science yielded few results, as there is little coverage [17]. A World

Bank study however [18] concludes that the economic potential of open data is very large indeed, and that these conclusions apply equally to both developed and under-developed economies. It suggests that governments should see themselves not only as supplier of open data but also as leaders, catalysts and users. A 2015 study for the European Commission [19] argued that a European open data portal would have the potential to generate a multi-billion euro bonus per year, including a cumulative efficiency benefit of 1.7Bn euros by 2020 [20]. Another report offers a deliberately conservative estimate of the opportunity costs (benefits foregone) for the European Union of not developing an open regime, where data is findable, accessible, interoperable and reusable (FAIR data), as at least 10.2 Bn euros and possibly as high as 26 Bn euros [21]. These considerations form a fundamental justification for the major European investment in an Open Science Cloud.

These are powerful arguments in favour of open science, although we acknowledge that the African are different from those of Europe or the Unites States. However, in the absence of directly applicable data, we suggest that they provide strong reassurance that investment in relevant capacities would be productive. It is the political dilemma that considerations of the future always hold, but one where inactivity is a strong, but in our view an ill-advised choice.

One of the elements of the open science enterprise that could be of great importance for Africa is the ethos and practice of sharing and collaboration inherent in the socially distributed, application-oriented, trans-disciplinary approach derived from the “mode 2” discourse [22] described in 2.1, and increasingly embedded in the open science paradigm. The 2024 STI Strategy for Africa [23] identifies two fundamental weaknesses of science systems in many African countries as weak intra-African collaboration and inadequate critical mass. These are precisely the weaknesses that the collaborative practice of open science could correct. Furthermore, individually weak systems can strengthen themselves through the efficiencies of shared resources by:

a) efficiencies of scale in planning, procurement and provision;
b) scaling-up through collaboration and shared capacities;
c) stimulating creativity through interaction of diverse groups;
d) amplifying impact through common purpose and voice;
e) building consortia and collaborations with a greater critical mass;
f) support from a shared capacity in cutting-edge data science.

There are of course dangers and difficulties [24]. These include potential loss of intellectual property to larger better-funded groups from beyond Africa, and the continuing difficulty of affordable access to the international scientific literature; relative scarcity of a high band-width internet; lack of open access policies to govern open science; and lack of standardization and interoperability amongst data repositories.

3.4 Timeliness: the African Continental Free Trade Area

In September 2019 the African Union launched the African Continental Free Trade Area (AfCFTA) comprising 54 out of 55 countries. Not only has it great potential to dynamise trade on the continent, but, as the European experience has shown, free trade and the mobility of ideas and people also stimulate social and cultural dynamism. These are precisely the qualities that science both thrives on and contributes to. An initiative to create an African open science area following hard on the heels of the AfCFTA announcement would represent a major statement of intent from Africa about a confident and creative scientific future. The synergies between these actions would have the potential to be powerful levers of social, cultural and scientific vitality and of economic development.

3. Relevance to a Science Granting Councils' initiative

Points of advocacy to government

- technologies unleashed by the digital revolution cannot be sidestepped
- national science systems must adapt to new demands and opportunities
- "open science" may be a cost efficient way of adapting to them
- powerful potential for synergy between AfCFTA and an African Open Science Area

4. The essentials of open science: frameworks, policies and tools

The capacity to conduct open science in a data-intensive age is fundamentally dependent upon the development of tools and procedures to acquire, store, communicate and analyse large and complex data fluxes, to manage these processes efficiently and in a highly structured fashion, and to communicate the results in ways that make them accessible to the largest number who may be able to use them for personal or collective benefit. Knowledge, scientific knowledge in our case, when released into the public domain, has long been regarded as a public good. Maintaining that public good in a data rich age is crucially dependent on our capacity to manage data and knowledge transfer in an efficient and coherent way. Otherwise, we risk drowning in a data deluge, and fail to realise it as a public good.

We now analyse what have come to be regarded as the essential tools and processes that need to be in place and the problematic issues that must be addressed if open science is to be efficiently delivered. It is not Africa-specific, but draws on evidence of good practice from wherever it is available. It sets the conceptual frame for chapter 5, which describes how these principles and practices have been implemented in open science systems. It is conventional to take open science as comprising open data and open access publishing, to which we add open to society as a necessary, outward facing attribute.

4.1 Open Data

If the opportunities of open science are to be grasped, the data, which are essential parts of the bedrock on which science is based, need to be efficiently and effectively managed. Data are key conduits leading to knowledge discovery and innovation, and need to be widely available for scrutiny to ensure the logical rigour of scientific claims. They need to be available for interdisciplinary integration and for reuse by the community. Unfortunately, the existing digital ecosystem surrounding scholarly data publication prevents us from extracting maximum benefit from our research investments, for example in text and data mining and because many scientific publishers still do not require data to be accessible or FAIR as a condition of publication (25). If we are to make best use of the data deluge rather than being confused or drowned by it, a series of technical demands need to be satisfied. We now set out the technical solutions that have been developed about how data resources should be managed, what defining characteristics of data need to be maintained in order that they can be productively shared and re-used by others, how data should be cited, and related legal issues. It is also critical that scientists are supported in ensuring statistical rigour in their analyses and in the use of some of the powerful techniques of machine learning that are able to discover deep structure in data.

4.1.1 Research Data Management

Such are the volumes and complexities of modern research data, that any organisation or institution that has a sustained need to utilise these resources will need to develop a strategy and a system for research data management (RDM). The creation of an efficient database, able continually to absorb new data and release data for use in ways prescribed by users is not a trivial matter. Several such systems have been tried and tested, all of which have similar characteristics, being based on the so-called data lifecycle. The influential Digital Curation Centre model of the data life cycle (26) is shown in figure 4.1.

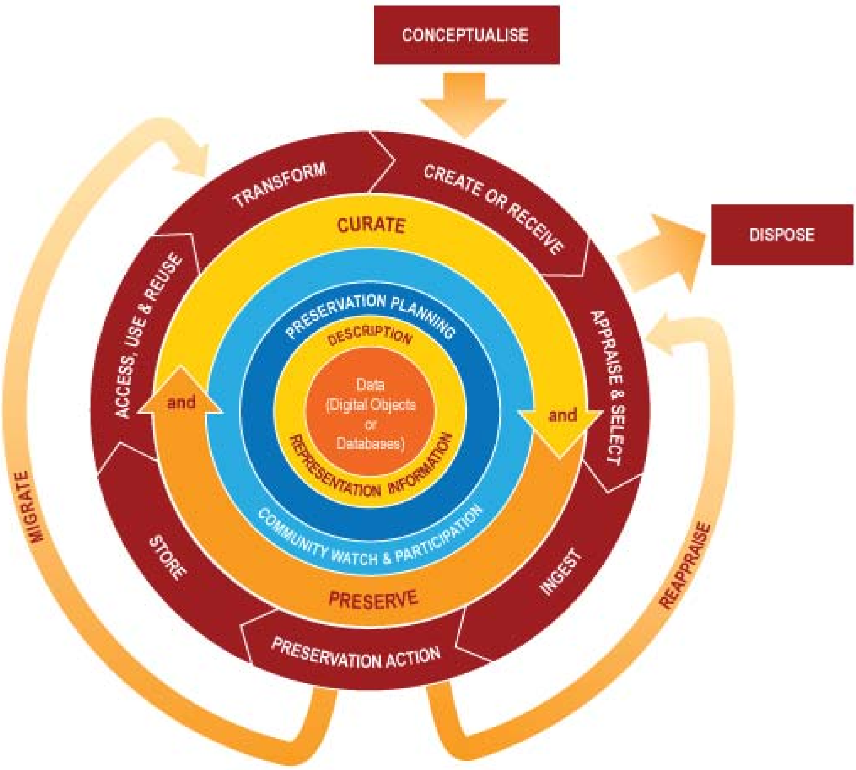


Figure 4.1: The Digital Curation Centre (DCC) data curation lifecycle model

The model provides a high-level graphical overview of the stages of successful curation and preservation of data from initial receipt of acquired data to its use or re-use in a scientific analysis. Starting with the data there are five key components in the management system:

a) Data, in binary digital form. It includes

- *Simple digital objects:* discrete digital items.
- *Complex digital objects:* combining digital objects, such as websites.
- *Databases:* structured collections of records or data stored in a computer system.

b) Description and Representation Information

- Assign administrative, descriptive, technical, structural and preservation metadata.

- Collect and assign information to understand the digital material and its metadata.

c) Preservation Planning

- Plan for preservation throughout the curation lifecycle of digital material.
- Include plans for management and administration of all curation lifecycle actions.

d) Community Watch and Participation

- Maintain community activities, and participate in the development of shared standards, tools and suitable software.

e) Curation and Preservation

- Manage and generate actions to promote curation and preservation throughout the lifecycle.

Once these management systems are in place, the life cycle for a specific scientific task is:

- Conceptualise: by planning data capture strategy and storage options.
- Create: digital objects and assign descriptors.
- Access and use: ensure routine access.
- Appraise and select: evaluate requirement for long-term curation and preservation.
- Dispose: of digital objects not selected for long-term curation and preservation.
- Ingest: transfer digital objects to a trusted digital repository or data centre.
- Preservation action: ensure long-term preservation and retention.
- Reappraise: digital objects that fail validation are further appraised and reselected.
- Store: keep the data in a secure manner as outlined by relevant standards.
- Access and reuse: ensure data are accessible to designated users use and re-users.

Understanding this life-cycle is of fundamental importance for the practical creation of an operational data management system. Figure 4.2 illustrates the DDC Research Data Service Model, a management system based on the lifecycle model in figure 4.1. This is a generic model that is adaptable to the needs of the organisation or institution, whether small or large, that needs to manage its research data. It is based on the need to manage both technical infrastructure and human resources.

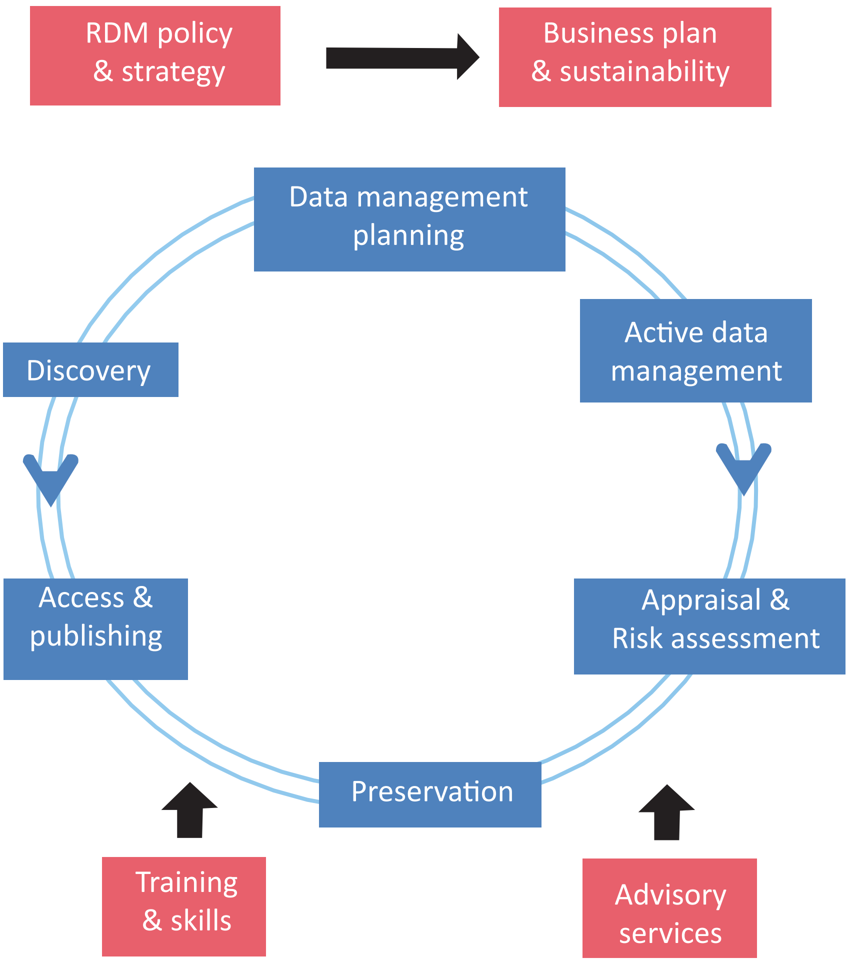


Figure 4.2. The Digital Curation Centre's Data Service Model

There are four fundamental attributes of a system that is able to service the data cycle shown in figure 4.1:

Purpose & strategy

It is vital that the purpose of the management system is clearly defined, whether it is to serve all disciplines or a sub-set, whether it is for a research group or institution, whether it is national or regional, and whether it is a federation of interoperable RDM systems. Whichever it is, coordination is essential because effective RDM is extremely difficult to achieve when components are designed in isolation. An open science system will be open to many different groups, and may face outwards towards society, and thus will need to embed this capability in its initial policy for design. Irrespective of its purpose, the system must be clear about:

- purpose and scope of the proposed service;
- identifying current provision;
- identifying feasible levels of service provision;
- roles and responsibilities identified in the policy;
- how policy is communicated to stakeholders;
- mechanisms to monitor and review.

Business plans and sustainability

Many current databases have had severe difficulties because of a reliance on short-term funding. Important databases have collapsed when such funding has ended. The business model must be based on the design purpose and will need to address:

- making the business case for the service, considering its value proposition to service customers;
- committee processes and timelines for securing resources for improved technical and human infrastructure;
- costs and benefits associated with RDM support provision;
- cost recovery models and research funder rules that govern what direct or indirect costs may be charged to research grants.

Advisory services

Many users will find management systems difficult to understand or to engage with unless the user interface is a simple one. The provision of online and face-to-face advice for researchers who need support with a particular aspect of their research data management is crucial. Important issues to resolve are:

- which staff deliver support to researchers across relevant professional service units, and what scope is there to join this up?
- on which topics is the advice provision strongest and weakest?
- which channels are used to connect researchers to any support already available, and what scope is there for using online connection more efficiently?

Training

The provision of basic training in RDM principles, practices and processes is vital. This may be done through online and/or face-to-face delivery of learning materials designed to meet the needs of both researchers and support staff. It should involve planning:

- what objectives does the training programme aim to address, e.g. which capabilities of the service will be improved;
- whose skills or competencies need to be developed, and what are they;
- what channels are used to connect staff and researchers with training opportunities;
- how can RDM be aligned with other learning approaches.

4.1.2 FAIR Data

It has been long recognised [27] that for data to be reused, particularly by those who are not the data originators, that it is not enough just to deposit the data in a repository and presume that others will be able to use it. For all but the simplest data, a great deal of metadata is required to make it (re-)useable. First it must be known to exist. Then it must be able to be retrieved from wherever it is found [28]. Then it must be able to be combined with data from other sources. Finally, a full description of the relationships of the data must be available so that it can be meaningfully re-used. A Royal Society report [29] articulated a view of what it called “intelligent openness” which required data to be Accessible, Assessable, Usable and Traceable. A fuller analysis by Force11 [30] enunciated the FAIR principles, that data should be Findable, Accessible, Interoperable, and Reusable. Box 4.1 describes what is meant by each of these terms [31].

For many high value datasets, there are well-organised, deeply integrated repository systems in areas such as genetics, space physics and astronomy, where these principles are readily and routinely applied. But there are many important datasets, from more traditional low-throughput bench science or routine low-tech observation, that are often of no less importance than their big data counterparts, but where the application of FAIR principles is less standardised and potentially more onerous, and for which more general purpose databases such as Figshare [32] or Dryad [33] have been developed.

A significant challenge is that of making scientific data “machine actionable.” For example, a machine may be capable of determining the datatype of a discovered digital object, but not capable of processing the data or determining the licensing requirements. The optimal state—where machines fully ‘understand’ and can autonomously and correctly operate-on a digital object—may rarely be achieved. Nevertheless, the FAIR principles provide ‘steps along a path’ toward machine-actionability. Adopting, in whole or in part, the FAIR principles, leads the resource along the continuum towards this optimal state [34].

BOX 4.1: REQUIREMENTS FOR "FAIR" DATA:
To be Findable: **F1:** (meta) data are assigned a globally unique and persistent identifier **F2:** data are described with rich metadata (defined by R1 below) **F3:** metadata clearly and explicitly include the identifier of the data it describes **F4:** (meta) data are registered or indexed in a searchable resource **To be Accessible:** **A1:** (meta) data are retrievable by their identifier using a standardized communications protocol **A1.1:** the protocol is open, free, and universally implementable **A1.2:** the protocol allows for an authentication and authorization procedure, where necessary **A2:** metadata are accessible, even when the data are no longer available **To be Interoperable:** I1: (meta) data use a formal, accessible, shared, and broadly applicable language for knowledge representation. I2: (meta) data use vocabularies that follow FAIR principles I3: (meta) data include qualified references to other (meta) data **To be Reusable:** R1: meta (data) are richly described with a plurality of accurate and relevant attributes R1.1: (meta) data are released with a clear and accessible usage license R1.2: (meta) data are associated with detailed provenance R1.3: (meta) data meet domain-relevant community standards

4.1.3 Data citation

Data citation is a key practice in support of data access, sharing, reuse, and of sound and reproducible scholarship. Many problems arise when research findings become disconnected from the underlying data that form the evidence for these findings. The most well-publicized of these problems is scientific fraud.

Within the social sciences, the vast majority of datasets produced by sponsored research are never deposited or shared [35], and, as a result, reproducing published tables and figures, and directly extending prior results is often difficult or impossible [36], [37], [38]. Similar problems exist in other fields. A recent study [39] of a sample of zoology articles found that less than 30% of even the most recent publications made data available, and that research data availability declined rapidly with article age, while loss of data increased.

The purposes of data citation as they have developed so far are:

- to facilitate description and information retrieval, using the principles that data in archives should be described as works rather than media, using author, title, and version;
- to support data access and persistence, associated with the principle that research data used in publication should be cited, and that those citations should include persistent identifiers, and should be directly actionable on the web;
- to support the use of citations for verification and reproducibility, including the principle that citations should support verifiable linkage of data and published claims.

Data citation principles as developed by Force11 are described in box 4.2 [40].

BOX 4.2 Joint Declaration of Data Citation Principles (2014)

1. **Importance.** Data should be considered legitimate, citable products of research. Data citations should be accorded the same importance in the scholarly record as citations of other research objects, such as publications.
2. **Credit and Attribution.** Data citations procedures should facilitate scholarly credit and normative and legal attribution to all contributors of the data, recognizing that a single style or mechanism of attribution may not be applicable to all data.
3. **Evidence.** In scholarly literature, whenever and wherever a claim relies upon data, the corresponding data should be cited.
4. **Unique Identification.** A data citation should include a persistent method for identification that is machine actionable, globally unique, and widely used by a community.
5. **Access.** Data citations should facilitate access to the data themselves and to such associated metadata, documentation, code, and other materials, as are necessary for both humans and machines to make informed use of the referenced data.
6. **Persistence.** Unique identifiers, and metadata describing the data, and its disposition, should persist, even beyond the lifespan of the data they describe.
7. **Specificity and Verifiability.** Data citations should facilitate identification of, access to, and verification of the specific data that support a scientific claim. Citations or citation metadata should include information about provenance and fixity sufficient to facilitate verifying that the specific time-slice, version and/or granular portion of data retrieved subsequently is the same as was originally cited.
8. **Interoperability and flexibility.** Data citation methods should be sufficiently flexible to accommodate varying practices amongst user communities, but should not differ so much that they compromise interoperability of data citation practices across communities.

DataCite [41] is the leading non-profit organisation that provides persistent identifiers (DOIs) for research data and other research outputs to ensure that the above principles can be put into practice. Organizations can become members in order to be able to assign DOIs to all their research outputs to ensure that they are discoverable and that associated metadata are made available to the community. DataCite frequently up-dates metadata Schema Documentation for the Publication and Citation of Research Data [42].

4.1.4 Text and data mining

Text and data mining (TDM) is the process of deriving information from machine-read material. It works by copying large quantities of material, extracting the data, and recombining it to identify patterns. There are four stages in the TDM process as shown in Figure 4.3.

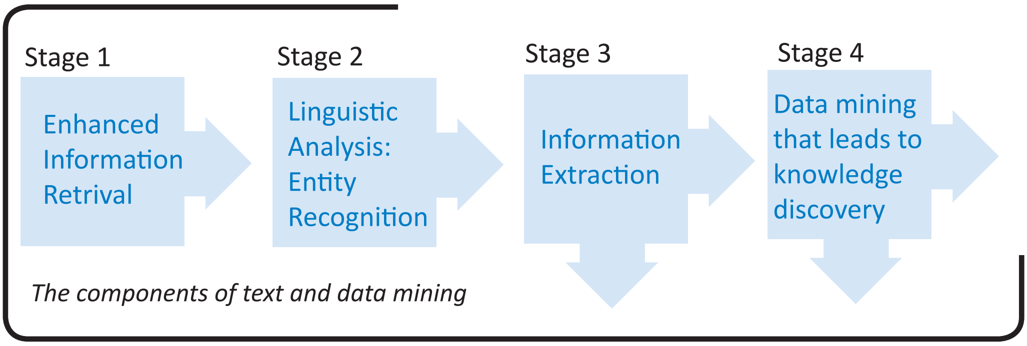


Figure 4.3: Processes in text and data mining [43]

Potentially relevant documents are identified and turned into a machine-readable format so that structured data can be extracted. The useful information is extracted (Stage 3) and then mined (Stage 4) to discover new knowledge, test hypotheses, and identify new relationships.

TDM will increase the progress of science exponentially. It has the potential to facilitate the discovery of cures for diseases such as cancer and Parkinson's. It has already been used to discover how existing drugs can be used to treat other conditions. It will also act as a foundation for innovation and new industry.

Ordinarily, authors are obliged to transfer their copyrights before publication to the commercial publishers and as a result they relinquish control over how publications are used. Hence it has not been possible to mine freely in legally accessed content made available by commercial academic publishers. This obstructs science and the distribution of scientific knowledge beyond the scientific community. It also impedes the use of TDM by private parties, depriving them of the ability to explore and innovate. Publishers have been resistant to free use of TDM, even to those who already have legal access to their journals and notwithstanding the fact that material has been freely given to them by scientists. There have been attempts to promote legislation to remove this barrier to scientific progress, but as yet without success.

4.1.5 The legal framework: copyright, licensing etc

Given the vastly different data practices and related ethics of ownership, curation, storage and dissemination in each discipline, it is important to assess differences in disciplinary approaches regarding data sharing and re-use, and to identify standards and related infrastructures that can foster communication and exchanges across fields while respecting diverse methodological traditions.

The Budapest Open Access Initiative [44] in which “open access” was defined as the “free availability of scientific literature on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.” It is a declaration that if implemented legally, would remove legal restrictions.

The use of recognized waivers or licenses that are appropriate for particular data is an emerging trend. Creative Commons licensing is an established best practice that is well-understood, providing a suite of licences that cover all needs. Open access journals are usually licensed under one of the six core Creative Commons (CC) licenses. Creative Commons offers six basic model clauses, two of which satisfy the above criteria of a free license: CC BY and CC BY-SA. CC BY- license lets others distribute, remix, tweak, and build upon an originator’s work, even commercially, as long as they credit the original creation. This type of license promotes open science and is therefore highly recommended. CC BY-SA -license lets others remix, tweak, and build upon prior work, even for commercial purposes, as long as they credit the originator and license their new creations under identical terms. All new works would carry the same license, so that any derivatives would also allow commercial use.

There are increasing numbers of articles and journals focusing on detailed descriptions of data and arguments about the value of the data for the future. Data journals do not host data themselves but recommend suitable repositories where data sets should be deposited, and then link to it. Notable examples of data journals are: Giga Science [45] Scientific Data [46] and Data in Brief [47].

It is generally the case, and we expect it to be the case in Africa, that where national research councils support a move towards open science they will issue policy statements for research grants applicants that commit them to publishing their results under open access conditions. For instance countries like Germany have such polices supported by an article in the Copyright Act [48]. It is important that African countries reconsider their legal frameworks to ensure that such enabling provisions for open science are in place. This is especially important for text and data mining.

4.1.5 Limits of Open Data

Legitimate reasons limiting data openness include the privacy of individuals or organisations, national security and safety. This calls for formulation and implementation of suitable procedures and policies that best protect the use of data in the context of developing open science [49]. The nature of competitiveness in a given field influences the researcher's willingness to collaborate and share research data with peers [50], an inherent barrier to openness. Collaborations between researchers and industrial partners with a view to commercializing the output also comes with constraints on sharing and dissemination of data resources and research findings

[51]. Unless revised in line with the requirements of a new open science paradigm and its limits, copyright law and other relevant intellectual property rights guidelines will negatively affect the dissemination of scientific results [52].

4.1.6 Data Analytics

Rejuvenated skills in statistical analysis are vital for handling large and complex data volumes where the pitfalls are serious for the unskilled. At the same time, massive, novel data resources have brought the approaches of artificial intelligence, particularly machine learning, into their own. These were developed some decades ago, with much hype, but because of the small data volumes available to them, they were only able to produce relatively trivial results. That has changed. Modern data resources are often able to satisfy the voracious appetites of these learning algorithms, which are now powerful tools in the armoury of science and of both public and private sectors. Machine leaning is at the heart of this new potential. Algorithms mimic human cognitive functions of pattern recognition, which are now supercharged through the data acquisition and processing power of modern digital devices.

4.1.7 Data Governance and Ethics

Data and AI technologies taken together are not just another utility that needs to be regulated once it is mature. It is a powerful force, a new form of smart agency, which is already reshaping our lives, our interactions, and our environments. In this fast-moving landscape of the data and AI world, governance challenges need to be addressed in a timely manner if an overall system of governance for data management and data use is to maintain public trust [172]. Existing data governance concepts, such as privacy and consent, are under unprecedented strain: their meanings in policy, law and public discourse have shifted, and will continue to do so in new and unpredictable ways. Personal data and its AI applications are able to use data that we freely give away by clicking the “Accept” box on a smart phone or laptop or every time we shop or fill in a form, in ways over which we rarely have control. Principles of good governance are essential in this new world of ubiquitous data if we are not to relinquish all standards of privacy and confidentiality without our consent, and thereby give free rein to cyber intrusion, cyber crime and cyber manipulation.

A set of high-level principles is needed to visibly shape all forms of data governance and ensure trustworthiness and trust in the management and use of data as a whole. New governance procedures are required to protect individual and collective rights and interests. They should ensure that trade-offs affected by data management and data use are made transparently, accountably and inclusively. They need to adopt good practices by learning from successes and failures to enhance existing democratic governance. The governance framework for data management and data use should perform three broad categories of functions:

- anticipate, monitor and evaluate;
- build practices and set standards;
- clarify, enforce and remedy.

Most countries have a range of actors already carrying out some of these important governance functions in their specific sectors or domains, but there is a clear need for a new body to steward the landscape as a whole, rather than being directly responsible for implementation within specific domains. The purpose of such national stewardship bodies would be to support delivery of the full breadth of critical functions in accordance with the principles set out above, and to relay them to branches of government that have the responsibility for sectoral oversight and regulation. A stewardship body would be expected to conduct inclusive dialogue and expert investigation into novel questions and issues, and to enable new ways to anticipate the future consequences of today's decisions. They should be independent, connected to diverse communities, expert across and beyond disciplines and tightly coupled to decision processes. They should be durable and visible, nationally focused but globally relevant.

An obvious area where there have been major efforts to address serious concerns is that of biomedical research and health systems in general [173]. There are significant advances in terms of personal data governance that have followed the creation of national or international ethical committees. Such committees assess any data collection initiative involving individual subjects, such that data protection and sharing in these cases are generally well regulated. However, global initiatives would improve governance of personal information, especially related to human genetic data use in research and the rapid evolution of sequencing techniques. For instance, the same consent already signed by study volunteers and approved by committees could become obsolete as a consequence of the exploitation potential of modern genomics data production machines. In such situations, investigators may need to return to patients to re-affirm their consent. Such a global perspective and management processes that include lawyers, biologists as well as data and biomedical technicians could help to frame more inclusive and sustainable governance systems. The World Health Organisation is currently considering such a development.

4.2 Open access and dissemination of scientific results

The public interest is almost invariably best served by the widest possible dissemination of scientific results as a means of stimulating innovation across society. It is an imperative that has been widely acknowledged in the United Nation's Sustainable Development Goals, and particularly Goal 9, which stresses the importance of industrial innovation and infrastructural development to ensure sustainable development for all, [53], [54]. For Africa, this imperative is two-fold. The first is for access to the internet, where the outlook is promising. The International Telecommunications Union statistics [55] show that globally in 2018, 3.9 billion people or 51.2% had access to the Internet, Africa's share being 24.4%, having grown from 2.1% in 2005, the highest growth rate internationally, although in the volume of use it still lags the rest of the world. Even if internet access of sufficient bandwidth is available however, scientific results must be available on the medium. The second priority therefore is to ensure that published scientific results are not only available but affordable. Affordability is a fundamental issue for Africa, as it is for the worldwide scientific community. Addressing the global problems that are a concern for us all requires global involvement. It is in the interests of science worldwide that their colleagues in Africa, and in the "global south" generally, are integral parts of the global scientific network, which they cannot be if excluded from access to scientific results from elsewhere. Conversely, although it is important for Africa, as it is for all science systems worldwide, to assess the extent to which current models of publications are consistent with their own interests, they must also strive to ensure that principles and processes of publication serve the whole international community, and not merely one segment. We need global solutions to the problem of affordability, which we explore below, not just African ones.

4.2.1 Recent history

With these issues in mind, it is useful to consider how current problems have arisen. The tools of the digital revolution have largely made conventional print-based approaches to dissemination of scientific work obsolete and should have led to a reduction in cost. Neither has the latter happened, nor have the current modes of dissemination adapted as well as they should to the opportunities the tools offer. As a consequence, there is an increasing body of opinion in the scientific community that regards the current system of scientific publication as dysfunctional. A little history is informative.

a) **The business model**

Until the 1960s/70s, most scientific publication was in the hands of learned societies through the medium of their journals. As learned societies were deemed to act in the interests of scientists in their various fields, it seemed natural that scientists should entrust copyright to the journals and freely offer their services to support editorial boards and refereeing processes. As commercial publishers began to invade this market at scale, they simply assumed the relationships of trust that had existed between scientists and learned societies were also open to them, even though their principal responsibility is to their shareholders, not to scientists. The outcome has been a business model of unique asymmetry. Scientists provide their work freely, or at their own cost, to publishers, give up copyright to publishers, staff publishers' editorial committees, provide peer reviews freely, and then buy back their published work at inflated costs. It has been calculated [56] that the actual cost of production of well-found articles is of the order of €300-€400. In practice, commercial publishers charge the order of 10x that amount. For example, Germany recently paid €26 million to the publisher Wiley to publish 9,500 open access articles a year over three years, at €2,750 per article. It has been calculated that their average real cost of production would have been no more than about €350. Such deals have produced for Wiley an operating profit margin of around 29.5 per cent, implying that about €7.7 million of that fee goes straight into its shareholders' pockets [57].

b) Impact factors

A trump card in the hands of commercial and non-commercial publishers that persuades researchers and their institutions that it is worthwhile to pay a premium for publication in a particular journal, rather than paying less for publication in a journal with equally high standards, is the so-called "impact factor". If it were not for this, there would be no reason to pay a premium. The San Francisco declaration of May 13, 2013 [58], signed by more than 150 scientists and 75 major scientific organisations worldwide called for a halt the practice of correlating the journal impact factor to the merits of a specific scientist's contributions. It argued that this practice created biases and inaccuracies when appraising scientific research, and that the impact factor should not to be used as a substitute "measure of the quality of individual research articles, or in hiring, promotion, or funding decisions"

[59]. Nonetheless, impact factors continue to have a stranglehold because of the desires of scientists and their institutions to target this proxy measure of excellence, irrespective of how good a proxy it is, and notwithstanding any pressure it may exert on scientists to “sex-up” their results to ensure publication in high-impact journals. Breaking this habit would do much to reduce the cost of publication

c) Other trends in publishing

There have been two other trends in the last 30 years. Firstly, towards concentration in the scientific publishing market, with a few commercial publishers buying-up smaller publishers, including the publishing arms of learned societies. All of the largest commercial publishers are now based in Europe or North America, and regularly report profit margins of over 30%, funded largely from the contributions of publicly funded libraries and researchers to which they offer bundled journal deals. This unique profitability has continued even as the former costly print-intensive role of publishers in typesetting and formatting has disappeared. At the same time, commercial scientific publishers are tending to re-brand themselves as technology companies, increasingly expanding into all parts of the scholarly research life cycle, including data analytics for ‘impact factors’, university rankings and management of research data. It risks giving bodies whose only accountability is to their shareholders an increasingly monopolistic stranglehold over many of the core components of the scientific enterprise [60].

Secondly, many university systems worldwide have adopted incentives for researcher recognition and advancement based on the number of citations gleaned by their published work. It has generated a global avalanche of publications, with over 3 million scientific papers published per year [61a], less than 5% of which receive significant numbers of citations. Coupled with the asymmetry of the business model, it has released a bonanza for publishers, with more than 30,000 scientific journals [61b], at excessive cost to the public purse, and with little incentive to innovate in the face of current profitability.

The open access movement seeks to replace this system, which it regards as exploitative [62] and inefficient, in ways that serves the public interest better in terms of affordability and functionality in creatively exploiting modern technology.

4.2.2 Open Access publishing

It is largely in response to the above trends that the movement towards open access (OA) publishing has developed. Besides making scientific results freely available on the internet, OA is also a means of ensuring that hypotheses and evidence are most widely accessible so that they might be scrutinised and tested as part of the process of self-regulation of science. The internet has the potential, par excellence, to make results widely available, comprehensible to other societal stakeholders (policymakers, business, professions, NGOs, citizen scientists and citizens), globally affordable, irrespective of the wealth of science systems and publics, and as a means of stimulating creativity to increase functionality (e.g. all papers online, all data online and ensuring that the two interoperate on a stable platform).

It is vitally important to Africa, as it is globally, to be able to negotiate the point of entry into this system at affordable cost in order to participate in the global nexus of knowledge, information, innovation and exchange. The issue of OA at affordable cost is fundamental. In practical terms this has taken two routes, the so-called green and gold [63].

a) **Gold open access**

This is when an author publishes in an online open access journal, with the advantage of making publications freely and immediately accessible. The open content license grants wide-ranging exploitation rights, whilst immediate availability enhances dissemination and the frequency of citation. Publication costs are borne by the author, or by their institution or funding body on their behalf. Other advantages are that of peer review before acceptance for online publication, the readership base associated with the publisher is available to the author, and an author benefits from measurement metrics such as the Impact Factor that adds reputation to their research (see however Box 4.3) [64]. Gold downsides include Article Processing Costs (APC) and the signing away of author's rights to the Online publisher.

b) Green open access

In Green Open Access, the peer-reviewed article is available on the publisher's website, but behind a pay-wall, although authors are permitted to upload their post-print versions on their own institutional repositories. It does not offer the same legal framework for content licensing as in the Gold case. Exploitation is only permitted within the confines of copyright law, which in principle requires an author's contract to be carefully reviewed to enable an article to be re-used in a way that fulfils all the legal stipulations [65]. Its advantages are: that published work is freely available through self-archiving; an author can make the work openly available on an OA repository or their own website while awaiting an open access publisher; it is not incompatible with peer review as most works that are self-archived are peer reviewed prior to publishing [66]. Its disadvantages are the need for additional statements about quality, and the potential to violate the rights of a publisher, with whom an agreement would be needed.

It is important to note that many authors routinely violate copyright agreements by placing published papers on online academic network sites, reflecting a rejection by authors of an existing system that is supposed to protect them. It is clear that the system is no longer fulfilling all the needs of its main market and audience: scholarly researchers and the institutions in which they work.

The debates about costs and open access have inevitably raised the level of rhetoric, to the extent that many rationally unsubstantiated claims have been made. Tennant et al [67] have usefully summarised these (see BOX 4.3).

BOX 4.3 - Ten myths about open access

1. Pre-prints will get your research "scooped". Preprints typically provide a time-stamp and a DOI, thereby establishing the priority of discovery.
2. Journal impact factors are a measure of quality. They are a flawed metric, never intended for this use.
3. Peer review proves you can trust an article. Current peer review has a poor record of finding error. Post-publication review is more efficient in this.
4. The quality of science suffers without peer review. There is no evidence for this. The integrity of the researcher is the key determinant of quality.
5. Open access has created predatory publishers. Wrong, they have flourished because of perversely excessive profits and the flawed market in scientific publishing.
6. Copyright transfer protects authors. They neither protect authors nor benefit scientific progress. They protect commercial publishers' profits.
7. Gold open access is synonymous with the article publishing charge (APC) business model. Most DOAJ-indexed journals do not have APCs. They are funded from other sources.
8. Embargo periods on Green OA are needed to sustain publishers. Traditional publishers can peacefully co-exist with zero-embargo, self-archiving policies.
9. Web of Science and Scopus are global databases of knowledge. Neither represent the sum of global knowledge, excluding much of Africa, Latin America, and South-East Asia.
10. Publishers add no value to the scholarly publication process. They are responsible for key functions, including peer review management and production and archiving final version articles.

4.2.3 The dilemma of affordability

Arguably, the cost of making the results of scientific work freely available as public goods [68] and of funding researchers to publish their work, should not be excessive. In other domains an impact of the digital revolution has been dramatically to reduce cost. In scientific publishing, costs have risen year by year, often at rates far in excess of inflation. It is not difficult to conclude that a grossly asymmetric business model is the cause, one that overturns normal relationships between supply and demand, and where producers (researchers), who are also consumers, have largely permitted commercial suppliers to manipulate the market, although there are increasingly hopeful signs of a "peasants revolt".

Given the inherent difficulties in realising the benefits of OA, a European consortium of funding agencies and councils (cOALITION S [69]) from twelve member states of the European Union have launched Plan S, whose aim is to accelerate the transition to full and immediate open access to all scientific publications by January 2021. The aim is that any publication created from data whose research is financed with public funds should be published/archived on an OA platform and be freely and openly accessible as a public good. Plan S explicitly outlines processes and procedures for compliance [70]. In the plan, the APC would not be borne by the author/s, who would retain their authorial rights, but by funding agencies. The plan advocates publishing and archiving through the gold and green routes. Lately, a new "diamond" route has been developed, in which the author and reader would neither pay for publishing nor reading. Plan S and its short timetable reflects a determined European attempt to achieve a "global flip" towards open access [71]. It requires that papers must be freely accessible from the day of publication with a CC BY licence. If widely implemented, it would mean that legacy publishers (those with long back runs of journals) would have to replace subscription revenues with article-processing charges (APCs). As legacy publishers dominate scholarly publishing, it would lead to a near universal pay-to-publish system, with APCs ranging from several

$100s to over $5,000 per article. The consequences for Africa would be that researchers could freely read research but be largely unable to publish. There have been similar responses in Latin America [72], and calls for regional solutions. India on the other hand has signed up to Plan S, though calling for caps on APCs.

cOAlition S has responded by saying that no one will be unable to publish for lack of funds. Does it mean that Europe would subsidise the global south? Even if it were to, it will continue to be necessary to pay subscriptions for content already pay-walled. Publish-and-Read agreements (PARs [73]) are an interesting development, where rights to access pay-walled content are combined with OA publishing rights. A similar approach is developing in North America where the University of California is seeking to force Elsevier in this direction. Egypt launched a similar portal called the Egyptian Knowledge Bank (EKB) in 2016 [74], which provides free-at-the-point-of-use access to content from international publishers to all 92 million Egyptian citizens.

The Global South has an advantage however as many journals are still government-funded and run by universities, not outsourced to for-profit companies, and are therefore much cheaper to operate. The dilemma is that the historically eminent journals of the Global North, that are largely in commercial hands, tend to dominate global attention to the detriment of the south. An alternative strategy would be to support existing APC free journals, create new ones for the publish element, and negotiate citizen-wide national licensing deals for the read element, using the many institutional repositories that have been established in universities [75]. Despite concerted effort to enlist membership from Africa, only one country, Zambia, has signed up to Plan S [76]. One reason for the resistance has been the view that Africa needs a home-grown alternative, although this is yet to materialise. Valuable insights by Dominique Babini [77] from Latin America may provide a useful exemplar for Africa.

These are issues that the Science Granting Councils will need to address if they are to pursue the opportunities offered by Open Science. It is fortunate that there is increasing dissatisfaction with the current commercial science publishing regime, no lack of radical ideas and an increasing number of potential allies. The International Science Council's new action plan [78] identifies scientific publishing as a priority issue. The Science Granting Councils should ensure that they engage with the ISC over this issue.

4.2.4 The contribution of African knowledge

Despite Africa's surging interest in the Internet and other digital computational technologies [79], its participation in the creation of culturally relevant knowledge is negligible in comparison to the global north. As an example, a landscape survey by the Academy of Science South Africa (ASSAf) on Open Science/Open Data initiatives in Africa [80] reported an estimate of around "0.74% of global scientific knowledge" as Africa's contribution. Several reasons could explain why this may be so: African scientific research outputs are not adequately visible on the internet as they are locked away behind pay-walls, and open access online journals have only begun to make an impact. The African Journal Online (AJOL [81] laments that "mainly due to difficulties of accessing them, African-published research papers have been under-utilised, under-valued and under-cited in both the international and the African research arenas". The internet offers ways of changing this, but many hundreds of worthy, peer-reviewed scholarly journals publishing from Africa cannot host their content online in isolation because of resource limitations and the digital divide.

Consequently, the bulk of global knowledge on the internet comes from the USA, Canada, Europe, China and Australia. It conflicts with the expectation that OA would maximise access and reduce inequalities among the scientific communities. Professor Sanchez-Azofeifa of the University of Alberta, Canada [82] has drawn attention to the fact that open access is now a very exclusive club, dominated by a few developed countries. Unless Africa can create more content, the non-African content will penetrate increasingly deeply into African scientific communities, to the detriment of African context. If OA and publishing are to unleash innovation and creativity in the African scientific community, there needs to be a distinctive African contribution that is highly relevant to its communities. In his keynote address to the Southern African Research Innovation Management Association (SARIMA) conference (March 23 – 25, 2017), Dr. Mangwende, NEPAD agency head of the STI cluster, argued that innovation and cultural context were interlinked, and that Africa could not innovate if she left her cultural context behind. "We need data to innovate, and Africa has a strong narrative and cultural context, so let us use this data to support Africa's development narrative" [83]. This must not however be a monocultural lens. Compared with other continents, Africa's cultural diversity is immense. This diversity should become a strength, not a weakness. Hence, deliberate open access policies across Africa, designed to stimulate participation and penetration of African produced knowledge and innovation are matters of priority. African repositories with major absorptive capacities should be developed, replicating the model of AJOL which has 500 journals across Africa. It would be a good starting point.

4.2.5 Barriers to efficient delivery

In summary there are a number of barriers to efficient delivery of OA that need to be addressed. Some are global problems, some are problems of development:

a) business models of publication with inadequate balances between publisher profit and scientific need;
b) inadequate protocols for access to data and publication;
c) infrastructural and network constraints across Africa, intermittent power supply generally in sub-Saharan Africa, which, unresolved have the potential to severely limit the potential of open science for Africa.

4.3 Open to society

4.3.1 Why?

The public good argument for open data and open access and their roles as stimuli to innovation is fundamental, and the ultimate justification for public funding of science. It is also fundamental to the concept of open science. If scientific knowledge created by the open processes discussed in the previous two sections (4.1 and 4.2) does not in practice find its way into the hands of societal stakeholders to whose work, innovative instinct or pleasure it is relevant, it would be merely “science talking to itself”. If the processes of open access scientific publication fulfilled the crucial condition of deeply disseminating scientific understanding into society, open science would indeed need to entail no more than open data and open access publishing; the duo that we primarily address in this paper. It is self-evident however that scientific publishing does not adequately fulfill this condition, and therefore that open science must be concerned with its openness to society beyond formal scientific publishing.

This issue is deeply relevant to the present era. Unarguably it is an era that needs to hear the voice of science more than ever to tackle many of the profound challenges that global society faces, many of them embedded in the sustainable development goals: but society has arguably become less inclined to listen. The present era is characterized by an increasingly fragmented and polarized political and media environment, in which science is less influential in shaping public opinion than appeals to emotion and beliefs based on personal experience or prejudice. While levels of public trust in science remain relatively high, pervasive digital technologies and the ubiquity of social media enable the widespread dissemination of fake news and of misleading and biased information. This in turn feeds new expressions of science denialism, casts doubt on the need for scientific understanding and interpretation, and threatens evidence-informed decision making in policy and public action. It poses a fundamental – and pernicious – attack on the public value of science, which in turn undermines efforts to build a robust global science system. This affects all scientific fields, all types of research, and all scientific communities around the world. It is of great concern, as our future health and survival depend on the adoption by governments of policies that have a sound scientific basis. The case for a publicly engaged open science is unanswerable.

4.3.2 Societal stakeholders and open innovation

The efficient dissemination of scientific knowledge into society is vital as feedstock for innovation, in government, in business and in society. It is most efficient when two conditions are met: that well-tested scientific findings are made rapidly accessible in the public domain, and that the knowledge is comprehensible to the largest number that may be able to use it in beneficial ways.

The term innovation has come to be used as a jargon term, restricted to commercial innovation. We use it here in its proper sense of creating something new. If society is to flourish and to overcome its many current challenges, the simulation of innovation must be directed to all parts of society. This larger sense of open innovation is important for our common future. Its commercial sense, of promoting an information age mindset towards innovation that runs counter to the secrecy and silo mentality of traditional corporate research labs, is equally applicable to government, and indeed to science itself. A consequence of this view is that knowledge should be openly available in ways that respect the needs and absorptive capacities of all sectors of society and the open data value chain needs to respect both supply and demand.

4.3.3 How?

The role of publicly funded scientists in business-facing innovation has developed greatly in recent years, particularly in universities, and has become widely accepted. The broader public engagement of science has also been strongly promoted. Initially this was badged as "public understanding of science", implying that the central issue was a public deficit in scientific understanding, and all that was required was for citizens to understand more science in order to accept scientifically based pronouncements. The error of this view is powerfully represented by the failure to persuade many in society on issues where the scientific evidence is strong. We now speak of public engagement, a two-way process of dialogue in which science engages more deeply with business, policymakers, governments, communities and citizens as knowledge partners in ways that increase both effectiveness and legitimacy. In the changing world described in 4.3.1, it is increasingly difficult for governments to act on major issues without deeper public consent.

The game-changing development that has the potential to enable such developments to flourish is the modern, global communication network. If the process of developing open science networks in Africa progresses, careful thought is needed about how this might happen, bearing in mind the capacity of the web to spread misinformation. A careful analysis of the parallel development of engagement processes that are sensitive to and capitalize on Africa's cultures is needed, though beyond the immediate scope of this report.

4.3.4 Citizen science

An important development of recent years has been that of so-called "citizen science". This has developed as a mode of scientific research conducted by non-professional scientists. It is frequently carried out in association with formal, professional scientific programmes or with professional scientists [84], [85], [86]. The degree of organisation, embedded within a professional effort, associated with it, or entirely independent, varies greatly, as does the degree of effort or sustained involvement from participants. The most popular are associated with nature in such programmes as iNaturalist [87], eBird [88] and Zooniverse [89].

There is significant African engagement with such initiatives and a number developing from within Africa involve a growing body of people and data, particularly in the domain of nature conservation, for example in the Tropical Biology Association-led "Citizen Science in Africa" programme [90]. An extension of the citizen science model into schools could have major impact on the science literacy of the rising generation. Important though these initiatives are however, they are only part of the "open to society" agenda, the main thrust of which is not interest in science for its own sake, but the needs of society where the engagement in science is a crucial component.

Box 4.4 Recommendations for Science Granting Councils on: Essentials of open science practice

Open Data

- Adopt and mandate standards and create advice capacity for research data management
- Plan for movement towards a FAIR data regime
- Adopt and mandate standards for data citation
- Mandate creative commons licensing for SGC-funded research
- Support an initiative to create a cutting edge, distributed AI capacity
- Create governance structures to oversee ethical data access and use

Open Access publishing

- Engage with international efforts on costs and access for:
- text and data mining
- public access to scientific publications
- scientists' access to publication vehicles
- Create a task force to devise an optimal publishing model for Africa

Open to Society

- Adopt a broad view of innovation priorities – business, governance, society
- Develop a citizen science strategy and its potential for schools

5. The potential of open science systems: case studies of platforms and commons

Chapter 5 describes the diversity of functions and related technical skills that are required to cope with the research data deluge and its diversity. It would be highly inefficient for every researcher or research group working in data-rich fields to develop their own capacities to handle their own data in the ways described above. To do so would either submerge the individual or group in a tangle of data, create confusion in data management thereby undermining the prospect of creating FAIR data, or deter them from working in a data-rich environment, in addition to creating a confusing plethora of incompatible data management systems. It has proved far more efficient at institutional, disciplinary, national, or international levels to scale up the effort and develop well-managed services in the form of open science platforms or commons that serve a wide community.

These recognize that the individual functions described in chapter 4 are inter-related, all parts of system of functions, rather than being stand-alones. Initiatives to create open science or open data platforms or commons are designed to provide more or less seamless provision of support, from IT infrastructure to high-level analytic and AI procedures, and in many cases, not merely in the provision of infrastructural support but also in direct involvement in thematic science priorities. They free domain scientists to concentrate on their immediate priorities rather than acting as amateur data scientists. They may operate at the level of individual disciplines or a wide range disciplines, or at national or regional levels. We here summarise a number of such platforms and the issues from their operational functions, costs, governance, principles and practices, impacts or anticipated impacts, to draw lessons that might be applicable to the African context and the roles of the Science Granting Councils.

5.1 International disciplinary group: bioinformatics ELIXIR programme

ELIXIR is an intergovernmental organisation that brings together life science resources from across Europe [91]. These resources include databases, software tools, training materials, cloud storage and supercomputing access. The goal of ELIXIR is to coordinate these resources so that they form a single infrastructure. This infrastructure makes it easier for scientists to find and share data, exchange expertise, and agree on best practices. Its long-term purpose is to help scientists gain new insights into how living organisms work. ELIXIR includes 23 national members and over 220 research organisations. It was founded in 2014, and is currently implementing its second five-year scientific programme. Its operational structure is based on a series of integrated platforms as follows:

Compute Platform develops ways that researchers across Europe can access, store, transfer and analyse large amounts of life science data.

Data Platform identifies key data resources across Europe and supports the linkages between data and literature e.g. by making it easier to move from a scientific paper to the dataset on which the paper was based.

Tools Platform provides ways for researchers to find the best software to analyse their data.

Interoperability Platform establishes Europe-wide standards that can be used to describe life science data, and makes different data sets easier to compare and analyse.

Training Platform helps scientists and developers find the training they need, and also provides that training.

Communities Platform develops communities, standards, databases and tools in selected life science domains (e.g. Marine Metagenomics, Human Data).

Its governance structure is defined by the ELIXIR Consortium Agreement as:
ELIXIR Board: the highest decision-making body.

Scientific Advisory Board: advises the Board on ELIXIR's scientific strategy and reviews Node applications.

Industry Advisory Committee: gives advice and guidance on industry needs.

Director: responsible to the ELIXIR Board for implementing ELIXIR's scientific programme.

Heads of Nodes committee: consists of the Director and the heads of the ELIXIR national infrastructures (Nodes). The Committee develops ELIXIR's scientific and technical strategy, including its scientific programmes.

The ELIXIR Hub is located at the Wellcome Genome Campus in Cambridge, UK. It accommodates executive management and administrative staff. It is responsible for developing and delivering the scientific strategy, coordinating the services run from the ELIXIR Nodes, supporting governance bodies, working with other biomedical science infrastructures to address the challenges of Big Data, leading communications and external relations activities, supporting the institutions within the Nodes and collaborating with national and European funders and policy-makers. The cost of the ELIXIR Hub over a five-year period, 2014-2018, was £5M.

An ELIXIR Node is a collection of research institutes within a member country, and is responsible for the resources and services that are part of ELIXIR. Each Node has a lead institute that oversees the work of that Node. The Norwegian node for example, comprises a lead institute at the University of Bergen, together with four other institutes. The Nodes build on the strengths of the scientific communities of that country. The European Molecular Biology Laboratory is an intergovernmental organisation and the only Node not associated with a specific country.

ELIXIR has a mixed funding model with contributions coming from a number of mostly public sources. The Hub is funded through membership fees paid by member countries. Nodes are typically funded through national-level investments, supporting national coordination, and the development and operation of services. Its science programmes compete for grant funding from the European Union, national funding bodies and some international funders (e.g. US National Institute of Health). Some Nodes are able to access European Union Structural Funds that are allocated to developing areas within the Union.

5.2 International Disciplinary Group: Pan African Bioinformatics Network for the Human Heredity and Health in Africa - H3ABioNet (89)

It is particularly helpful to contrast the preceding European effort to create a major open science bioinformatics enterprise with an analogous effort in Africa. H3ABioNet [92] was established to develop bioinformatics capacity in Africa and specifically to support genomic data analysis by H3Africa researchers across the continent. It develops human capacity through training and support for data analysis, facilitates access to informatics infrastructure by developing or providing access to pipelines and tools for human, microbiome and pathogen genomic data analysis. Its mandate is to develop and roll out a coordinated bioinformatics research infrastructure that is tightly coupled to a sophisticated pan-African bioinformatics training programme (90).

The development and application of effective genomic medicine is heavily dependent upon the ability to aggregate and analyze large data sets and to interpret and disseminate knowledge across multiple biomedical disciplines. In Africa, there are few centres of expertise where large numbers of clinicians, genome scientists, and bioinformaticians are situated to jointly perform competitive genomic medical research. As part of a strategy to develop critical mass through intra-African collaboration, as echoed in the STI 2024 Strategy for Africa (see also 3.3), African bioinformatics groups have, over the last 10 years, been collaborating to develop the capacity to perform globally competitive research on public and local data sets, in spite of the geographical distances separating them [94]. These efforts recently received a major funding boost that catalyzed the nascent African genomics research community through the creation of H3Africa Bioinformatics Network (H3ABioNet), which was established with a grant from the US National Institutes of Health (NIH) Common Fund, as part of its contributions to the Human Heredity and Health in Africa initiative [95], [96].

The consortium is based on a system of collaborating nodes. The network, which is run from a central node at the University of Cape Town, consists of more than 30 nodes across 15 African countries [97] with one partner in the United States and one in the United Kingdom. The institutions range in their current capacity from full nodes with a track record in bioinformatics research, training, and support; through associate nodes with some bioinformatics activities; to development nodes with little or no bioinformatics capacity. Altogether, the network funds more than 40 staff and students and includes more than 80 additional members who contribute to H3ABioNet activities. The nodes collectively provide excellent expertise in different areas of bioinformatics including functional genomics, human population genetics, GWAS and NGS analysis, microbiome analysis, SNP linked protein structure analysis, and biomedical and clinical data storage and management.

The consortium faces a number of high priority challenges that need to be overcome to enable genomics research and competitiveness on the continent. These include, poor internet connectivity, data access, transfer and remote computing; lack of significant computing infrastructure for data storage and processing; lack of bioinformatics skills in clinical genetics and genomics teams performing genomics research; and disparate pockets of bioinformatics expertise across the continent. There is an important contrast here with the ELIXIR programme. Although both are designed to work on analogous issues for which the approaches of open science are essential, ELIXIR can depend on high levels of computational, networking and cloud capacities that are provided by European states and the European Union as a matter of course for their science systems, whereas H3ABioNet has to confront these issues itself and throughout its network, and to perennially make the case for their development. With ELIXIR, the case is already accepted at national and European Union levels such that their requests for development are accepted as parts of ongoing science system planning processes.

Major objective of H3ABioNet therefore are to develop human resources through the training of bioinformaticians and researchers in computational techniques and to develop a robust, continent-wide research infrastructure that provides access to bioinformatics tools, computing resources, and technical and data management expertise. Network activities are being achieved through dedicated working groups and task forces comprising representatives from multiple countries. Full nodes, including those situated abroad, are helping to build capacity in the less resourced nodes, thus, ensuring the transfer and dissemination of knowledge and skills within Africa. Some nodes have already or plan to set up their own bioinformatics centres dedicated to training and research in bioinformatics.

Long-term sustainability is a key objective of the network but is not realistically achievable within the first five years of its existence. However, the project has increased computing facilities and provided for eBioKits in many nodes, which will remain in place beyond the end of the five-year project. In Egypt, H3ABioNet funds facilitated establishment of an eBioKit-based computer laboratory connected to the internet. Joint collaborative project funding proposals are being developed using bilateral agreements between some of the participating countries, and multi-institution research project proposals have been submitted to funding agencies in response to specific calls.

5.3 International, multi-disciplinary open science system: European Open Science Strategy (https://ec.europa.eu>research>openscience)

The European Union research strategy recognises an ongoing major transition in how research is performed and how knowledge is shared. In response it has adopted an ambitious strategy that seeks to make open science a reality across all its member states.

It contrasts with the two previous examples in being a top-down policy-driven initiative in contrast to being science-driven, although scientific researchers are involved in advising on its policies. The other contrast lies in its being designed to address the interests of a wide range of varying needs from the whole science community such that no single science agenda that is able to attract enthusiasms of a well-defined disciplinary group is particularly targeted. The Commission does however have a powerful means of persuasion in the form of:

- its annual science budget of multi-billion euros [98], part of which can be targeted on developing take-up by researchers of its open science priorities;
- its ability to provide access to high end computing and cloud facilities which can again be conditional on adherence to open science practices;
- its power to require data created through Commission-funded programmes to be deposited in open access repositories and to publish in open access journals as conditions of grant, and as will be mandated by its projects.

The Commission has set in place several open science policies and mandates in its framework programmes, particularly:

- strategies for promoting economic growth, job creation, transnational cooperation, access to and transfer of scientific knowledge;
- policies promoting open access to scientific publications and data;
- recommendations to EU Members States to implement open access policies;
- strategies to promote citizen science in more inclusive, transparent and accessible ways.

The problems that its policies face are common ones:

- incentives in national systems that do not align with EU priorities;
- difficulties in achieving interoperability between diverse data streams and managing heterogeneous data systems that are particularly prevalent in some disciplines;
- discrepancies between per capita funding and the maturity of different European science systems;
- governance that is contained within the European Commission, and which does not necessarily reflect national priorities.

The strategy's s component parts are:
The Open Science Policy Platform with the role to advise the Commission and act as a consulting body for all European open science policies and the development of a Science Policy Agenda to radically improve the quality and impact of European science across member states and internationally.

The European Open Science Cloud designed to provide a public data repository which conforms to open science values. It is projected to become a reality by 2020. It aspires to be Europe's virtual environment for all researchers to store, manage, analyse and re-use data for research, innovation and educational purposes. It is also intended that data submitted to the system should progressively conform to FAIR data principles.

Open Access Publication policies that require all projects receiving Horizon 2020 funding to make sure that any peer-reviewed journal article that they publish is openly accessible, free of charge.

The EU Citizen Science Platform is designed to support the activities of individuals and groups wishing to undertake citizen science projects. It will be interesting to observe how the projects undertaken on this platform evolve.

The structure of governance of the European Open Science Cloud (EOSOC) is planned to be based on three layers:
Strategic Layer comprising a board that combines state-of-the-art expertise on scientific cloud infrastructures with the Funders and Policy Makers. It will therefore include EU Member States and Associated Countries representatives. It will mainly make strategic decisions on the development and evolution of the EOSOC.

Executive Layer comprising an executive board to manage day-to-day operation of the EOSC and procurers, and designing and planning work-related future developments. It is the only full-time staffed layer, will be supported by Working Groups, and will have the responsibility of ensuring that user needs are met and strategic requirements addressed.

Stakeholder Layer organised in the form of a stakeholders' forum to provide a medium for stakeholders (users (consumers), providers and Intermediaries of EOSC Resources). This would have the main role of discussing, supervising and channelling communication between the EOSC and the communities across all three layers.

5.4 Potential lessons for an African initiative

It is important to recognise that the purpose of this report is to explore whether there are benefits to African science and its application in developing open science approaches, and if so to suggest how this might best be done. The examples above are of systemic rather than piecemeal approaches. They are based on strong collaboration and common purpose. The way that they integrate resources from an international network to support a common infrastructure embeds cost efficient economies of scale that is a potentially attractive example to Africa and to the interests of the SGCI. They are a demonstration of the case made for cost effectiveness in 3.3. They have also been cost effective in generating good science and stimulating innovation. How is this best achieved?

5.4.1 Enthusing scientists

Even when they work in teams, as they increasingly do, scientists tend to be driven by an individualistic curiosity for discovery in their chosen fields, not by the desire to use novel technologies nor develop new ways of collaboration. They are means to ends rather than ends in themselves. In the cases of ELIXIR and H3ABioNet, successful bottom-up developments have occurred because of the overt potential of openness and a shared technological capacity to achieve scientific ends. The European Open Science initiatives represent work in progress where it is too early to judge success. In some overtly data-intensive fields, such as high-energy physics or cosmology, the Open Science Cloud is a powerful stimulator of enthusiasm because of the immediate potential it offers for discovery. In others, the pathways to scientific discovery through open science are more arduous because of the technical complexity of the data-intensive challenge. This is where the financial leverage of the European Commission is a potent driver of behaviour.

In the African case, the financial leverage of the Science Granting Councils, though significant, is proportionately less than its European Commission counterpart, as much funding for science comes from outwith the continent. We have little doubt but that scientific potential must be a key driver, in which the development of open science practices goes hand-in-hand with funding of priority issues in ways that strongly favour intra-African collaboration and deliver the benefits summarised in 3.3. Obvious science priorities are such as those identified in the 2024 STI Strategy for Africa, such as the burden of disease [99], sustainable agriculture, resilient cities, disaster risk reduction etc., where intra-African collaboration supported by open science processes have the potential to create the virtual critical masses of effort and engagement of funders that could yield substantial benefits in these fields as well as creating powerful capacity in African science systems. Such programmes, embedded in open science practices, should be designed not only to deliver value in themselves in the specific field, but also to act as demonstrators to the wider African community of governments, policymakers and scientists of the value of the open science approach and more important still, to act as stimulus to vitalize African science.

5.4.2 Enthusing engagement through inclusive structures

ELIXIR and H3ABioNet have a similar structure: a central hub with responsibility to plan overall strategy, service governance and coordinate the network; and national nodes that support work in national institutions in ways that are sensitive to the level of national science system maturity. This structure seems highly relevant to the African need, delivering a common supra-national strategy whilst being sensitive to national issues. For nodes, a careful strategic balance would need to be struck between national priorities, the potential of a network to deliver the high-level, long-term benefits of stimulation of intra-African collaboration and the development of virtual critical masses.

The hub/nodes structure could also be one that is well attuned to funding potential. With major international funders such as the World Bank and Development Agencies potentially being able to support a hub together with national contributions, and nodes being funded nationally and through external funders that traditionally fund specific countries. Such a structure could also map well onto the coordination patterns necessary for essential collaboration with National Research and Education Networks (NRENs).

5.4.3 Addressing specific barriers

There are a series of barriers that are Implicit in the examples above, and in the long experience internationally of providing common resources for science, that need to be taken into account when creating such systems.

a) **Trust/Competence.** It is inevitable that some degree of centralization of effort is required, whether a particular facility (e.g. high performance computing, cloud management, expensive experimental facility) is located in one place or whether (as in ELIXIRor H3ABioNet) it is a node in a network of nodes. There are several fundamental requirements:
   - that the centre or node has relevant expertise;
   - that it is trusted by partners;
   - that it has a highly professional management;
   - that is has efficient and effective governance structures.

   The centre or node must operate for the benefit of the partnership and not primarily for its own local advantage. Its remit and structure of governance must act and be seen to act to the benefit of the network and in response to agreed priorities. Access to the facility's capacities should favour all partners equally.

b) **Connectivity.** The developing network should therefore prioritise minimal levels of effective connectivity. This is a major challenge for Africa where connectivity is a non-negligible barrier for access to science materials because of low internet access rate. E.g., in Sub-Saharan Africa, only 22% in average (range = [1 - 59%]) of the population have access to the internet [100].

c) **Data protectiveness.** Not all institutions or states in Africa are ready to go from protecting data to offering open access. Even if open science has recognized benefits, it is also the case that individual researchers will feel threatened by such openness. Researchers generally share their data if they have guarantee or if they feel to be in a win-win collaboration (e.g., recognition of their work in the resulting publications, exchanges in terms of knowledge, experimental protocols or equipment allowing them to raise their technical level).

d) **Intellectual property (IP).** Different states have different IP regulations or laws that influence data use and sharing, and varying approaches to copyright. Ideally they need to be homogenized or to be brought under a common convention that minimizes barriers to exchange and sharing.

e) **Confidentiality.** Data on human subjects in particular is a sensitive matter that will need strong, agreed regulation within any open science initiative (see 4.1.7).

f) **Language.** Sub-Saharan Africa is home to four non-African languages, English, French, Portuguese, Arabic and many African language groups. An open science initiative will need to take this issue seriously, partly because much of Africa's meaningful production of knowledge for innovation cannot be readily separated from its indigenous linguistic and cultural contexts (see discussion in 4.2.2).

g) **Security.** Protective security measures are vital in preventing unauthorized access to computers, databases and websites and in protecting data from corruption. It will be necessary to build conventional and secure data sharing infrastructures to promote exchange [101].

h) **Incentives.** Many of the current incentives for academics are effective barriers to open science. They are discussed further in 7.3.2.

Box 5. Recommendations for Science Granting Councils on:
A systemic approach to open science in which its tools are embedded

- Collaborative systems (platforms or commons) are powerful mean of delivering the essentials of open science practice
- They create:
  - Efficient and accessible services
  - Economies of scale
  - Impact and voice through intra-national coordination of effort
- They should also focus on:
  - Major programmes of pan-African relevance to enthuse scientists and create virtual critical masses
  - A hub and network of nodes to ensure both national commitment and effective coordination
  - Recognise and address the practical issues that perennially arise in providing services for science

6. The evolving landscape of open science in Africa

Chapter 6 describes the purposes and structures of currently operational open science enterprises and draws a number of lessons such as establishment of collaborative systems that form a hub that is efficient, focuses on key agenda of the region and at the same time creates economies of scale. These practices are potentially applicable to an open science initiative that the Science Funding Councils might choose to launch which would build on some of the experiences of open science in Africa and apply lessons learned elsewhere. Figure 6.1 shows examples of open science initiatives in Africa. We now assess the evolving landscape in sub-Saharan Africa on which an initiative would need to build.

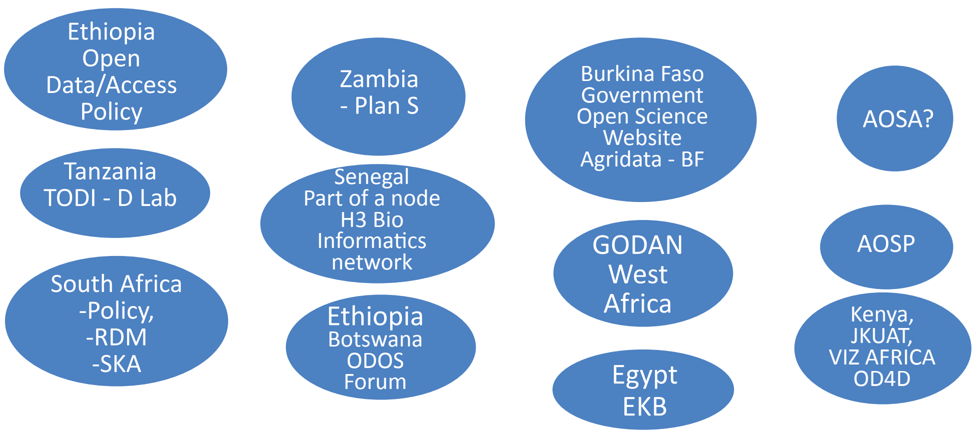


Figure 6.1 Examples of Open Science Initiatives in Africa

6.1 Science in sub-Saharan Africa

Sub-Saharan Africa spends less than 1% of the global expenditure on research and development (Figures 6.1 and 6.2). Latin America and the Caribbean account for 3%; Europe 27%; Asia 31%, and North America 37% [102]. The result is that Africa contributes a meagre 0.74% (Bank, 2014) to the global research output, although a World Bank Study showed a marked increase from 0.44% in 2003 to 0.72% in 2012. African countries' expenditure on research and development is low as a percentage of their GDP [103], which contrasts with their developmental goals and aspirations [104]. Although Sub-Saharan Africa gained an additional percentage point of world population between 2007 and 2013 (to 12.5%), its gross domestic product (GDP) grew by just 0.3% and gross domestic expenditure on R&D (GERD) by just 0.1% [105]. This has left very little lee-way for the funding necessary to enhance the performance of their universities and other higher learning institutions and to develop the capacities that contribute to development and innovation in their economies.

However, several countries have seen strong growth in their scientific production, including Ethiopia, Ghana, Mozambique and Rwanda. Although South Africa accounted for 46% of sub-Saharan Africa's publications in 2014, low-income countries such as Benin and Gambia have scientific productivity levels (articles per million inhabitants) comparable to those of middle-income economies. Ethiopia (0.61% in 2013), Kenya (0.79% in 2010) and Mali (0.66% in 2010) have all increased their R&D effort (GERD as a percentage of GDP) in recent years to the level of a middle-income economy. Malawi's reported commitment of 1.06% of GDP to R&D is questionable given current constraints on national budgets. However, if these statistics were to be relied upon, Malawi would claim the highest ratio in Africa with its scientists publishing more in mainstream journals – relative to GDP – than any other country of a similar population size [106].

Although these latter developments are encouraging, a consequence of low investment is that sub-Saharan countries consume research outputs from outside the continent, but contribute very little from their own resources. They depend heavily on international collaboration and visiting academics for their research output.

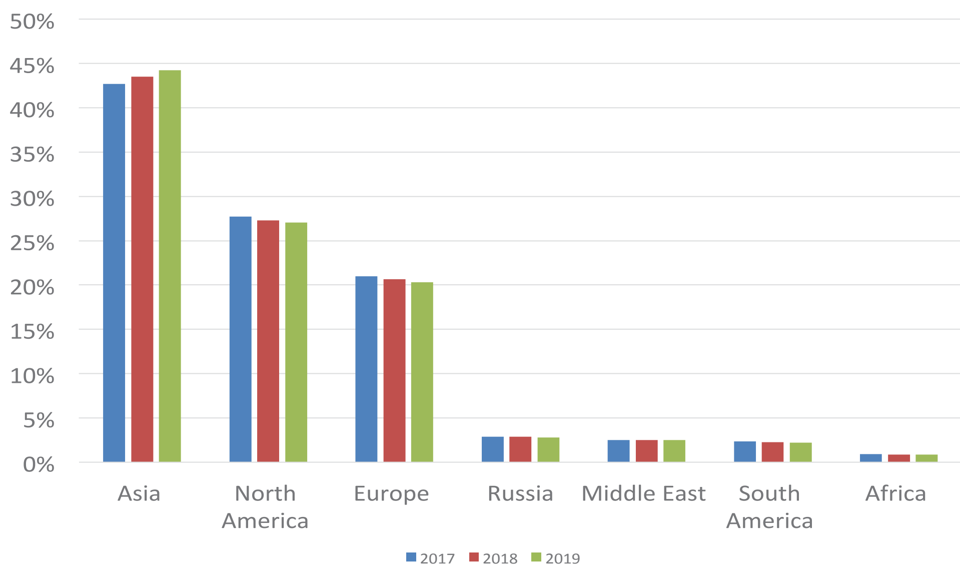


Figure 6.2. Regional contributions to global R&D expenditure

Such collaborators bring grants and technical expertise to complement the work of local counterparts. In 2012 for example, southern Africa, east Africa, and west and central Africa produced 79%, 70%, and 45% of all their research output, respectively, through international collaborations (see also Fig. 6.4). Ironically, intra-Africa collaboration remains poor. World Bank data show that collaboration among local researchers in sub-Saharan Africa range from 0.9% in west and central Africa to 2.9% in southern Africa [107]. Observers of African researchers attempting to work together cite several barriers to intra-Africa collaborations that span from the geographical, to the political, linguistic [108] and financial.

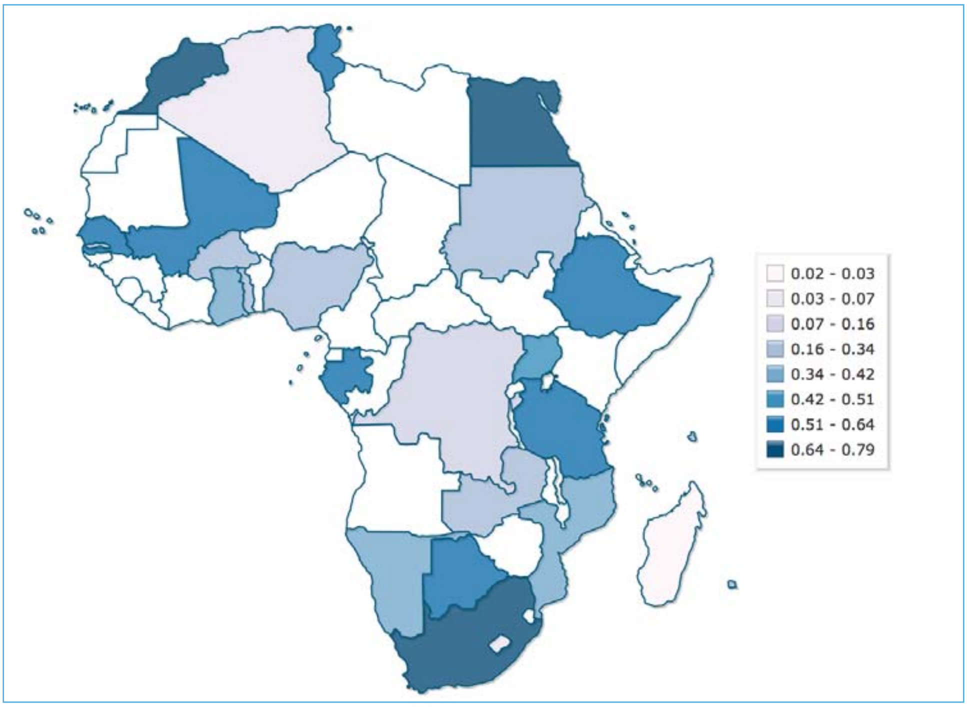


Figure 6.3. Gross Expenditure on R &D as a proportion of GDP by country in Africa

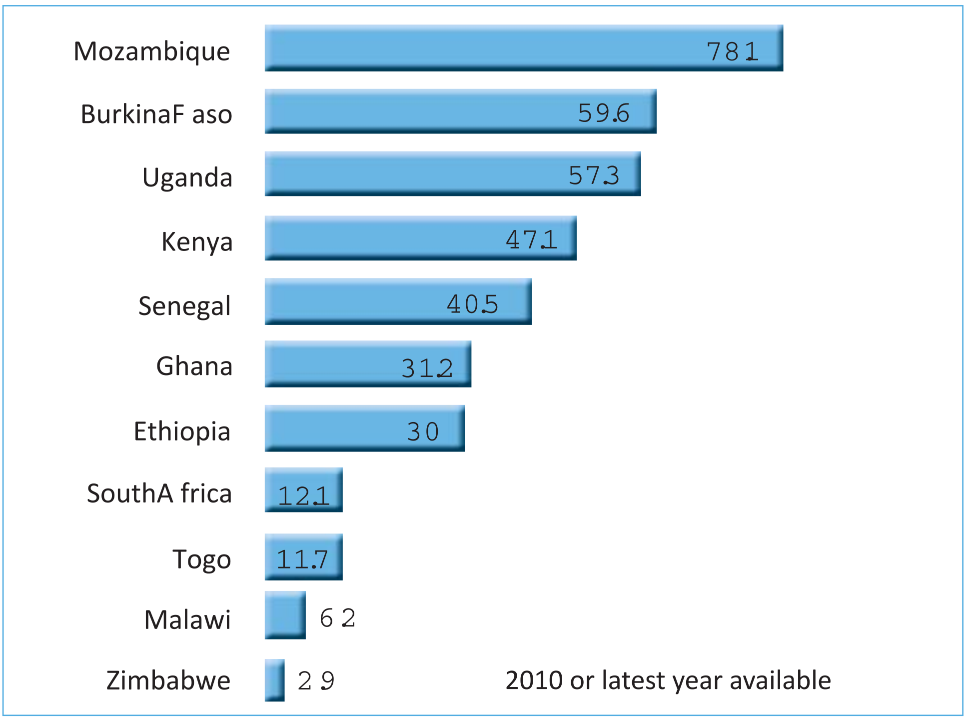


Figure 6.4. Proportion of funding for selected African states R & D from outside the continent (World Bank)

Notwithstanding the sustained efforts of some donors in support of specific countries, or on specific science priorities, the overall impact is arguably less than it might be. Donor aid is likely to reflect the donor's priority rather than necessarily responding to the recipient's top priority, much is relatively short term and project-based rather than sustained and strategic, and many donors required a recipient contribution, typically of the order of 20-40%, which many potential recipients are unable to afford. It would be a great improvement if the Science Granting Council's initiative could develop by presenting a coherent view of an optimal medium to long-term science strategy that might best serve Africa's needs, by engagement with external donors through a joint forum. These are issues to which we return in chapter 8.

That being said, there are scientific highlights in sub-Saharan Africa. South Africa stands out as having the greatest number of researchers per million inhabitants and by far the greatest output in terms of scientific publications and patents. With nearly a third of their publications in chemistry, engineering, mathematics and physics, South Africa and Mauritius stand out as being more akin to developed countries than the other countries of the Southern African Development Community (SADC) or of sub-Saharan Africa as a whole, where research tends to favour bio-medical science and geosciences [109], reflecting priorities for health and primary mineral extraction respectively. The African Institute for Mathematical Sciences (AIMS) [110] is a pan-African initiative that has been highly creative in forming excellent scientists. The Next Einstein Forum [111], that is pan-African in inspiration, shows that Africa has the talents that it needs. What seems to be needed are institutional frameworks and science structures across Africa that create opportunities and provide continuing support to nourish the talents of the continent and support intra-African initiatives. Figure 6.5 illustrates important elements of the developing institutional framework in Africa. Additionally, it is vital to have recognition by African governments, both individually and collectively, that support for the science base and associated education is the single most vital investment that African society can make for its future vitality. It is here that the Science Granting Councils can be so vital, in representing a powerful and coherent view to their governments and to external donor agencies of the priorities for African science, whilst coordinating their mutual priorities within the states that they represent.

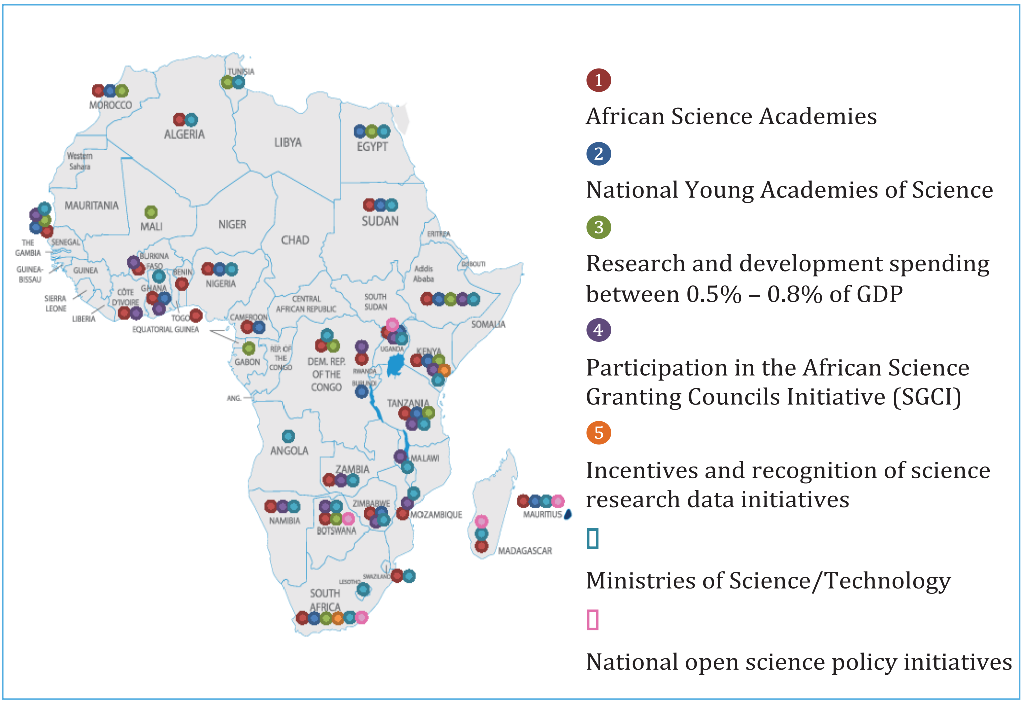


Figure 6.5. Institutional commitments to science in Africa (African Open Science Platform pilot)

6.2 Sub-Saharan Africa's status in the open science movement

The pattern of engagement with the global open science movement and Africa's place in this developing patchwork (2016) is shown in Figure 6.6.

Figure 6.6. The Global Open Data Barometer 2016 (108)

The basis on which the African assessment was created is:

a) Operational open science projects of international significance:

- The H3ABionet project (H3ABioNet, 2019 that we describe in 5.2.
- The South African National Biodiversity Institute (SANBI) is a major node of the Global Biodiversity Information Facility [112] which hosts biodiversity information to make it freely available on the internet so that policy makers, managers and researchers can make well-informed decisions that contribute to sustainable development
- . DataFirst is the only African database that has the CoreTrustSeal of the International Science Council's World Data System [113].

It provides a trusted repository service for SA and other African users, with training and research on the quality and usability of data.

- AAS Open Research [114] is a platform for rapid publication and open peer review for researchers. It enables researchers to publish any research they wish to share, supporting reproducibility, transparency and impact. It uses an open research publishing model, including all supporting data, reanalyses, replication and reuse. Key benefits include: all types of research can be published rapidly, standard research articles, clinical trial findings, systematic reviews, study protocols, data sets, results, and case reports. It supports research assessment based on the intrinsic value of the research, not the venue of publication, and reduces the barrier to collaborative research through data sharing, transparency and attribution.

b) Active, sectoral initiatives with the potential to contribute to a major development:

- ICT development: NRENS, SADC cyber-infrastructure roadmap, high-performance computing facilities in 10 countries.
- Data science courses in 15 HE institutions, of which 6 are in SA.
- Open Access/Data declarations or agreements endorsed by 12 governments.
- 63 Research data repositories, of which 24 registered with re3data.org.
- Open data awards in 2 countries.

c) Projects in development with major potential:

- The African component of the Square Kilometre Array (SKA) is based in South Africa and involves 8 African national partners (South Africa, Botswana, Ghana, Kenya, Madagascar, Mozambique, Namibia, Zambia) [115]. It is developing an African Data Intensive Research Cloud and the associated skills needed to cope with the vast big data streams to be produced by the astronomical programme.

- Indigenous Knowledge and Climate Change Adaptation Research Project among the Griqua and Nama peoples in South Africa [116]. It includes participatory action research ("PAR") design and methods with the aim of promoting open science by reducing the power relations within and between researchers/researched. PAR takes a "bottom-up" approach by developing partnerships with communities to identify key issues of importance and find means of conducting research, interpreting results, and acting on the findings [117].

- The African Open Science Platform (AOSP) has the mission to put African scientists at the cutting edge of contemporary, data-intensive science. It is developing an integrated approach involving a federated hardware, communications and software infrastructure, including policies and enabling practices to support open science in the digital era, and a network of excellence in open science that supports scientists and other societal actors in accumulating and using modern data resources to maximise scientific, social and economic benefit. It plans for an operational launch in 2020 (see Box 6.1).

- The World Bank project: The Digital Economy for Africa [118]. This is a continent-wide initiative that has five pillars, including Digital Infrastructure, Digital Skills, Digital Platforms, Digital Finance and Digital Entrepreneurship. The World Bank has committed to lend $25 billion up to 2030 to contribute to the overall goal of making every African individual, business and government "digitally enabled".

The fundamental question is whether the globally weak scientific performance of Africa could be radically improved through the adoption of a pan-African open science initiative. Could a powerful open science ethos have a major impact? Its essence would be to provide a nurturing frame for developing creative common strategies, removing national boundaries as siloes for scientific policy and practice and stimulating intra-African collaboration as a means of creating virtual critical masses of researchers on important common problems. It is an issue we will return to in chapter 8. We also comment that it was precisely such a collective approach that enhanced the creativity of Europe to become a scientific super-power.

In this context we now explore the open science landscape through the lens of the three domains described in chapter 4 that we consider to be the building blocks of open science for the modern age: open data, open dissemination of scientific results and open to society. Sections 4.1, 4.2 and to some extent 4.3 provide a template and check-list against which the state of open science in Africa can be assessed.

BOX 6.1: THE AFRICAN OPEN SCIENCE PLATFORM

Its building blocks are:

- a federated hardware, communications and software infrastructure, including policies and enabling practices to support open science in the digital era;
- a network of excellence in open science that supports scientists and other societal actors in accumulating and using modern data resources to maximise scientific, social and economic benefit.

These objectives are to be realised through six related strands of activity:

Strand 1: A federated network of computational facilities and services.
Strand 2: Software tools and advice on policies and practices of research data management.
Strand 3: A Data Science and AI Institute at the cutting edge of data analytics.
Strand 4: Priority application programmes: e.g. cities, disease, biosphere, agriculture.
Strand 5: A Network for Education and Skills in data & information.
Strand 6: A Network for Open Science Access and Dialogue.

6.3. Open Data

We briefly map the open data landscape by considering four essentials: principles and policies, infrastructure, skills, and processes and procedures.

6.3.1 Open Data policies

According to the Registry of Open Access Repository Mandates and Policies (ROARMAP), there are currently 31 institutional open access policies registered across Africa. Examples of an institutional open research data policies are those of: the Jomo Kenyatta University of Agriculture and Technology (JKUAT) Open Research Data (ORD) Policy; the regional open data policy of the City of Cape Town, and the South Africa Open Data policy, which addresses government data. A further, intergovernmental, discipline-specific policy level, is exemplified by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES).

Examples of international agreements that are in favour of open science developed or signed by African states include:

- Africa Data Consensus International (G8) Open Data Charter (adopted by Sierra Leone; endorsed by institutions worldwide)
- UN Agenda 2063 (55 African member states)
- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (17 African countries; 51 signatories)
- Budapest Open Access Initiative (signatories from multiple African countries)
- Cape Town Open Education Declaration (32 African countries; 533 signatories
- Dakar Declaration on Open Science in Africa (Sci-GaIA) (12 African signatories) (ASREN; CUBRe (Nigeria); DIT (Tanzania); Eko-Konnect (Nigeria); GARNET (Ghana); MaliREN (Mali); NgREN (Nigeria); RENU (Uganda); RITER (Côte d'Ivoire); TogoRER (Togo); Vice Chancellors of Ghana; WACREN)
- ISC Accord on Open Data in a Big Data World (signatories Kenya-JKUAT and others)

- Kigali Declaration on the Development of an Equitable Information Society in Africa (signed by 27 African countries and 4 intergovernmental organisations)
- Open Data Barometer (28 African countries) (Tunisia, Egypt, Morocco, Benin, Cameroon, Ethiopia, Malawi, Mozambique, Rwanda, South Africa, Togo, Zimbabwe, Botswana, Côte d'Ivoire, Ghana, Mali, Namibia, Senegal, Swaziland, Uganda, Burkina Faso, Democratic Republic of the Congo, Kenya, Mauritius, Nigeria, Sierra Leone, Tanzania, Zambia)
- Open Knowledge Foundation Network (OKFN) (Established groups - Burkina Faso; Incubating groups - Cameroon, Ethiopia, Senegal; Affiliates - Nigeria; Hibernated groups - Egypt, Kenya, Morocco)
- African Center for Technology Studies (ACTS) Charter (4 African signatories - Ghana, Kenya, Malawi, Uganda)
- Good Governance Africa (unknown)
- The Principle of Universality of Science and Academic Freedom (28 African members): Angola, Botswana, Burkina Faso, Cameroon, Côte d'Ivoire, Egypt, Ethiopia, Ghana, Kenya, Lesotho, Madagascar, Malawi, Mauritius, Morocco, Mozambique, Namibia, Nigeria, Rwanda, Senegal, South Africa, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zambia, Zimbabwe
- Open Government Partnership Declaration (12 African participants: Morocco, Tunisia, Senegal, Sierra Leone, Liberia, Côte d'Ivoire, Ghana, Burkina Faso, Nigeria, Kenya, Malawi, South Africa0
- Nairobi Data Sharing Principles (Kenya, Madagascar, other)

The National Research Foundation (South Africa), as a science funder, issued in 2015 a Statement on Open Access to Research Publications from the National Research Foundation (NRF) [119] making depositing data sets available a requirement (119).

6.3.2 Open Data infrastructure

Open data infrastructure has two fundamental components, access to powerful computational and storage capacity through high bandwidth networks, and effective systems of data repositories. The former is a necessary pre-requisite, the latter is a powerful source of strength, both are needed if African scientists are to engage effectively with data-driven science and African society is to benefit from the application of that science to the opportunities and challenges that it faces.

Networks must enable access to a spectrum of shared resources including the provision of Cloud systems connected by large bandwidth Wide Area Networks (WANs) that host software systems that enable data analysis and provide access to massive data collections. African higher education and research institutions rely on National Research and Education Networks (NRENs) to provide connectivity and specialised services. They are key parts of the landscape and promote collaboration among member academic institutions and in sharing infrastructure, content and high-end ICT talent. They vary in their level of maturity as shown in the table 6.1. Level 6 NRENs offer numerous value-added services such as videoconferencing, federated identity management and wireless roaming services. There is a well-established culture of collaboration amongst NRENs. The pan-European GÉANT programme is also working to strengthen Europe's links with the African continent and to provide African research and education communities with a gateway for global collaborations.

NRENs also provide services to schools and Technical and Vocational Education and Training (TVET) institutions, e.g. SABEN (South African Broadband Education Networks) [120]. School networks can also request to be connected via SABEN, but need to provide funding for this. Another partnership between an NREN and schools is the KENET Schools Connectivity Initiative (SCI) [121] that coordinates various commercial, educational and government organizations interested and willing to provide Internet access and promote the use of ICT in Kenyan schools. The SCI is a platform through which public and private sectors partner in an effort to provide scalable and sustainable ICT and Internet access to schools. The SCI model is based on a holistic approach that integrates Internet connectivity, Internet access, relevant educational content and capacity building for teachers.

Level 0	Central African Republic, Djibouti, Republic of the Congo, Lesotho, Libya
Level 1	Angola, Comoros, Eritrea, Seychelles, South Sudan, Equatorial Guinea, Guinea Bissau, São Tomé and Príncipe
Level 2	Botswana, Democratic Republic of the Congo (2.5), Malawi (2.5), Mauritius, Rwanda, Somalia, Swaziland, Zimbabwe, Cape Verde, Chad, Gambia, Guinea, Liberia, Sierra Leone, Mauritania
Level 3	Benin, Burkina Faso, Cameroon, Gabon, Ghana (3.5), Mali, Niger, Togo
Level 4	Burundi, Ethiopia (4.5), Madagascar, Mozambique (4.5), Namibia, Sudan (4.5), Tanzania (4.5), Côte d'Ivoire, Nigeria, Senegal, Morocco (4.5), Tunisia (4.5)
Level 5	Uganda, Zambia
Level 6	Algeria, Egypt, Kenya, South Africa, [Zambia – 2019]

Table 6.1 Levels of maturity of NRENS in African states. Level 0 is the lowest, Level 6 the highest.

High capacity networks are being established through AfricaConnect2. This comprises three geographical areas (clusters) and involves the respective regional NRENS (Figure 6.6):

- ASREN in North Africa (connecting the Arab countries as well as Algeria, Djibouti, Egypt, Libya, Mauritania, Morocco, Somalia, Sudan, Tunisia)

- WACREN in West and Central Africa (Ghana, Côte d'Ivoire, Togo, Niger, Nigeria, Cameroon, Mali, Chad, Guinea, Sierra Leone, Burkina Faso, Senegal, Gabon, Benin)

- UbuntuNet Alliance in Eastern and Southern Africa (Burundi, Democratic Republic of the Congo, Ethiopia, Madagascar, Kenya, Malawi, Mozambique, Uganda, Rwanda, Somali, Sudan, South Africa, Tunisia, Namibia)

There are serious problems that stand in the way of developing the NRENs in ways that would permit them to provide the level of networked services that African needs. They receive limited support for operational expenditure (OPEX) from their governments and are poorly understood by the telecom and Internet community, where they are regarded merely as specialised Internet service providers that have to compete with very large telephone companies. Big Data requires sufficient bandwidth and stable and reliable Wide Area Network (WAN) connections, whereas universities and research institutions have very low WAN and Internet access budgets and many areas in Africa struggle with ageing and unreliable power infrastructure and frequent power outages.

Cloud services require expensive high-speed network access and should be connected to NRENs, but this is undermined by the lack of funding for hardware in support of data sharing and commercial ISP offers that are too expensive for Africa. Researchers are often unaware of the availability of Open Source Software tools/applications to collaborate and to share data as part of open science. Data security is a huge concern, but despite this, many researchers store data on their own laptops and workstations running operating systems that are highly vulnerable to viruses that aim to expose their data.

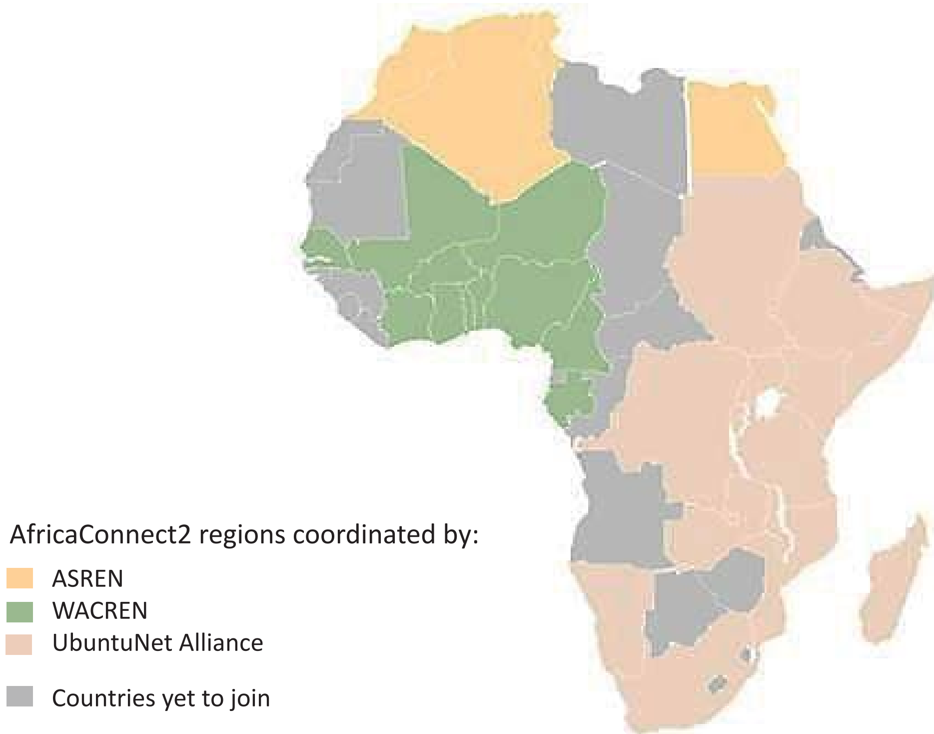


Figure 6.7. National Research and Education Networks' (NRENs) Alliances across Africa. (Africaconnect2.net, n.d.)

The potential for development of the present computational and communication network is very clear. The cable communication network around Africa is as good as that serving any other continent (see figure 6.9). It is the internal network that is deficient and requires major up-grades. Whilst telecommunications companies see a great opportunity in Africa for an extension of their services, these are not likely to be of the character required by the research community. Experience shows that for a cost-efficient research system to develop, key parts of any network need to be publicly managed in ways that do not make the system dependent entirely on any single providers' services. Extension and development of the NRENs, coupled with efficient wide area networks and Cloud computing are clear priorities.

In this context, professionally managed, trustworthy data repositories are vital. The development of open databases in Africa falls well short of what is needed for an effective open data ecology that is able to support competitive data-intensive scientific enquiry. A devastating example of this lack was provided by the 2013-2016 Ebola outbreak (figure 8a-b). In response, and in search for effective treatments, clinical, epidemiological and laboratory data on tens of thousands of patients were collected and then analysed in many collaborating bodies worldwide, and scattered across many databases held by these different organisations, very few of them in Africa. Combining databases in Africa, which has known more than two dozen Ebola outbreaks in the last 40 years, would have been a far better approach, to permit routine strategies for outbreak identification, control and characterization to be developed and applied. Sadly, no such facility was in place, although this may now be corrected through the development of an Ebola Data Platform [122].

The government-led response to the West African Ebola outbreak included many different international organisation

Figure 6.8 a) International response to the 2013-2016 Ebola outbreak in west Africa

When the outbreak ended and organisations left the region, the data was scattered globally

Figure 6.8 b) Data flight from Africa following resolution of the epidemic

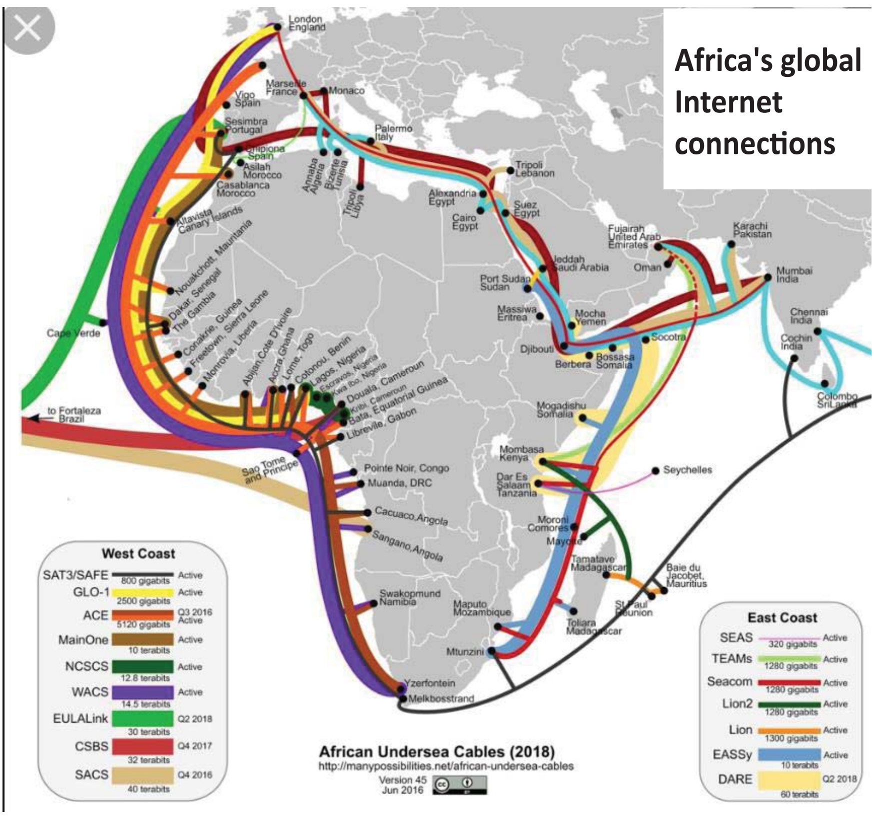


Figure 6.9 Major internet connections around Africa

6.3.3 Capacity building in open data skills

Capacity building in skills relating to the 'science of data' (data analysis, visualisation, statistics etc.) and those relating to data management (planning for curation, annotation and metadata etc.) are critical if OS in research and innovation for development in Africa is to have any realistic impact. Skills in statistical analysis, database design, data management, data analysis and integration, visualisation and interpretation [123] are rare on the African continent.

As noted in 6.2, slightly over a dozen HE institutions have developed data science courses in Africa. Generally, universities and colleges in Africa have undergraduate or certificate courses in computer science that mainly focus on basic computer hardware and software applications. To gain the benefits of open science, Africa must develop its human capital in data science to utilise data for innovation and development. Data science courses require a different focus from generalised computer science courses and deliberate curriculum design, coupled with emphasis that capitalizes on the deluge of digital data.

There are a number of initiatives in data science that form a potential basis on which a more ambitious and influential effort at appropriate scale and critical mass could be built.

a) **Data Science Africa (DSA)** [123]. (Datascienceafrica.org, 2019). The aim of DSA is to train participants in machine learning and data science methods and providing an avenue for researchers to present work that demonstrates the application of these techniques to problems relevant to the African context.
b) **CODATA-RDA Data Schools Initiative.** [125]. The aim of the school is to teach data skills to researchers in every field and to those advanced in their careers as a form of continuing professional development.
c) **University Masters and PhD courses.** Pan African University has introduced MSc and PhD in Data Science [126a]. The University of Cape Town (UCT) introduced an MSc in Data Science in 2017 [126b]. The University of the Western Cape (UWC) introduced

an MSc in Data Science in 2019 (Quantify Your Future, 2019). The University of Pretoria runs a post graduate course in Data Science [126c].

d) **Online courses.** Principa is an international data analytics company based in South Africa that provides valuable online courses [127].

e) **Data carpentry training:**

   i) In 2015, Integrated Digitized Biocollections (iDigBio) [128], conducted the first data carpentry workshop in Africa in Nairobi. The focus was to provide training and skills in generating better data. If this workshop were to be a yearly feature on the African open data-training calendar, it would add to the training infrastructure of data science short courses. However, it rotates biannually among developing countries in the world.

   ii) South Africa has institutionalized data carpentry training [129] and frequently hosts training for trainers in Unix shell, version control with Git, and a programming language (Python or R) [130]. Trainers also frequently conduct training in data organization, (clean-up, analysis, and visualization) and library carpentry lessons on concepts software development and data science to library contexts. Other training initiatives [131] aim to build bridges for digital and computational literacy.

   iii) Most frequently, the data carpentry workshops across Africa are being conducted through the NRENs [132], for the sole reason that they have the basic infrastructure to support such capacity building skills training and development in data skills.

A common trend amongst these initiatives is that most trainers/ facilitators come from outside the continent. Although their contribution is welcome, a system based on this cannot be built to the scale that is required. A more strategic approach is required at several levels in:

- developing systems of summer schools based on training the trainers;
- online courses for in service training;
- more widespread development of university courses at post-graduate level.

The Science Granting Councils could play a significant role here in stimulating the creation of course content for the different levels and mapping potential demand and provision as a basis for concerted investment and action amongst governments. Although the lack of data science (including data curation) and software engineering skills are problems worldwide, they are particularly acute in Africa. It is also important to recognise the need for a spectrum of skills as a vital adaptation to the new digital world for:

- governments and funding agencies;
- primary, secondary, further and higher education
- citizens who need to be prepared for lives as responsible citizens in a data-rich world.

6.3.4 Open Data processes and procedures

These processes and procedures are those of research data management, FAIR data, data citation, licensing and analytics. It is important to stress that the challenge that these rigorous processes entail are not easy for even the best funded and supported scientific systems, and many are struggling to cope.

Research libraries across Africa are important key stakeholders in terms of data curation and RDM training and implementation. The current uptake is slow, and libraries should become part of broader conversation involving areas of science that are inherently data-intensive.

A valuable questionnaire study of the provision of data services by research and university libraries in southern Africa has been conducted [133] to establish the readiness of libraries to engage collaboratively with their stakeholders in providing institution-wide services and systems for research data management, including data support services, archiving, organisational structures, staffing and training, funding, outreach and partnerships, and challenges and data management. It is clear that many services are still at an early stage of development, with some countries and institutions not yet ready to implement any type of service. A small number of institutions have policies in place or are planning to implement some in the near future. In some cases, existing institutional repositories are seen as a possible extension of data management services, and in some, recruitment of specialised personnel is taking place. Many are up-skilling traditionally trained library staff. The report suggests the need for advocacy and awareness-raising about research data management with libraries taking a leading role in spearheading data management and providing training and the technical support needed to store and retrieve research output and data sets.

Processes such as the use of unique identifiers and identity management are well implemented on ISBN and ISSN levels, but more awareness needs to be created about digital object identifiers (DOIs) and research IDs (ORCIDs). It is clear that the level of awareness of need as yet falls well below issues such as FAIR data, or the need for support in such advanced analytics as machine learning.

The issue of awareness is crucial. However, we argue that it is awareness of how much better a researcher's science could be if they were to subject themselves to the disciplines of data-rigour that is the most powerfully persuasive argument. We have argued above (5.4) that science-driven awareness of the potential of data is the correct route, and add here that the example of enterprises such as H3ABioNet need to be promoted as models of productive scientific enquiry in a data-rich age.

A number of organisations already implement open access policies (see ROAMAP), but at an institutional rather than national level. They would need to be aligned with broader open science and IP policies should countries adopt them.

6.4 Open Access publishing

Initiatives from the Directory of Open Access Journals (DOAJ) [134] and African Journals Online (AJOL) [135] are making progress on working with journals to make the transition to dissemination. DOAJ records the number of journals currently listed in 16 countries as: Algeria (19), Angola (1), Cameroon (1), Egypt (50), Ethiopia (3), Ghana (5), Kenya (5), Libya (2), Mali (1), Mauritius (2), Morocco (12), Nigeria (6), South Africa (85), South Sudan (1), Tunisia (6), Uganda (1).

Scholarly journals are slowly adopting policies for open access, for self-archiving in institutional repositories, and for data curation. Recent mega-journal initiatives in the continent include: AAS Open Research (funded through AESA, AAS & NEPAD) and Scientific African (published by Elsevier, owned & managed by Next Einstein Forum (NEF).

Great progress has been made in making research output available in the form of research articles (second copies), theses and dissertations that are available through institutional repositories (IRs). openDOAR lists 165 IRs from African countries (Eastern Africa 60; Middle Africa 1; Northern Africa 30; Southern Africa 44; Western Africa 30). ASSAf, in collaboration with the AAU, have developed IR criteria for a trusted IR, to guide IRs. High quality IRs are being harvested through DATAD-R (AAU).

6.5 Open to Society

We have argued (4.3) that greater openness to society has become a necessary and increasingly important dimension of open science. What activities does it comprise, how important are they for Africa, and what is currently happening in this domain in sub-Saharan Africa that could be built on in an open science initiative? Put simply, an open science initiative inspired by transdisciplinary values (the co-design and co-production of knowledge between scientists and other societal actors) should strengthen the engagement of science with all those parts of society where it can add value. Its value lies in the creation of networks of mutual learning and production of solutions-oriented, actionable knowledge, in collaborations that are profoundly enabled by digital technologies.

There are two important dimensions. Firstly, there should ideally be pervasive network access, easier in urban areas, more difficult in rural areas, but where new wireless solutions for remote access are increasingly available [136]. Secondly there need to be structures and processes that facilitate engagement with a diversity of societal actors, where machine learning approaches can be powerful means of identifying, creating and managing networking between knowledge partners in a controlled way. Relatively formal structures of engagement with business and commerce have been developing in many African universities in recent years. The ISC's International Network for Government Scientific Advice (ISC-INGSA) has developed rapidly, whose INGSA-Africa chapter [137] might be a basis for an INGSA node in an open science network.

A fundamental dimension is a wide, open science engagement with citizens and communities, including those that are marginalised, a dimension that plays an important role in contextualising science (see 7.6). It is one that has developed strongly in recent years. The ISC programme, funded by SIDA, "Leading Integrated Research for Agenda 2030 in Africa" (LIRA) seeks to increase the production of high-quality, integrated (inter- and trans-disciplinary), solutions-oriented research on global sustainability by early career scientists in Africa [138]. It has demonstrated its capacity to bring real benefit to communities on crucial issues of practical relevance to them, such as water supply, health, urban sustainability and disaster risk. There are many NGO initiatives that focus on the interface between social need and technical/scientific process. The Open Data4Development (OD4D) programme funded by IDRC seeks to create locally-driven open data ecosystems around the world [139]. The citizen science movement is growing in Africa and the shift towards transdisciplinary science at an international level has strong participation from African scientists, particularly in areas such as public health, urban sustainability and agriculture.

The role of an open science initiative could be as a coordinator of federated projects and programmes such as those illustrated above, to maximise impact, increase intra-African collaboration, enhance mutual learning, provide and manage access to a wide variety of skills and knowledge, and to facilitate provision and use of modern digital tools. Careful and inclusive planning could put Africa in a leading position in a domain that is vital to social transformation, to the achievement of sustainability and to defending the value and values of science in a "post-truth world".

6. Context for a Science Granting Councils initiative

The sub-Saharan science landscape
Weaknesses
- Low investment in R&D – 16.7 % of global population less than 1% of global R&D spend
- Majority of science funding from outwith the continent
- Few centres of critical mass
- Very low level of intra-African collaboration
- Universities struggling financially
- Few high performance or cloud computing facilities, and inefficient wide area networks
- Piecemeal patterns of open science policies and few common standards

Strengths to build on
- Square Kilometre Array (SKA) collaboration between African states developing powerful computational and cloud capabilities
- Some high quality database centres
- Ambitious Platform developments (e.g. H3ABioNet)
- Major World Bank investments in digital skills education

Opportunities
- Strong circum-continent internet connections
- NRENS as potential framework for a strong intra-continental network
- Open science and society initiatives by overseas development agencies (e.g. SIDA & IRDC programmes)
- Potential for a coordinating role for SGCs

7. Enablers and inhibitors of open science in sub-Saharan Africa

We now discuss key enablers and inhibitors that need to be exploited and overcome respectively if a successful open science enterprise is to be created in sub-Saharan Africa. Technology is a key enabler of open science, whilst some pre-existing policies, processes and habits that were more or less well adapted as enablers for “pre-digital science” have become inhibitors of innovation in a digital era, and need to change. They include national policy frameworks, some incentives and norms of scientific behaviour, technical skills and outmoded cultural assumptions.

7.1 National policy frameworks

National governments and their funding agencies should consider, both individually and collectively, adopting policies that enable and encourage open science. Without a framework of regulation or legislation to unlock data and stimulate sharing of scientific knowledge, significant progress would be difficult. For research undertaken in universities, a typical process [140] has been for national funding bodies to require, by regulation, data acquired in research that they have funded to be made open, with a prescribed deadline for submission to a trusted data repository and in a format prescribed by regulation or negotiation. In addition, many governments have adopted an open government charter [141] that requires them to open some of their data holdings, and national statistics offices now collaborate internationally in developing open data practices [142]. The principles underlying such developments should ideally be “openness as a default position” or “as open as possible, as closed as necessary", although the latter formulation begs the question, who decides and what are the criteria? The extent to which the private sector monopolizes data, much of which is publicly sourced, is a matter of increasing international concern, and under review by the International Science Council. Companies such as Google and Facebook are now facing pressure to recognise that they do not own much of the data that they routinely acquire from public or private sources. An African contribution to this discussion is essential.

Policies are also required for science management, funding, intellectual property, and copyright. It is particularly important that IP protection is well balanced between protecting the rights of originators and stifling innovation. A number of organisations already implement open access policies, at organisational level (31 OA policies from Africa registered on ROARMAP), though this also needs to be done at national and intra-national levels. Relevant policy statements that have been advocated for Africa comprise [143]:

- Adopt Findable, Accessible, Interoperable and Reusable (FAIR) Data Principle
- Observe Data Justice when distributing data, selecting procedures for distributing data and finally using data.
- Establish open access to publications through repositories and journals.
- Support submission of data to a repository before submitting the respective manuscript analysing the data.
- Develop shared and interoperable data infrastructures.
- Encourage use of recognized waivers or licenses that are appropriate for data
- Public and private funders should adopt obligatory green, gold or a hybrid of green and gold open access policies with their respective implementation measures.
- Offer incentives to acknowledge open practices in publications.
- Encourage open peer-review models.

7.2 Resistance to openness and sharing

Although many scientists support the OS agenda in principle, they are often resistant in practice. It is important to distinguish between three related issues:

7.2.1 The data supporting a published truth claim

The reproducibility crisis of recent years (see section 4.1.3) reflects in part a widespread failure to make the data and metadata underlying a published truth claim openly available. This subverts a process that is at the heart of the scientific enterprise. The motivation for such failure is frequently that authors wish to mine the same data for further publication. Nevertheless, it is malpractice and should be non-negotiable. Funders, scientific bodies and particularly science publishers should work to ensure essential compliance with what is a fundamental scientific norm.

7.2.2 Other data from publicly-funded research

The attitude implicit in the behaviour of most publicly-funded researchers is that that they “own” the data they have collected or have caused to be collected. In contrast, the international accord on open data [144], endorsed by over 120 major scientific bodies world-wide, enunciates the principle that – “Publicly funded scientists have a responsibility to contribute to the public good through the creation and communication of new knowledge, of which associated data are intrinsic parts. They should make such data openly available to others as soon as possible after their production in ways that permit them to be re-used and re-purposed” [145]. This implies that researchers do not own their data. They are data custodians on behalf of taxpayers who have funded the research, and their responsibility is to ensure that the maximum benefit is derived from this data, whether by them or others.

It is our view that this ethos is growing, but most strongly in those areas of science where collaborative, sharing enterprises have shown the power of openness in creating new scientific understanding (e.g. crystallography, bioinformatics, linguistics, Earth science, etc).

The SGCI should take note of this in promoting joint programmes of Africa-relevant open science (see 5.4.1).

7.2.3 Asymmetric benefits of N-S collaboration

There is concern that one of the consequences of adopting an OS agenda in Africa would be to enhance a process that has been experienced in recent decades whereby collaborative research between African and Northern Scientists has led to data migration from Africa and the loss of intellectual property, including from indigenous sources. It has been referred to as “helicopter science”, where collaboration with global north partners, funded by northern agencies, are frequently dominated by northern scientists, who fly in, collect data from their African partners, then fly out. Collaborations have proliferated in recent decades as international agencies have stepped up funding for research in Africa, particularly in the field of health. Yet many African scientists have often been little more than data-collectors and laboratory technicians, with no realistic path to develop as research leaders. However, overseas funders are increasingly prepared for African agencies to influence the agenda [146]. The Science Granting Councils should consider an intervention with the purpose of agreeing a concordat with overseas funders to ensure that collaborations support the career development of African scientists.

7.3 Incentives & motivations

7.3.1 The challenge of change

It is important to recognise the impacts on well-established personal and institutional habits created by the technologies of the digital revolution and the open science transition. Many of those habits, such as those surrounding scientific publication, represent adaptations to modes of communication and working that are well-suited to paper-based and pre-digital technologies that have become almost obsolete, rather than matters of unavoidable scientific necessity, and can create a barrier to open science innovation. However, changing embedded habits is not easy. It is vital to reconsider the incentives for change, and how those incentives can draw upon deep motivations that are shared by many or most scientists.

OS fundamentally threatens the comfort zone of researchers, institutions, governments and international funders who have had long-held traditions on how to conduct science and how to handle and treat data from the scientific process. Systems of accountability cut out the public, being considered as a matter a between the researchers, publishers and universities alone. The dominant mode of work until recently has been that of researchers working in isolation or in small, closed groups sharing lab notes, with results being published in pay-walled journals, inaccessible to the average citizen. However, the edifice of open science is built on sharing scientific activities, knowledge and data beyond the nexus of the researcher and pay-walled journals [147]. From the perspective of the traditional researcher/university/government, OS threatens a loss of power and control over information, data and management of the research process.

The change in mind-set and of practice expected of participants in the new open science paradigm is radical in destabilising the status quo [148]. It is understandable therefore that some in the African scientific community, like their counterparts in other continents, should be lukewarm or even trenchantly resistant to OS. In this setting it is crucial to understand not only where established patterns of incentive are barriers to change and where they need to change, but also how open scientific approaches can speak to the fundamental motivations of scientists and their institutions.

7.3.2 Incentivising change

In recent decades, for good or for ill, research has become perceived by universities, which contain the majority of public sector researchers, and their academic staffs, as the predominant determinant of reputation. Reputations of both scientists and their institutions have been predicated on the basis of metrics of research income, numbers of citations, publication in so-called “high-impact” journals, prizes and the academic league tables that purport to reflect university excellence. Three immediate questions arise:

- are these proxy metrics appropriate?
- are they barriers to desirable change?
- do they have perverse consequences?

Proxy metrics tend, almost inevitably, to become targets, which suffer from the consequences of "Goodhart's law" [148], that "any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes", which has been re-stated [150] as "when a measure becomes a target, it ceases to be a good measure", primarily because they can be and are "gamed." Exactly that has happened, suggesting the need for new measures but with the rider that they too are likely to become inappropriate targets.

These metrics have become barriers to change by concentrating, at the level of researchers, on the performance of the individual rather than the team, and at the level of the university, of the performance of the university team rather than the wider scientific group of which the university team or individual may be a member. They both militate against the intra-African collaboration which we have argued could be a powerfully positive impact of OS.

A major unintended and perverse outcomes of the power of the publication metric has been the massive growth in the number of published scientific articles, of which only a very small proportion gather significant indices of impact, together with rich market pickings for commercial scientific journals that feed on the demand. It absorbs a massive share of universities' potentials, to the detriment of their primary role as educators of the next generation and distinctive contributors of their knowledge base to innovation across the whole social, economic and political spectrum.

It imperative that incentives are developed that are appropriate to the evolution of science [151]. Systems of so-called altmetrics are being developed that permit recognition and visibility in the scientific community whilst encouraging collaboration with other researchers, and regaining authorial rights to their work and data stored online. The International Science Council is shortly to announce major projects on metrics and science publishing that will address these issues. It would be appropriate to ensure African engagement with this project to ensure that the distinctive concerns and voice of Africa, and indeed of the global south, are heard.

7.3.3 Motivating change

A fundamental lesson in the management of scientists and science systems is that scientists are enthusiasts. They are profoundly motivated by the opportunity for discovery in their chosen fields. Incentives are the stick, but self-motivation is the carrot, and much more nutritious. It is one of the clear lessons to be drawn from the examples of open science systems in chapter 5.

As suggested in 5.4, multiple benefit would be realised by funding intra-African collaboration (benefit 1) on major issues for Africa (benefit 2) that require cutting edge, inter-disciplinary/transdisciplinary work (benefit 3) that needs operational open science approaches for its success (benefit 4), and that then inspires emulation in other fields of science (benefit 5).

7.4 Skills and capacity building

Skills and educational programmes in data science and engineering and data management in the broadest sense are fundamental to the effective exploitation of the digital revolution and the adoption of open science in Africa as a powerful means of energising its scientific effort. Such is the volume and diversity of digital data streaming into storage systems from a large variety of sensors and sources, far greater than previously known, that rigorous control and management of these data have become a fundamental issue for modern science and for the public and private enterprises for which such data is crucial to their future success.

Although the lack of data science (including data curation) and software engineering skills are problems worldwide, they are particularly acute in Africa, which has not been able to train and produce enough data analysts and scientists and other support staff able to acquire and process large data sets, to identify patterns, establish relationships and solve problems [152]. The gap between Africa and much of the rest of the world is widening. The use of resources is not optimised, training institutions function in silos, and African students are only exposed to data science during tertiary level education [153]. Rationalised and coordinated training schemes and common, perennially up-dated curricula are essential.

There is a particular need in research, governmental and private sectors for:

Data stewards who handle and manages data and whose responsibilities include planning, implementing and managing research data input, storage, search, and presentation for the whole data management lifecycle.[1]

Data scientists who have expertise in the overlapping regimes of business needs, domain knowledge, analytical skills, programming and systems engineering, and managing end-to-end scientific processes through each stage of the data lifecycle, up to the delivery of scientific and business value to science or industry. [2]

Primary factors that hinder the development of these skills are:

- lack of political/managerial leadership and awareness of the need for investment;
- lack of training opportunities and acknowledgement of courses by national accreditation agencies;
- Inadequate infrastructure: slow and unstable connectivity, unreliable power supply, obsolete computer infrastructure from medium-scale server infrastructures to small numbers of workstations, lack of centralized and secure data storage.

[1] *The working definition of data steward adopted in this framework is the Edison definition for a data steward on p. 21 of the Data Science Framework document presented at the Malta workshop June 8-9 2017.*

[2] *The working definition of data scientist adopted in this framework is the Edison definition of a data scientist on p. 9 of the Data Science Framework document presented at the Malta workshop June 8-9 2017.*

Overcoming these barriers would benefit from:

- developing a federated pan-African strategy and actions;
- developing agreements with a consortium of funders for a decadal support programme;
- enhancing and coordinating supportive international collaboration;
- funders making provision for capacity building as a part of grant allocation;
- institutions making provision for capacity building as part of institutional budgets;
- including data science training as part of Continuing Professional Development (CPD).

7.5 Data analytics and machine learning

Re-invigoration of skills in statistical analysis is vital for handling large and complex data volumes where the pitfalls are serious for the unskilled, and training and degree offerings must ensure that they are embedded in relevant programmes. A further major priority derives from the impact that machine learning in particular is having on cutting edge scientific research, on governmental and business processes, and in providing efficient and cost-effective solutions for a wide variety of complex problems across the whole breadth of human concern. Such is its ubiquitous applicability, that scientists and researchers from almost all fields need to understand, at least in schematic form, how learning algorithms work, and to be able to use them.

A crucial issue for Africa is and will be, how to create, manage and apply high level skills in machine learning for a wide and diverse community, whilst also maintaining a cutting-edge presence in this rapidly developing field. It is possibly that the African Institute for Mathematical Sciences, which has a distributed presence in Africa, could fulfil this latter role. Deployment of state-of-the-art service, training and educational functions for excellent scientists in their field, whether it be biology, philology, economics or chemistry, should also include support in ways that do not require such scientists to become AI experts in order to use AI technologies with rigour.

7.6 Contextualising Open Science

A major problematic issue for Africa, and it is true elsewhere, is how to adapt to a major global movement in ways that are responsive and sensitive to the regional context and culture. We identify three aspects of this problem for Africa.

a) The strength of the relatively well-funded science systems of the global north has been such that the global science agenda has been dominated by issues defined through a northern lens, and amplified by the so-called "high impact journals" that are largely in the hands of northern commercial publishers. It is important that African science builds on and develops a sense of African priorities and adds its voice (see 3.3d) to the voice of the north in identifying and framing truly global priorities. This is not to imply that there is an African "truth" and a northern "truth", but that there are different experiences, which may perceive different priorities in the search for truth. It is an essential issue in the project to de-colonise human affairs [154].

b) Scientific publishing is a vital means of articulating the scientific voice, but the asymmetry of access to mainstream publications as indicated above, either as reader or author, diminishes the extent to which that voice is articulated or heard. It should be a major priority for the Science Granting Councils, as discussed in 4.2.4 to explore how an African science publication strategy might be developed to serve the needs of the continent and thereby the global scientific community of which it is part.

c) We regard openness to society as particularly important in the African context, as it is a means whereby the sense of African priorities alluded to in c) can be drawn out. It is critical to craft the case for in an African context and to engage at the outset with the inhibitors highlighted above from the perspectives of the different language communities and indigenous knowledge contexts. This is important because the language of the Internet and the language of communication that facilitate information, knowledge and innovation exchange in the African scientific community and citizenry is generally tied to monolingual legacies of former colonial powers.

The push for bilingual and multilingual awareness [155] in pedagogy and general communication is emerging on the African continent. Scientific activity and outputs are slowly being made available online and in data repositories. The implication of this is that, besides the bilingual combinations, African indigenous languages are finding space on the Internet and may eventually assume the role of conduits for African scientific activities, outputs and repositories. The discourse on open science in research and innovation for development in Africa must therefore be understood from this heterogenous mix of countries that share one underlying goal: to use science, research and innovation to spur development and to improve their people's lives.

7.7 Perspectives on Open Science from Science Granting Councils

A questionnaire was circulated to a group of Science Granting Councils, primarily those drawn from the SGCI, with the intention of understanding the potentials that they see in open science as a means of delivering their mission, and understanding what they believe, based on their experiences, to be key enablers and inhibitors of the process of embedding Open Science nationally or regionally. The full questionnaire is shown in appendix 10.3, together with a summary and statistics of responses. The questionnaire was introduced with by two related hypotheses:

- The fourth industrial revolution is powered by the tools of the digital revolution.
- A collaborative "open science" area would be an efficient response to this challenge.

All respondents agreed with these hypotheses, which validate our decision to place the creation of an open science area at the centre of our recommendations, and in the policy brief derived from this report.

Respondents were also asked to rank the priorities shown below, on a scale from 1 (low) to 10:

- Wide area networks
- Open science policies
- Incentives for researchers
- Capacity building
- Cloud computing
- High Performance Computing (HPC)
- Multi-national mission-led programmes (e.g. STISA2024)
- Institutional commitments (e.g. universities)
- Commitment of external funders etc.]
- Collaboration among the 15 SGCI member countries

The average ranking are shown in figure 7.1, and rankings by state are shown in the appendix, figure 10.1.

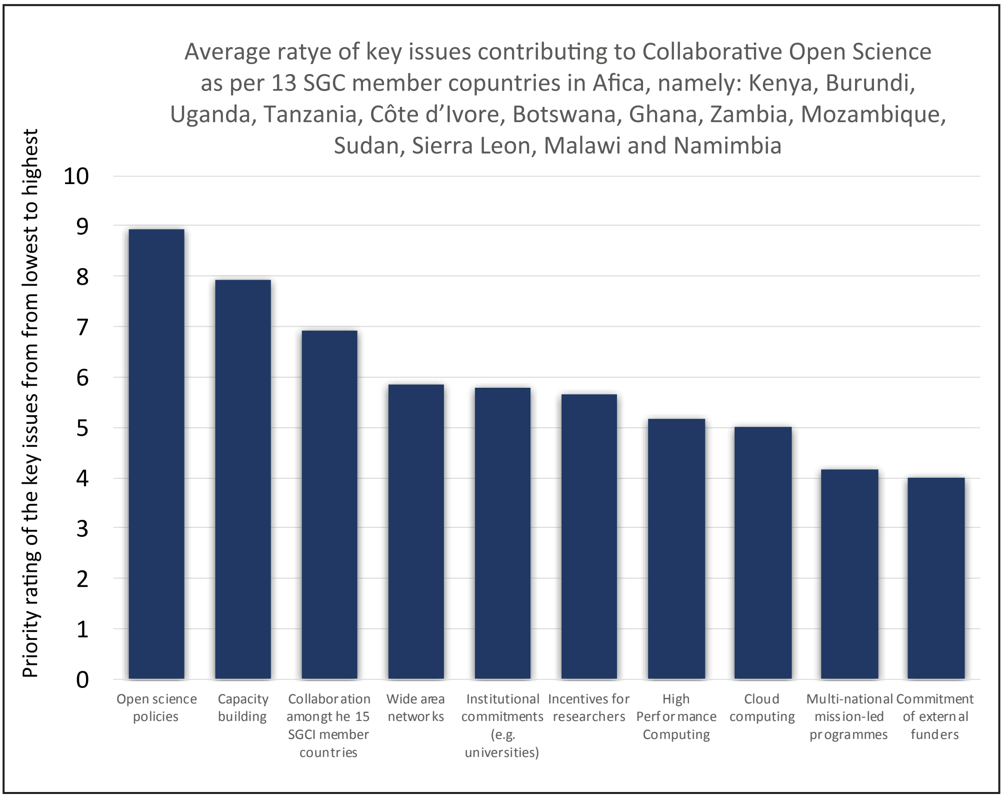


Figure 7.1. Average ranking of open science priorities.

The ranking is somewhat artificial in that is difficult to rank one priority higher than another if both are unavoidable necessities in achieving a particular aim. Thus, for a successful open science initiative as advocated here, both appropriate policies and infrastructure are essential, although they are separated in the ranking. In broad terms, vital soft infrastructures, policies – capacities – collaboration are ranked as immediate pre-requisites without which an initiative is unlikely to take off. It is heartening that external funding, though likely to be an important part of a successful initiative, is most lowly ranked, possibly reflecting confidence in an African commitment, and looking towards a time when external funding is a luxury rather than a necessity. In the recommendations in chapter 8, we do stress the vital importance of ITC systems and the programmes that will stimulate their use as well as addressing structural problems in African science systems. But we also argue that the most effective and efficient route to progress is through collaboration, and that shared open science policies are critical enablers, ranked here as the highest priority.

The questionnaire reproduced in appendix 10.3 also poses a series of questions about issues that arise for the process of creating of an open science area that would seek to realise the benefits that Granting Councils agree are embedded in the two initial hypotheses. Table 10.1 in the appendix summarises responses to issues that are central to any attempt to realise an open science area. The responses were highly informative, illustrating with clarity the perceptions amongst the Councils about key contextual issues. Responses are grouped together in broad categories:

Principle barriers to achieving collaborative open science in Africa

- Lack of understanding and commitment at political and policy levels.
- Lack of appropriate infrastructure, of human capacities, of policy and regulation at national levels.
- Career-related concerns amongst researchers, who are currently motivated by incentives that are inimical to open science.

Roles for an SGC initiative in developing collaborative open science

- As lobbyist and advocate for open science to governments.
- As coordinator in implanting initiatives, and a supporter, facilitator and funder of such initiatives.
- As creator of awareness and culture change amongst researchers and of incentives to drive collaborative open science.

Merits of a collaborative open science area:

- Cost-effective.
- Innovative.
- Maximises utilisation of research output.
- Speedy publication, wide dissemination and easy access to knowledge.
- A means of strengthening the common voice.

Demerits of a 15 SGCI collaborative open science area:

- Varying levels of capacity.
- Lack of mutual trust.
- High cost.
- Poor broadband access.
- Lack of confidentiality.
- Low impact factors of open access journals.

These responses are invaluable in identifying the open science aspirations of the Councils, the benefits that they seek to obtain from any initiative, and the barriers and perceived disadvantages that any initiative would need to overcome. They are important in framing the analysis and recommendations in chapter 8.

Box 7. Recommendations for Science Granting Councils on: Priorities for addressing enablers and inhibitors of open science

Policies & practices:

- Commit themselves to collective action in developing open science
- Develop common policies for:
  - Intellectual property
  - Fair data
  - Open access publication
  - Shared and interoperable infrastructure
- Work with institutions to create Africa-appropriate metrics for researcher evaluation
- Evaluate the needs for open science and data analytic skills and discuss with stakeholders how they might best be satisfied

Culture change:

- Endorse the Science International accord on open data and work with stakeholders to stimulate a culture of data sharing
- Develop a concordat with international funders for balanced international collaboration involving African scientists
- Stimulate a conversation on how open science might best be contextualised in the African setting

8. Possible roles and responsibilities of and recommendations for the Science Granting Councils

The questionnaire responses from 13 SGCs revealed unanimity for strong action to exploit the digital revolution, enthusiasm for open science as a means of doing so, practical realisation that change would be difficult, but a view that joint action by SGCs could be an important contribution to achieving necessary change. We now build on these conclusions to suggest a way forward.

8.1 Principles

The African Science Granting Councils, and similar bodies elsewhere, have a pivotal role because of their unique, intermediary position in national science systems. On one hand they both represent and influence government policies for science: on the other hand they both influence and respond to the priorities of the scientific community. They have a crucial role to play when confronting the epochal challenges such as that of the digital revolution, primarily because as intermediaries, they are able to deploy a binary strategy to stimulate:

- high-level, governmental and intergovernmental impact, without which the resources required for effective action will not be mobilised; and
- demonstrable grass-roots level utility, without which any mobilized resources will be under-utilized.

The preceding chapters have presented the evidence and developed the arguments which we draw on here as the basis for responses to the questions posed in our remit as set out in chapter 1, to advise on the roles and responsibilities that the Science Granting Councils might take in the development of open science in Africa.

It is important to be clear about the purpose of the interventions that we advocate, which are based on the arguments in chapters 4-6. We argue that the aspiration for the near future of African science should be to mobilise the resources needed to operationalise the policies, infrastructures and practices that are needed for a powerful open science capacity. These foundational essentials are:

- **Shared, pan-African Policies** for the purpose and practice of open science.
- **Access to state-of-the-art computation and communication systems,** with major distributed nodes of capability that will stimulate and serve a growing network.
- **Grand challenge research programmes** that focus on major issues for Africa that stimulate take-up of data intensive capacities, the creation of virtual critical masses and enhanced intra-African collaboration in an African open science area.
- **Major database centres** that serve the above objectives and are powerful resources for open science.
- **An internationally competitive artificial intelligence/machine learning capacity** to inspire and serve the open science community.

We regard the questions posed by the Science Granting Councils as potential preliminaries for a strategy to achieve the above objectives. We now respond directly to each of these questions in turn and then make relevant recommendations that arise from our analyses which are then mapped on to the SGCs framework for action.

8.2 Responses to Science Granting Councils questions

Question 1: *What roles could Science Granting Councils play in fostering Open Science in research and innovation for Africa's development and how can they effectively play this role within the OS ecosystem?*

Worldwide, the analogues of the Council have been grappling with the same question. Their varied responses reflect institutional cultural habits, the level of national funding for research, the level of relevant infrastructure, and the perceived benefits of particular modes of adaptation. These latter include policies and regulations for open data and open access publication, policies for human capacity development, major funded programmes to simulate data-intensive science, and open science platforms at disciplinary or national level.

The dilemma for the African Granting Councils is the relatively low level of resource that is available to most of their number in order to stimulate and fund such changes, and the relatively low level of IT infrastructure provision that is required to support them, as summarised in chapter 6. In this setting, it is incumbent on Councils to use their unique intermediary positions to stimulate change and to avoid the dangers of a looming knowledge divide. Their influencing, convening and coordinating roles should include:

- influencing national government policymakers;
- influencing national science systems including institutions and their researchers;
- using their collective aggregate resources strategically (assuming the acquiescence of governments);
- coordinating policy and action;
- making an integrated case for African science priorities and drawing on support from international development bodies such as IDRC, SIDA, DFID etc;
- accessing the expertise of the international science community as represented by the International Science Council (ISC) and its data bodies (CODATA, WDS), and the Research Data Alliance,

and participating in the work and influencing the policies of the Global Research Council.

We have been impressed by an issue flagged in the African Union's STISA2024 report, that enhanced intra-African collaboration may hold crucial potential for the progress of African science and its application. A way to achieve such collaboration and the critical masses that are increasingly needed in modern science, would be for SGCs to act as a collective in promoting and coordinating a major open science initiative that could have the economies of scale and impacts referred to in 3.3.

Question 2: *What tools, interventions, policies, incentives, infrastructure and frameworks are required to foster OS in research and innovation for development? Which of these are of immediate relevance and importance for Africa's Science Granting Councils?*

The issues relevant to these questions are discussed in chapter 4, which sets out the basic toolkit that is currently regarded as intrinsic to open science together with some of the problematic issues that require attention; in chapter 5, which sets out the systemic concept of platform or commons that have proven able to create a framework within which open science services can be most efficiently delivered; in chapter 6, which describes the mosaic of open science initiatives that have or are being developed in sub-Saharan Africa, and draws attention to the overall lack of coherence; and in Chapter 7, which addresses the policies, incentives and some of the processes that are required to ensure efficient delivery.

Interventions with government

The government-facing role of SGCs should be to make the case for:

- the vital role of the science system in creating and exploiting opportunities in the 4th industrial revolution;
- adapting science systems to the new paradigm of open science to maximise their creative potential;
- endorsing collective action by SGCs as a cost effective means of maximising impact and developing powerful synergy with AfCFTA .

Policies

a) ***Open Data.*** There is general consensus about the technical issues that require specific policies if an open data regime is to be implemented (section 4.1). They include policies for research data management, data citation, whether open data should be mandated as a condition of public funding, when it should be released and in what form it should be released (e.g. FAIR data), how claims of intellectual property and copyright should be adjudicated, the limits of openness, and policies for privacy, safety and security. A common framework of standards should ideally be developed to regulate the ethical use of open data, which would need to be explored in the context of national legislation. Given our suggestion that intra-African collaboration should be regarded as an important priority for any initiative, intra-African variations in the extent and character of policy and regulation (see ch. 6) are an issue to be addressed.

b) ***Scientific publication.*** The issue of scientific publication is more problematic (section 4.2). Affordable access for both readers and authors is a major issue, and, given its global ramifications, should be pursued both through discussion between African stakeholders and as part of the International Science Council's new project on scientific publishing [156]. Discussions with Latin American colleagues would be useful in this regard [157].

Incentives

Having incentives for individuals and institutions that are aligned with their purpose is essential to successful achievement of that purpose. Section 7.3.2 argues that current incentives for research have not only led to increasing systemic dysfunction but are also inhibitors of open science practices through improper use of metrics. These are issues that are increasingly discussed internationally, and with much work on alternative metrics, though these too can be gamed, and should be developed with that in mind [150]. Ultimately, discussion is rooted in the fundamental purpose of research and universities in society, against a backdrop of reputation and branding. If the SGCs were to decide on a collective open science initiative [158], the issue of incentives would be a vital topic for resolution, both within Africa with key stakeholders, including universities, and as part of the international effort about to be led by the International Science Council [159].

Infrastructure

Building on current capacities to create an effective pan-African, networked computational and communication system that is managed to provide efficient services to the scientific community, including cloud and high performance computing, is a vital pre-requisite to enable modern open science. The major internet connections around the shores of the African continent (Fig. 8.2) are as good as any other continent or region. It is the inter-state and internal connections that require investment, as described in chapter 6. This should be a fundamental priority for the collective SGCs in an open science initiative.

Frameworks

We have argued (ch. 5) that because open science tools and processes interact dynamically, the most efficient way of deploying them is as component parts of a common platform that provides seamless services to its members. It is important to note that there are levels of activity within such a framework, from overall strategy, to technical coordination of all services, to delivery of individual services.

"Innovation for Development"

Notwithstanding the political and economic importance of the concept of the 4th industrial revolution, it is important to recognise much wider dimensions of innovation than the conventional industrial. Research-supported innovation needs to occur at all levels of society, from individual citizens to citizen groups to all levels of governance, and that an open science initiative should be inclusive of all levels.

Question 3: *What are the key enablers and inhibiters for mainstreaming and implementing OS policies, initiatives and activities in Africa; and how can they be sustained and resolved respectively?*

This overlaps with our responses to question 2, so there will be some repetition although within a different formulation. We comment that the open science movement of the last decade has been conceptually driven both by active researchers and by science policymakers and influencers (particularly science academies and government agencies), and has gathered increasing momentum. It is a response to the unprecedented opportunities created by the digital revolution, the progressive replacement of the "lone scientist" by teams for whom sharing is second nature, and increasing awareness of the global challenges facing humanity that require a collaborative, international effort.

Surveys [160] demonstrate an increasing will to pursue the route of open science, with every prospect that it will become embedded as a fundamental norm of 21st century science; in the mainstream and therefore sustainable. However, that potential is at least slowed down, and possibly undermined by major inhibitors, which are, primarily:

a) metrics for individual performance and criteria for advancement that are ill-adapted to teamwork and sharing (7.3.2);
b) the burden of complex tasks required for efficient management, sharing and re-use of data (4.1);
c) affordable access to scientific publications by both readers and authors (4.2);
d) access to adequate computation and communication networks [161].

We have suggested pathways to the resolution of these issues in 8.2:

- for a), work on alternative metrics and criteria (recommendation 2c);
- for b), creation of a framework to deliver seamless services (recommendation 1a);

- for c), work nationally and internationally to devise workable access (recommendation 2b)
- for d), discussions with NRENS, HPC and Cloud groups about a federated solution (recommendation 2d).

Question 4: *Who are the key players? How is OS governed? How are the rules, roles and responsibilities determined in the co-creation and utilization of open knowledge? What are the experiences across the 15 SGCI countries?*

Key players

If the collective convening power of the Councils could be mobilised in support of a bold and ambitious strategy, its success would depend upon attracting the commitment and understanding the motives of a wide range of stakeholders which, because this is such an important issue, are described in detail in appendix 2 and summarised here as comprising three groups:

a) **Policymakers and influencers,** primarily governments and their agencies, dominantly motivated by concern for innovation and development; national academies and university representative bodies (e.g. AAU) with motivations for excellence in science systems.

b) **Practitioners,** primarily comprising researchers motivated by scientific opportunity; universities motivated by reputation, funding and attractiveness to staff and students; and private sector companies motivated by innovative capacity, the supply of innovative personnel and the creation of markets for their products.

c) **International supporters,** primarily comprising international funders of research, particularly in the fields of development and health; and international scientific bodies.

A crucial issue for Africa is that a large proportion of funding for its science originates from outwith the continent, from charitable bodies and foundations and from national development agencies.

Whilst there are some funders that have sustained predictable patterns of funding for particular regions and purposes over long periods, the overall pattern of funding does vary considerably in its geographic and thematic focus. It would be of great benefit if the Science Granting Councils could act as an interface with international donors and science bodies of community c) above, able collectively to express continental priorities with greater coherence, and if funders were to act in more coordinated ways in responding to these priorities.

Governance

There are several levels and dimensions of governance, which will vary according to the scope and ambition of any initiative. A governance approach for an African open science area would be the most ambitious, and we follow the recommendation of chapter 5 that such an option should have a coordinating hub, with national, potentially specialised nodes, able to ensure a distributed capability.

Effective governance of OS requires recognition of the responsibilities and contributions of the key players described in 8.4.1. We have stressed the efficiencies and potentials that would be released for Africa by developing a collective, inter-state approach, which we presume would require a dialogue between the Councils and Governments, particularly if recommendation 1a were adopted by the Council for further work.

If recommendation 1a were to be progressed by the Councils, and based on experience elsewhere (ch. 5), the following might be an appropriate structure, reflecting the different but complementary roles of Governments, the SGCs, technical experts and users:

a) An African Open Science Commission at the interface with governments and Granting Councils to agree on policies and priorities for open science. It might include representation from the AfCFTA secretariat.
b) An African Open Science Oversight Board, as a Granting

Councils' body, but with membership including key stakeholders (8.3), including high-level representation from business, and with responsibility to oversee and propose evolution of an open science strategy.

c) Coordination and overall management of the technical operation of the system would be in the hands of the coordinating hub, overseen by a Technical Advisory Board. This would require high levels of expertise, experience and politically-aware judgement.
d) Each national node would have its Management Committee, responsible to the hub for the delivery of agreed services, and representing national priorities to the Technical Advisory Board.

If the Councils were to take this route, it is possible that an Oversight Board could, with the agreement of its current Advisory Council, take over or merge with the African Open Science Platform to give the Platform the breadth of institutional leadership that it needs, as the operational arm of an African open science strategy, bearing in mind that AOSP is pan-African in spirit, not just sub-Saharan.

Open knowledge creation

It is important to recognise the need for high-level governance and coordination of a shared system. This should not be confused with project–level governance which must be adapted to specific circumstances of the project purpose, including engagement with communities in ways that require inclusive issue- and community-specific governance arrangements (e.g. [162]).

Question 5: *What are the pros and cons of OS? Is OS increasing marginalization or bridging the divides? How can OS benefit excluded/ vulnerable groups?*

Many of the most exciting initiatives in open science have been grass-roots-driven efforts [163], [164]. The data sharing processes that are at the heart of the modern concept of open science were developed in some disciplines, such as crystallography, linguistics, genomics, when long, standard, data series became available, and the scientific benefits of sharing became apparent to their scientists. A position paper by the International Union of Crystallography [165] spells out these benefits with clarity and in detail, and contains a ringing challenge to the scientific community – *"we urge the worldwide community of scientists, whether publicly or privately funded, always to have the starting goal to divulge fully all data collected or generated in experiments"*. A further intervention from the same source made the powerful statement that "the science is in the data" [166]. Such grass-roots positions, increasingly linked to open access publishing and forms of societal openness have been taken up by representative science bodies, such as the Royal Society of London in their 2012 report, Science *as an Open Enterprise* [37], by intergovernmental bodies such as the G8 [167] and currently by many funding agencies in national science systems.

The fundamental argument for open science is that it is a means of deploying a collective intelligence in understanding nature and society, and of using that understanding to address fundamental issues for human society. Knowledge, science, is a public good, and publicly funded scientists have a responsibility to contribute by maximising the efficiency of discovery through collaborative working, communicating that knowledge in a comprehensible form and engaging with society in seeking its beneficial use.

The countervailing arguments tend to be conservative or radical [168]. The conservative critique defends the right of the individual against the collective. This argument was trenchantly stated in an editorial published in the *New England Journal of Medicine* [169] which described the "emergence of a new class of research parasites", which also commented that some of these parasites might seek to examine whether the original study was correct, a response that implicitly but directly conflicts with a fundamental principle of scientific rigour (ch. 2, p.6). However, this position is rational at the level of the individual when the current mode of assessment of scientists is based on publication in “high impact journal” (roundly condemned by the San Francisco Declaration) and citation statistics, which we argue are ill-adapted to the needs of modern science and which in effect enshrine the “scientific paper” as the sole goal of science.

The radical critique [170a] argues that the release of vast troves of data, papers or research results which, although potentially beneficial to science as an enterprise, simply exacerbates the trend towards the increasing marketization and corporatization of science that disproportionately benefit large corporations. Tyfield [170b] argues that open science opens the door to:

- capture of publically-funded research value by commercial platforms;
- introducing yet more “metrics” of productivity to “incentivize” scholars to work harder, and simply replace one form of game playing by another;
- focussing on system-wide progress of science, ignoring costs and benefits to individuals, whether scientists or non-scientists.

In the African case, we add to these dangers those of the exploitation of African scientific resources by better-funded researchers from the global north, and the marginalisation of African needs in evolving science-publishing regimes.

These critiques rightly challenge the developing open science movement to resist the increasing privatisation of knowledge, to maintain or redevelop a “human centred” science and to adapt to the needs of different communities, whether small or large. We suggest that commons- or platform-based systems (chapter 5) are effective ways of doing these things, provided that the scientific collective voice is a strong one, thereby laying great emphasis on the role of governance, in which users have at least partial ownership or control, rather than simply being passive drones in an “on-demand” economy.

Overt engagement with society is a crucial element in ensuring that the voice of excluded or vulnerable groups are part of the enterprise. We have no instant solution here, but suggest that an African initiative could demonstrate how this component of open science could become a powerful reality.

Recommendation 1:

Depending on the scale, enthusiasm and commitment to collective action that are possible to mobilise by exploiting their intermediary role, we suggest that the Councils should consider options for action as follow:

Option 1a). *Promote creation of an African open science area designed to offer the range of services and capacities typical of the open science platform or commons systems described chapter 5 and with the intention of yielding the benefits described in 3.3. We note (3.4) the timely recent creation of the African Continental Free Trade Area (AfCFTA), much of the potential strength of which lies in the mobility and collaboration that it will enable. It is precisely these attributes that have been shown to be effective in stimulating dynamism in science systems. Such an action would be timely and creative, resonating with the establishment of AfCFTA, and with the potential for profound mutually beneficial synergy. It would resonate not only within Africa, but globally.*

Option 1b). *Develop a strategy to coordinate and complement existing open science activities (chapter 6) in sub-Saharan Africa, primarily concerned to coordinate policies and incentives as described in 8.3.2 and 8.3.3 below.*

Option 1c). *In this scenario, the responsibility for open science policies and incentives would remain largely with national systems, but would be open to a collective, focussed strategy comprising programmes designed to have major impact on two key areas:*

- *Enhancing computational and communications (IT) capacities by federation and expansion of existing capacities;*
- *Creating major data-intensive programmes on intra-African priorities with the intention of stimulating creative use of enhanced IT capacities and building intra-African virtual critical masses.*

Recommendation 2:
The Science Granting Councils should commission and become involved in expert reviews of key open science issues as follows:

2a) A task group should be created to explore the potential for convergence of national policies, regulations and standards for open science, and the possible role of the SGCs in facilitating convergence.

2b) The SGCs should stimulate a discussion of how African priorities for scientific publishing can be achieved, and to ensure that these concerns contribute to a global review being led by the International Science Council.

2c) The SGCs should stimulate a discussion of the impact of metrics for research on the research process, and how they might be improved to satisfy African priorities. They should ensure that these concerns contribute to a global review being led by the International Science Council.

2d) The SGCs should commission an expert group with the task of identifying cost-efficient means of federating current computational, cloud and communication capabilities, and extending and efficiently managing them. (Potential funders should be involved in this process, including the World Bank, which could, for example, be the source of long-term loan finance).

***Recommendation 3:** The SGCs should seek to develop structured relationships between key players:*

3a) Initiate conversations with those actively involved in open science processes and strategies, possibly through the African Open Science Platform, to seek maximum synergy and collective impact and to reduce unnecessary duplication.

3b) The SGCs and other relevant partners should seek to create a forum together with major funding agencies (possibly build around the SGCI) to identify a more strategic approach in supporting science in Africa. This would be particularly important if the SGCs adopted recommendation 1a.

3c) If recommendation 1a were accepted, an approach should be made by SGCs to governments, for creation of an inter-governmental statement or concordat that:

- *recognizes the vital importance of science in enabling them to exploit the potentials of the 4th industrial revolution;*
- *commits them to working together in promoting Open Science as a vehicle for achieving this;*
- *embeds in policy, and in national law if necessary, high-level agreements (e.g. IP, open data, standards, rules of access for common infrastructure) that are necessary as a frame for funding and for operational open science activity.*

8.4 SGCs proposed Call to Action Framework

Actionable items from the recommendations are shown below within the SGCs proposed actions framework.

Table 8.1 Action Framework

Action Items	Action
Interventions	• Create and exploit opportunities in the 4th industrial revolution. • Adapt science systems to the new paradigm of open science. • Endorse collective action by SGCs as effective & cost efficient. • Collaborate and negotiate with publishers to implement open access as the default standard. • Create a comprehensive and transparent system for gathering and sharing information on the costs and conditions of academic communication. • Create a funding mechanism to explore paybacks to open science.
Policies	• Develop a common framework of standards to regulate the ethical use of open data. • Enforce publication of data and code concurrently with publication of concepts based on them. • Clarify IP protection. • Make open data the default standard for all publicly funded research. • Establish standards on privacy by design. • Strengthen intra-African collaboration in OS initiatives. • Review and reform reward systems. • Develop assessment and evaluation criteria that promote OS. • Adopt a positive, integrated approach in career progression systems to remove obstacles to open science practices. • Raise awareness and promote open science in universities and other knowledge institutions. • Develop plans for capacity building in data stewardship and data science. • Encourage the sharing of expertise that enables disciplines/ regions to learn from each other.

Incentives	• Champion and lead realignment of funders and research organisations to cater for both article processing charges (APC) and subscriptions charges. • Support discipline-based foundations that help flip subscription journals to FAIR open access by providing funds for APCs. • Advocate open access practices. • Provide start-up funds for alternative open access publishing models. • Encourage FAIR data sharing by valuing data stewardship and efforts to make data available.
Infrastructures	• Support national institutions to emplace institutional data policies that outline roles and responsibilities for research data management and data stewardship. • Make development of Data Management Plans (DMPs) a precondition for funding. • Introduce incentives for FAIR data sharing by valuing data stewardship and efforts to make data available and by acknowledging and rewarding those who compile the data. Require data to be cited according to international standards. Encourage the sharing of expertise that enables disciplines/ regions to learn from each other • Set up and manage local and national e-infrastructures and facilitate researchers in the selection and use of services.

9. References

[1] African Technology Policy Studies Network & The Scinnovent Centre. Concept Note and Terms of Reference for Authors. "Open Science in Research and Innovation for Development. 2019. https://atpsnet.org/wp-content/uploads/2019/06/FINAL-Concept-Note-and-TOR-for-the-Open-Science-Masterclass-Paper.pdf

[2] L. Chan., A.Okuna., R. Hillyer., D. Albornoz., and A. Posada. (eds) Contextualising Openness. University of Ottawa Press and International Development Research Canada,(354pp). 2019. https://idl-bnc-idrc.dspacedirect.org/handle/10625/58201

[3] A. Einstein Quotes. BrainyQuote.com, Brainy Media Inc, 2019.https://www.brainyquote.com/quotes/albert_einstein_100017

[4] M. Hilbert and P. Lopez. The World's Technological Capacity to Store, Communicate, and Compute Information. Science. (Vol.332, Issue 6025). 2011. DOI: 10.1126/science.1200970 https://science.sciencemag.org/content/332/6025/60/tab-pdf.

[5] M. Baker, '1,500 scientists lift the lid on reproducibility', Nature, 533 (7604), 2016. 452–454. doi: 10.1038/533452a.

[6] M. D'Ancona. Post-Truth: The New War on Truth and How to Fight Back. Ebury Press. 2017.

[7] H. Ledford. Us trust in scientists is now on par with the military. Nature, News, 06, August 2019. https://www.nature.com/articles/d41586-019-02389-8?utm_content=buffera08e5&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer

[8] B.R. Jasny. Lies spread faster than the truth. Science, Vol. 359,Issue (6380), 1114-1115). 2018. https://science.sciencemag.org/content/359/6380/1114.8/tab-pdf

[9] T. Hey, S. Tansley, and K. Tolle. The Fourth Paradigm. jim gray on escience: a transformed scientific method. Microsoft Research 2009. https://www.immagic.com/eLibrary/ARCHIVES/EBOOKS/M091000H.pdf

[10] M. Gibbons, M. Trow, P. Scott, S. Schwartzman, H. Nowotny and C. Limoges. The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies. 1995. https://www.researchgate.net/publication/225088790_The_New_Production_of_Knowledge_The_Dynamics_of_Science_and_Research_in_Contemporary_Societies

[11] H. Nowotny, P. Scott and M. Gibbons. "Mode 2 Revisited: The New Production of Knowledge", Minerva. 2003. https://link.springer.com/article/10.1023/A:1025505528250

[12] STISA 24 STISA 2024 – A guide for Africa to exploit its opportunities http://www.nepad.org/news/stisa-2024-guide-africa-exploit-its-opportunities

[13] International Science Council Plan, Advancing Science as A Global Good. 2019-2021 https://council.science/cms/2019/09/Advancing-Science-as-a-Global-Public-Good_ISC-Action-Plan-1.pdf

[14] UNESCO. (2015). UNESCO Science Report: Towards 2030. 2015. https://en.unesco.org/unesco_science_report

[15] M. Hilbert. Technological Information inequality as an incessantly moving target: The redistribution of information and communication capacities between 1986 and 2010. Journal of Association for Information Science and Technology. (65.4. 821 – 835) 2014. doi:10.1002/asi.23020. https://www.martinhilbert.net/technological-information-inequality-as-an-incessantly-moving-target/

[16a] J.P. Tennant, F. Walder, P. Masuzzo, L.B.Collister and C.H.J. Hartgerink. The academic, economic and societal impacts of Open Access: an evidence based review. 2016. doi: 10.12688/f1000research.8460.3 https://f1000research.com/articles/5-632/v3

[16b] J.P. Tennant. The academic, economic and societal impacts of Open Access. Blog. Open Science.com 2016. https://www.scienceopen.com/document?vid=2dda4be2-86ef-4183-8be4-d2ebd4f59d00

[17] J. S. Ahinon and J. Havemann. Open Science in Africa. 2018https://elephantinthelab.org/open-science-in-africa/

[18] World Bank. Open Data for economic growth. 2014. http://documents.worldbank.org/curated/en/131621468154792082/Open-data-for-economic-growth

[19] European Commission. Creating Value through Open Data. Capgemini. 2015. https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

[20] Capgemini. New European Data Portal is a key step in seizing a multi-billion euro potential of Open Data. 2019. https://www.

capgemini.com/news/new-european-data-portal-is-a-key-step-in-seizing-multi-billion-euro-potential-of-open-data/

[21] Union Commission. Cost of not having FAIR research data: Cost benefit analysis for FAIR Research data. PWC. 2018. https://op.europa.eu/en/publication-detail/-/publication/d375368c-1a0a-11e9-8d04-01aa75ed71a1

[22] Nowotny Helga, Peter Scott and Michael Gibbons (2003) “Mode 2 Revisited: The New Production of Knowledge”, Minerva.

[23] African Union. NEPAD https://edu-au.org/strategies/stisa-2024

[24] J. S. Ahinon and J. Havemann. Open Science in Africa. 2018. https://elephantinthelab.org/open-science-in-africa/

[25] Roche, D. G., Kruuk, L. E. B., Lanfear, R. & Binning, S. A. Public Data Archiving in Ecology and Evolution: How Well Are We Doing? PLOS Biol. 13, e1002295. 2015.

[26] S. Higgins. The influential Digital Curation Centre model of the data life cycle. 2008. http://www.dcc.ac.uk/resources/curation-lifecycle-model.

[27] The Royal Society. Science as an Open Enterprise. 2012. https://royalsociety.org/topics.../science...enterprise/report/

[28] M.D. Wilkinson., et al. The FAIR Guiding Principles for scientific data management and stewardship. 2016. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18

[29] The Royal Society. Exploiting data revolution is key to scientific and economic progress. 2012. https://royalsociety.org/news/2012/science-open-enterprise/

[30] Force11. Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing Version B1.0 https://www.force11.org/node/6062

[31] Scientific Data. The Fair Guiding Principles for scientific data management and stewardship. Sci data. V.3. 2016. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4792175/

[32] Figshare. Store, share, discover research. https://figshare.com

[33] Dryad. For your research data. https://datadryad.org

[34] M.D. Wilkinson., et al. The FAIR Guiding Principles for scientific data management and stewardship. 2016. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18

[35] A. Pienta. “LEADS Database Identifies At-Risk Legacy Studies.” ICPSR Bulletin 27(1). 2006.

[36] W.G. Dewald, J.G. Thursby, and R.G. Anderson. “Replication in empirical economics: The journal of money, credit and banking project.” American Economic Review. (76/4:587-603). 1986.

[37] Altman, M., J. Gill, and M.P. McDonald. (2003). Numerical issues in statistical computing for the social scientist. John Wiley & Sons.

[38] D.S. Hamermesh. “Viewpoint: Replication in Economics,” Canadian Journal of Economics. 2007.

[39] T.H. Vines, A.Y.K. Albert, R.L. Andrew, F. D barre, D.G. Bock, M.T. Franklin, K.J. Gilbert, J-S Moore, S. Renaut and D.J. Rennison “The Availability of Research Data Declines Rapidly with Article Age” Current Biology (24/1: 94 – 97). 2014.

[40] Force11. Joint Dclaration of Data Citation Principles – Final. https://www.force11.org/datacitationprinciples

[41] DataCite https://datacite.org

[42] DataCite Schema Schema. DataCite Metadata Schema 4.3. (16 August) 2019. https://doi.org/10.14454/7xq3-zf69.

[43] D. McDonald and U.Kelly. Value and Benefits of Text Mining. Jisc Reports. 2012
https://www.jisc.ac.uk/reports/value-and-benefits-of-text-mining

[44] The Budapest Open Access Initiative
https://www.budapestopenaccessinitiative.org/read

[45] Giga Science (BioMed Central BioMed Central, http://gigascience.biomedcentral.com

[46] Scientific Data. Nature Publishing Group, http://www.nature.com/sdata/

[47] Data in Brief, Elsevier, http://www.journals.elsevier.com/data-in-brief/

[48] OECD. Making Open Science a Reality. OECD Science, Technology and Industry. Policy Papers No.25. 2015. http://dx.doi.org/10.1787/5jrs2f963zs1-en https://www.oecd-ilibrary.org/docserver/5jrs2f963zs1-en.pdf?expires=1571939511&id=id&accname=guest&checksum=9C72B09E1BD539949C8A80AECEC28F00 .

[49] S. Jamil. Study on Open Science: Impact, Implications and Policy Options. Luxembourg: European Commision. 2015. https://doi.org/10.2777/237283

[50] B. Fecher, S. Friesike and M. Hebing. What Drives Academic Data Sharing? Plos One. 2015. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0118053

[51] E. Carvalho, A. Batilana, J. Simkins and R. Pietrobon. Ethical and legal constraints on data sharing between countries in multinational epidemiological studies in Europe report from a joint workshop of the European League Against Rheumatism standing committee on epidemiology with the "AutoCure'' project. Research Gate. 2008
PubMed. DOI: 10.1136/ard.2007.080598

[52] M.W. Carroll. Sharing Research Data and Intellectual Property Law: A Primer. 2015 doi: 10.1371/journal.pbio.1002235

[53] African Union, Economic Commission for Africa, African Development Bank and United Nations Development Programme. 2017. Africa Sustainable Development Report - Tracking Progress on Agenda 2063 and Sustainable Development Goals

[54] United Development Programme. Goal 9: Industry, Innovation and Infrastructure. https://www.undp.org/content/undp/en/home/sustainable-development-goals/goal-9-industry-innovation-and-infrastructure.html

[55] ITU News. New ITU statistics shore more than half the world is now using the Internet. August. 2019. https://news.itu.int/itu-statistics-leaving-no-one-offline/

[56] J.Tennant. 'Transformative' open publishing deals are only entrenching commercial power. 2019. https://world.edu/transformative-open-access-publishing-deals-are-only-entrenching-commercial-power/

[57] J.Tennant. 'Transformative' open publishing deals are only entrenching commercial power. 2019. https://world.edu/transformative-open-access-publishing-deals-are-only-entrenching-commercial-power/

[58] DORA. San Francisco declaration of May 13, 2013. https://crln.acrl.org/index.php/crlnews/article/view/9104/9996

[59] Alberts, Bruce (May 17, 2013). "Impact Factor Distortions". Science. 340 (6134): 787. doi:10.1126/science.1240319. PMID 23687012).

[60] A. Posada, G. Chen. Inequality in Knowledge Production: The Integration of Academic Infrastructure by Big Publishers. ELPUB 2018, Jun 2018, Toronto, Canada. ff10.4000/proceedings.elpub.2018.30ff. ffhal-01816707f.

[61a] Working Group 3. Researcher Recognition and Recording Nationl Platform Open Science. https://www.openscience.nl/files/openscience/2019-02/notitie-erkennen-en-waarderen-van-onderzoekers_en-gb.pdf

[61b] Working Group 3. Researcher Recognition and Recording Nationl Platform Open Science. https://www.openscience.nl/files/openscience/2019-02/notitie-erkennen-en-waarderen-van-onderzoekers_en-gb.pdf

[62] Making the Choice: Open Access vs Traditional Journals. AJE Scholar.https://www.aje.com/arc/making-the-choice-open-access-vs-traditional-journals/

[63] Lexi Rubow, Rachel Shn Brinna Schofield and Samuelson Law (2015). Understanding Open Access. https://authorsalliance.org/wpcontent/uploads/Documents/Guides/Authors%20Alliance%20-%20Understanding%20Open%20Access.pdf

[64] J. P. Tennant., et.al. Ten Myths around scholarly publishing. ResearchGate. 2019.
https://www.researchgate.net/figure/The-ten-myths-discussed-in-this-article_fig1_331656205

[65] Lexi Rubow, Rachel Shn Brinna Schofield and Samuelson Law (2015). Understanding Open Access.
https://authorsalliance.org/wpcontent/uploads/Documents/Guides/Authors%20Alliance%20-%20Understanding%20Open%20Access.pdf

[65b] Posada, A., and Chen, G. 2018. Inequality in Knowledge Production: The Integration of Academic Infrastructure by Big Publishers. ELPUB 2018, Toronto, Canada. ff10.4000/proceedings.elpub.2018.30ff. ffhal-01816707f).

[66] Lexi Rubow, Rachel Shn Brinna Schofield and Samuelson Law (2015). Understanding Open Access.
https://authorsalliance.org/wpcontent/uploads/Documents/Guides/Authors%20Alliance%20-%20Understanding%20Open%20Access.pdf

[66b] Boon, S. 2017. 21st Century Science Overload. Canadian Science Publishing - Blog. http://blog.cdnsciencepub.com/21st-century-science-overload/

[67] J. P. Tennant., et.al. Ten Myths around scholarly publishing. ResearchGate. 2019.
https://www.researchgate.net/figure/The-ten-myths-discussed-in-this-article_fig1_331656205

[68] J.E. Stiglitz and J.K. Rosengard. Economics of the Public Sector. Academia. 2015. https://www.academia.edu/37552359/Economics_of_the_Public_Sector_-_Joseph_E._Stiglitz

[69] Science Europe. Plan S. Making Full and Immediate Open Access a reality. Principles and Implementation. https://www.coalition-s.org/principles-and-implementation/

[70] Science Europe. Plan S. Making Full and Immediate Open Access a reality. Principles and Implementation. https://www.coalition-s.org/principles-and-implementation/

[71] Poynder2019.https://blogs.lse.ac.uk/impact of social sciences/2019/03/06/plan-s-and-the-global-south-what-do-countries-in-the-global-south-stand-to-gain-from-signing-up-to-europes-open-access-strategy/

[72] H. Debat, and D. Babini. Plan S in Latin America: A precautionary note. Peer J Preprints7:e27834v2https://doi.org/10.7287/peerj.preprints.27834v2

[73] Open Science. Publish-and-Read agreements (PARs) 2018. https://openscience.com/tag/publish-and-read-agreement/

[74] Egyptian Knowledge Bank (KBE). 2016 http://www.sis.gov.eg/section/496/4576?lang=en-us

[75] Ranking Web of Repositories. 2019. http://repositories.webometrics.info/en/node/30

[76] https://www.coalition-s.org/coalition-s-welcomes-its-first-african-member- and-receives-strong-support-from-the-african-academy-of-sciences/).

[77] International Science Council. Blog Interview with Dominique Babini. 2019. https://council.science/current/blog/plan-s-and-open-access-interview-with-dominique-babini

[78] International Science Council Plan, Advancing Science as A Global Good. 2019-2021 https://council.science/cms/2019/09/Advancing-Science-as-a-Global-Public-Good_ISC-Action-Plan-1.pdf

[79] M. Kende. Promoting the African Internet Economy (Nov) 2017. https://www.internetsociety.org/resources/doc/2017/africa-internet-economy/

[80] Academy of Science of South Africa (ASSAf). Mapping the current landscape of Open Science/Open Data initiatives in Africa: draft report (Online). 2019. DOI [to be added

[81] Why is AJOL needed? https://www.ajol.info/index.php/ajol/pages/views/about-AJOL-African-Journals-Online

[82] M. Banks. Journal editors share their views on the future of open access publishing. Professor Sanchez-Azofeifa of the University of Alberta, Canada. 2019. https://physicsworld.com/a/journal-editors-share-their-views-on-the-future-of-open-access-publishing/

[83] STISA 2024 – A guide for Africa to exploit its opportunities http://www.nepad.org/news/stisa-2024-guide-africa-exploit-its-opportunities

[84] T. Gura. "Citizen science: amateur experts". Nature. 496 (7444: 259–261.). 2013. doi:10.1038/nj7444-259a. PMID 23586092

[85] E. Hand. "Citizen science: People power". Nature. 466 (7307: 685–687) . 2010. doi:10.1038/466685a. PMID 20686547.

[86] C. Doyle; R. David; Y.Li; M. Luczak-Roesch; D. Anderson; C.M Pearson. Using the Web for Science in the Classroom: Online Citizen Science Participation in Teaching and Learning. 30 June 2019. doi:10.1145/3292522.3326022. ISBN 978-1-4503-6202-

[87] iNaturalist Connect with nature. https://www.inaturalist.org

[88] eBird. Discover a new world of birding. https://www.ebird.org

[89] Zooniverse. People Powered Research. https://www.zooniverse.org.

[90] Tropical Biology Association-led "Citizen Science in Africa" programme. http://www.tropical-biology.org/specialist-course/citizen-science-africa-making-work-conservation/

[91] ELIXIR https://elixir-europe.org

[92] H3A BioNet. https://www.h3abionet.org

[93] N. Mulder., et al. H3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa. Genome Research, (26/2, 271-277). 1916. doi: 10.1101/gr.196295.115

[94] Genome Research https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4728379/#MULDERGR196295C18

[95] The World Bank / World Development Indicators, Last update on 10/07/2019 https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=ZG&most_recent_value_desc=false&view=map

[96] National Institutes of Health (NIH) Common Fund. https://commonfund.nih.gov

[97] A. Ghouila, G.H.Siwo, J.B.D, Entfellner, S. Panji, K.A. Button-Simons , S.Z. Davis, F.M. Fadlelmola, The DREAM of Malaria Hackathon Participants. In Ferdig M.T and Mulder N. Hackathons as a means of accelerating scientific discoveries and knowledge

transfer. Genome Research. 2018. https://genome.cshlp.org/content/28/5/759.long

[98]EU budget 2014-2020. Programmes' Performance Overview. https://ec.europa.eu/budget/library/biblio/documents/2019/Programmes_performance_overview.pdf

[98b] TrustAfrica Annual Report 2015.
www.trustafrica.org › component › item › 3271-annual-report-2015

[99] AU STISA 2024, AU Agenda 63. http://www.nepad.org/news/stisa-2024-guide-africa-exploit-its-opportunities

[100]E. Opiyo. Why Internet Use is low in Africa. 2016. https://www.scidev.net/global/icts/news/internet-africa-local-content.html

[101]Forcepoint.https://www.forcepoint.com>cyber-edu>data-security

[102] S. Fonn, L.P. Ayiro, P.Cotton, A. Habib, P. M.F. Mbithi, A. Mtenje, B. Nawangwe, E.O. Ogunbodede, I. Olayinka, F. Golooba-Mutebi, A. Ezeh. Repositioning Africa in global knowledge research. 2017. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(18)31068-7/fulltext

[103]World Bank. A Decade of Development in sub-Saharan African Science, Technology, Engineering and Mathematics Research. Washington DC: World Bank Group, 2014. http://documents.worldbank.org/curated/en/237371468204551128/A-decade-of-development-in-sub-Saharan-African-science-technology-engineering-and-mathematics-research

[104]J. Mouton, J. Gaillard and M. van Lill. Functions of Science Granting Councils. In Sub-Saharan Africa. In Knowledge

Production and Contradictory Functions in Higher Education. N. Cloete, P. Maassen and T. Bailey. African Minds Higher Education Dynamics Series (Vol. 1). 2015. http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers15-08/010065145.pdf

[105] UNESCO Science Report 2015, Towards 2030. https://en.unesco.org/unesco_science_report

[106]UNESCO Science Report 2015, Towards 2030 https://en.unesco.org/unesco_science_report

[107Trust Africa. Post 2015 Dakar African Higher Education Summit Webinar Series - TrustAfrica. 2015. http://www.trustafrica.org/Fr/les-ressources2/actualites/item/3331-post-2015-dakar-african-higher-education-summit-webinar-series

[108] UNESCO Science Report 2015, Towards 2030 h t t p s : / / en.unesco.org/unesco_science_report

[109]J. Mouton, J. Gaillard and M. van Lill. Functions of Science Granting Councils. In Sub-Saharan Africa. In Knowledge Production and Contradictory Functions in Higher Education. N. Cloete, P. Maassen and T. Bailey. African Minds Higher Education Dynamics Series (Vol. 1). 2015. http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers15-08/010065145.pdf

[110] https://www.AIMS.ac.za

[111] Next Einstein Forum. https://nef.org

[112] GBIF – Global Biodiversity Information Facility. https://www.gbif.org

[113] DATA FIRST. https://www.datafirst.uct.ac.za

[114]African Academy of Sciences. AAS Open Research. Immediate and Transparent Publishing. 2019. https://aasopenresearch.org

[115] The Square kilometer Array. https://www.ska.ac.za

[116] OCSDNET – Open and Collaborative Science in Development Network. https://ocsdnet.org

[117]R. Kahn., and F. Mann. Developing Research Partnerships. Journal of Social Issues, (8/3, pp.4-10). 2010.

[118] Digital Economy 4 Africa. https://olc.worldbank.org/content/digital-economy-africa

[119] National Research Foundation- NRF. http://www.nrf.ac.za/media-room/news/statement-open-access-research-publications-national-research-foundation-nrf-fundedFunded Research,

[120] SABEN – South Africa Broadband Education Network. http://www.saben.ac.za

[121] Kenya Education Network. https://schools.kenet.or.ke/

[122] IDDO. Infectious Diseases Data Observatory 2019. https://www.iddo.org

[123] Data analytics. http://www.dataanalytics.org.uk

[124] Data Science Africa 2019. https://www.datascienceafrica.org

[125] CODATA-RDA Data Schools Initiative. 2019. http://africanopenscience.org.za/?p=1157) (https://www.rd-alliance.org/application-deadline-10-may-aosp school-research-data-science-and-codata-rda-data-schools-initiative).

[126a] University of Rwanda. African Centre of Excellence in Data Science. http://aceds.ur.ac.rw

[126b] University of Cape Town. Statistical Sciences. http://www.stats.uct.ac.za/stats/study/postgrad/masters/data-science.

[126c] University of the Western Cape. https://www.uwc.ac.za/Faculties/NS/News/Pages/Launch-of-the-Data-Science-programme.aspx .

[127] Principa. https://insights.principa.co.za/where-to-learn-essential-data-science- skills- online

[128] Integrated digitized Biocollections. https://www.idigbio.org/content/idigbio-and-data-carpentry-go-africa.

[129] Software Carpentry. https://software-carpentry.org/lessons/

[130] Software Carpentry Instructor Training. https://software-carpentry.org/blog/2018/03/Carpentries-Instructor-Training.html

[131 Carpentry Connect Johannesburg 2018. https://carpentryconnectza.org

[132] Data Carpentry Ethiopia. https://www.lentilcurtain.com/posts/data-carpentry-ethiopia/

[133] E. Chiware, and D. Becker. Research trends and collaborations by applied science researchers in South African universities of technology: 2007–2017. The Journal of Academic Librarianship, (44/4, pp.468-476). 2018. http://hdl.handle.net/11189/7099

[134] Directory of Open Access Journals. DOAJ. https://doaj.org

[135] African Journal Online. AJOL. https://www.ajol.info

[136] International Telecommunication Union.

[137] International Network for Government Scientific Advice (INGSA-Africa chapter) www.ingsa.org/chapters/ingsa-africa

[138] International Science Council/ Leading Integrated Research for Agenda 2030 (LIRA) https://council. science/ what-we- do/ funding-programmes/leading-integrated-research-for-agenda-2030-in-africa

[139] Open Data 4 Development. https://www.od4d.net

[140] J. M. Wafula. Africa Open Science Policy Framework (2019). In Press

[141] L. Nordling. Africa's fight for equality. Nature. (News, 05 May, 2015). 2015.

[142] International Data Charter- Principles. https:// opendatacharter.net/principles/

[143] Registrary of Open Access Repository Mandates and Policies. ROARMAP. http://roarmap.eprints.org

[144] Hackmann, H. and Boulton, G. (2015). Science for a sustainable and just world: a new framework for global science policy?. UNESCO Science Report: 2015 Towards 2030.

[145] The World Academy of Science TWAS. Open Data in a Big Data World. https://twas.org/sites/default/files/ open-data-in-big-data-world_short_en.pdf

[146] FORUS. Promising Approaches to Financing Development in the 21st Century. Research Report. May 2019. http://www.forus-international.org/en

[147] M.P. Jisc. An Open Access to scientific information: facilitating knowledge transfer and technological innovation from the academic to the private sector.Pasteur4OA. 2015. http://www.pasteur4oa.eu/sites/pasteur4oa/files/resource/Brief_OA%20and%20knowledge%20transfer%20to%20the%20private%20sector.pdf

[148] S.E. Ali-Khan, L. W. Harries and E.R. Gold Motivating participation in open science by examining researcher incentives. eLIFE.2017;6:E29310. 2017. doi: 10.7554/eLife.29319

[149] W. Koehrsen. Towards Data Science. Unintended Consequences and Goodhart's Law. Fe. 24. 2018. https://towardsdatascience.com/unintended-consequences-and-goodharts-law-68d60a94705c

[150] C. Goodhart. "Problems of Monetary Management: The U.K. Experience". In Courakis, Anthony S. (ed.). Inflation, Depression, and Economic Policy in the West. pp. 111–146. 1981.

[151] Reports – MLE Open Science – Altmetrics and Rewards. 2018. https://rio.jrc.ec.europa.eu/en/policy-support-facility/mle-open-science-altmetrics-and-rewards

[152] J. Mwelwa. Africa Open Science Capacity building and skills development Framework (2019). In Press.

[153] J. Mwelwa., I,. Smith and P. Molutsi. Advancing African Open Science through a network for Education and Skills.

2018. https://www.jointmindsconsult.com/wp-content/uploads/2017/01/MWELWA-AND-SMITH-2008.pdf

[154] A.Okune. Decolonising Scholarly Data and Publishing Infrastructures. University of California. Irvine. 2018. https://ocsdnet.org/decolonizing-scholarly-data-and-publishing-infrastructures/

[155] J. Mwelwa, J. and B. Spencer. A bilingual (Bemba/English) teaching resource: Realising agency from below through teaching materials designed to challenge the hegemony of English. Language Matters, (44/3, pp. 51-68). 2013. https://doi.org/10.1080/10228195.2013.840011

156 International Science Council. Starting again with scientific publishing. Q and A with Alex Freeman. Blog. 2019. https://council.science/current/blog/starting-again-with-scientific-publishing-qa-with-alex-freeman

[157] International Science Council Plan, Advancing Science as A Global Good. 2019-2021 https://council.science/cms/2019/09/Advancing-Science-as-a-Global-Public-Good_ISC-Action-Plan-1.pdf

[158] New Incentives and Rewards for better and Open Science: 'The Dean's Perspective' https://www.intensivistendagen.nl/wp-content/uploads/2019/04/Frank-Miedema-Science-in-transition.pdf

[159] International Science Council Plan, Advancing Science as A Global Good. 2019-2021 https://council.science/cms/2019/09/Advancing-Science-as-a-Global-Public-Good_ISC-Action-Plan-1.pdf

[160] Elsevier and Ipsos MORI Research Futures – Drivers and Scenarios for the next decade. Full

Report. 2019 https:// www.elsevier.com/__data/assets/pdf_file/0011/847955/Research_Futures_full_report_Feb2019.pdf

[161]OECD. Making Open Science a Reality. OECD Science, Technology and Industry. Policy Papers No.25. 2015. http://dx.doi.org/10.1787/5jrs2f963zs1-en https://www.oecd-ilibrary.org/docserver/5jrs2f963zs1-en.pdf?expires=1571939511&id=id&accname=guest&checksum=9C72B09E1BD539949C8A80AECEC28F00 .

[162]C. Neylon. Governing Open Science; an introduction. 2019. In Chan, L., Okuna, A., Hillyer, R., Albornoz, D. and Posada, A. (eds) 2019. Contextualising Openness. University of Ottawa Press and International Development Research Canada, 125-132. 2019.

[163]N.D. Bowman and J.R.Keene. A Layered Framework for considering Open Science Practices. 2018. https://www.tandfonline.com/doi/full/10.1080/08824096.2018.1513273

[164]Wiley Researchers. What Does Open Science Require? Cultural Change. Interview of Henriikka Mustajoki, Head of Development at the Federation of Finnish Learned Societies by Chris Graf – Director, Research integrity and Publishing at Wiley. https://www.wiley.com/network/researchers/licensing-and-open-access/what-does-open-science-require-cultural-change

[165]M.L. Hackert., L. Van Meerfelt., J.R. Helliwell, J.R., and B. MaMahon. 2016. Crystallography and Open Data in a Big Data World. International Union of Crystallography. 2015.

[166] J.R. Helliwell.,J. McMahon.,J.M Guss., and L.M.J. Kroon-Battenburg. The Science is in the data, International Union of Crystallograpy Journal (4/6, 714-722). 2017. (doi: 10.1107/S2052252517013690)

[167] G8 Open Data Charter
https://opendatacharter.net › g8-open-data-charter

[168] A. Lancaster. Open Science and its Discontents. Ronin Institute: Reinventing Academia. 28 June, 2016. http://roninstitute.org/open-science-and-its-discontents/1383/

[169] D.L. Longo,and J.M. Drazen. Data Sharing. The New England Journal of Medicine. (374, 276-277).2016. https://www.nejm.org/doi/full/10.1056/NEJMe1516564

[170a] D. Tyfield.Transition to Science 2.0: "Remoralizing" the Economy of Science. Economic Aspects of Science, (7/1).2013.
https://spontaneousgenerations.library.utoronto.ca/index.php/SpontaneousGenerations/article/view/19664

[170b] D. Tyfield.Transition to Science 2.0: "Remoralizing" the Economy of Science. Economic Aspects of Science, (7/1).2013.
https://spontaneousgenerations.library.utoronto.ca/index.php/SpontaneousGenerations/article/view/19664

[171] Plan Ceibal – Uruguay. https://www.ceibal.edu.uy/en/institucional

[172] British Academy and Royal Society. 2017. Data management and use: Governance in the 21st century.
https://royalsociety.org/~/media/policy/projects/data-governance/data-management-governance.pdf

[173] http://p3g2.org

10. Appendices

10.1: The Report Team

Geoffrey Boulton is Regius Professor of Geology Emeritus in the University of Edinburgh, UK, and a former Vice Principal. He is a Fellow of the Royal Society and the Royal Society of Edinburgh, Scotland's national academy. He is a member of the Governing Board of the International Science Council and a member of the AOSP Advisory Council. He recently stepped down as President of CODATA, where he was a champion of Open Data and Open Science. He has been a member of the UK Prime Minister's Council for Science and Technology and Chair of the Royal Society's Science Policy Centre. He has chaired relevant influential reports, for the Royal Society (Science as an Open Enterprise) and for Science International (Open Data in a Big Data World), which received over 120 endorsements from major science bodies worldwide.

Cheikh Loucoubar is a mathematician with PhD in statistical genetics at the Institut Pasteur de Dakar, heading the biostatistics, bioinformatics and modelling group. He teaches statistical genetics, biostatistics and R statistical software in master programs at the University Gaston Berger of Saint Louis and University Cheikh Anta Diop of Dakar. He researches on statistical methods and tools for family-based genetics of infectious diseases. His team focuses on applied mathematics for biomedical research, programming, database systems and web application development. He his Co-Pi in H3ABioNet (Human

Health and Heredity in Africa, The Bioinformatic Network) running the Dakar Node.

Joseph Mwelwa is PhD- and senior partner at Joint Minds Consult, a research, education and training institute he founded to provide capacity development and support in educational research, policy, curriculum development, training, knowledge development and management and, post graduate student support in sub-Saharan Africa. He has been a member of the African Open Science Platform Technical Advisory Board where he developed a report on capacity building and co-authored the Joint Minds Consult position paper that guided Botswana towards an Open Science and Open Data Strategy. He has also conducted workshops on developing national discourse on Open Science and Open Data strategies in South Africa, Ghana and Ethiopia using Botswana as a case study.

Joseph Muliaro Wafula holds a PhD, is Associate Professor of computing at JKUAT, founder Director of the ICT Centre of Excellence and Open Data (iCEOD) and a cyber security expert and data engineer. His recent works include review of the Global Biodiversity Information Facility's activities and accomplishments; development of an African Policy Framework and Policy Roadmap for open Science and open data; and a performing skills audit and strategy development to fill the gaps in Northern Corridor Integrated Project (Kenya, Rwanda, South Sudan, and Uganda). He is a member of CODATA International and of the editorial board of the Data Science Journal. He is the lead Author of the African Open Science Policy Framework- A guide for African States on Open Science Policies and Practices.

Nicholas Ozor holds a B. Agric. (Nig., First Class Honours); MSc, Agricultural Administration (Nig., Distinction) Ph.D, International & Rural Development and Agricultural Extension (Reading, UK & Nigeria respectively) and is the Executive Director of the African Technology Policy Studies Network (ATPS). He was formerly a Senior lecturer in the Department of Agricultural Extension, University of Nigeria, Nsukka. Dr Ozor leads many internationally funded research projects bordering on science, technology and innovation (STI). He is a member of many professional organizations and has published over 100 articles in reputable international journals, book chapters, and other multimedia. Dr Ozor has raised over US$ 50 million in grants to support development work in Africa. He sits on the Board of many international bodies including UNESCO.

Maurice Bolo holds a PhD in Science, Technology and Innovation Policy and over 15 years' work experience. A Visiting Research Fellow in the Department of Policy and Practice (DPP) of the Open University (UK) and a Research Associate at the Innogen Institute (Edinburgh, Scotland), Dr. Bolo has a vast international consultancy experience.

10.2 Key Stakeholders

This appendix summarises the various groups in Africa that have a stake in the operation of science systems, identifies their motivations and interests and whether and how they might need to be engaged with an open science initiative.

10.2.1 Policymaker and policy influencers

a) ***National governments.*** A two-fold case should be made to national governments, based on the inevitability of confronting the imperatives of digital science as a key to the 4th industrial revolution and the economic and social benefits it offers; and the potential impact on science to be derived from a collaborative, intra-African open science initiative. There is a strong argument that the developmental goals that are crucial to Africa's future will depend upon the evolution of a bold and vigorous African science community, in which the diversity of Africa becomes a strength rather than a weakness. The willingness of governments to permit national Science Granting Councils to coordinate their activities and funding with common purpose is one that would need to be tested. Such a case would require a cost-benefit analysis, where the capacity of governments, strengthened by acting in concert. An approach the World Bank for loan finance for vital infrastructure and for planning support could be a considerable asset. In addition, governments may be required to pass or amend legislation as a national frame for open science, whilst the Granting Councils may need to make regulations or create policies to create the environment within which open science can flourish on such matters as privacy, intellectual property, access to services, publishing etc.

b) ***Academies.*** National academies and Africa-wide academies (the African Academy of Sciences and the Network of African Science Academies) play influential roles in representing and influencing science at national and international levels and have a high international profile. The Science Granting Councils' engagement with relevant academies would be important in developing a common, consensual approach that might be important in influencing national governments and international donors.

c) ***Public Sector data holders.*** Much of the data held by public sector and national statistics bodies is also of great value to the research community, and its contribution to understanding, particularly of social phenomena (e.g. resilient cities, disaster risk, precision medicine, agriculture etc), is substantial. It would be helpful in exploring the accessibility for research of such data if the Councils' could fund a survey of data holdings, of accessibility, of legal use and the policies of African governments about such data. The availability and use of such data could be of great value in solutions-oriented work on SDGs.

10.2.2 Practitioners

a) Universities. There are three key related issues:

- They need to be persuaded of the long-term benefits of a strategy that would be likely to strongly perturb patterns of institutional and personal funding, and of the policies and incentives that influence their behaviour. Many of the norms and habits of academic researchers are challenged by open science principles, although recent years have seen substantial changes of attitude as the open science movement has gained momentum. An important role for the Councils would be to convene and lead deliberative dialogue with scientists and their institutions with the objective of evolving a committed, shared purpose. If successful, it would be a strong success factor.

- Universities may be, or may develop into important nodes of activity in open science. For example, they may be centres for high performance computing, data science analytic and AI skills, cloud facilities and significant databases. An open science initiative would need to engage with and rely on such facilities. If the Councils were to act in a planning and coordinating role, they would need to work with universities in discussing how this might best be done.

- Universities are the obvious locations for much of the higher education and training that is required in data science and technology. It is important to recognize that the purpose of this is not solely to train specialists for the science system, but also to create a pipeline of skills for public and private sector roles. For longer term sustainability of skills and knowledge development, innovations in the school system from primary through secondary level, with incremental studies on data science and technology required by open science. (See recommendations in the UK system and Plan Ceibal – Uruguay) [171].

- Again, the Council's could play a role in mapping intra-African capacities, working with expert groups (University Centres-CODATA-RDA-WDS) in defining the curriculum and potentially seeking support from international donors in building up the manpower potentials at national level. There is of course a major manpower-planning role to be taken up by national governments.

b) ***The Private Sector.*** The private business sector is a major beneficiary of the digital technologies that are the drivers of the 4th industrial revolution. Its major concerns are access to the skills described in this report and to the solutions and approaches that open science is designed to deliver. The perspective of private sector, as a driver of national economies, is therefore an important consideration in developing an open science initiative. In the event of an SCGI initiative, there should be early engagement with private data sector as described below:

- **Public data acquired by the private sector.** The dramatically increasing role of the private sector in the world of data is striking. The acquisition of the copyright to publicly funded data by scientific publishers, and their business model, is a major development of considerable concern given the precedent of their excessive profits from journal publishing. It is important that this process is understood in relation to the extent that it affects ownership and access to data in Africa. As we have already commented, the paucity of African data holdings, which, if not corrected, is a serious barrier to entry into a data-intensive world.

- **Commercial science publishers**

  We have already commented on this issue (4.2) and expressed our concern about a business model for subscription of open-access journals. The ISC is launching a project on scientific publishing. It is important that African representatives take part in it.

- **Private data in private hands**

It is not necessarily true that the public/private interface is impermeable to data flow across it. There are several business sectors where a creative and productive flow of data and ideas across the public/private boundary takes place. Given the importance of innovation for Africa, both in social and economic terms, it is important to examine this interface in ways that establish where processes are sub-optimal, governed by rules that inhibit benefit. This could be one of a series of small research projects that would be of great value in determining how to maximize scientific, social and economic benefit from Africa's data resources and identifying where blockages to benefit occur.

- **Computing and network technology companies**

Commercial equipment service providers are important in the provision of existing infrastructural components and will be important in further developments involving hardware systems. "Cloud" systems connected by large bandwidth Wide Area Networks (WANs), are important in hosting software systems that enable data analysis and in providing access to massive data collections. Private companies such as Microsoft, Amazon, Mozilla and Google are potential partners in providing required e-infrastructure services, though care will be needed to avoid becoming dependent on any particular service provider whose business model may diverge from what is needed. It is crucial to be an "intelligent consumer" that understands the technical issues sufficiently deeply to be able to engage with commercial providers in identifying optimal solutions, rather than the most profitable solution for the supplier. A collective approach to commercial suppliers can be highly advantageous in cost effective procurement.

10.2.3 International supporters

These have long played an important role in supporting the development of African scientists. They may support the development of individuals through scholarships and fellowships, often held in the funder's country, and through funding research projects including projects involving joint work between African scientists and those from the funder's country. They may focus on specific fields of research or they may be prepared to give support across a wide range of disciplines, but they overwhelmingly fund science (natural, social, medical or engineering). They may fund institutional developments by supporting the development of university research and by developing and improving processes within national or pan-African academies or Science Granting Councils. Principal groups include:

- Governmental and Intergovernmental agencies. The World Bank has a major project in Africa that supports activities relating to the digital economy. Governmental agencies that have potential to support science system development include the International Development Research Centre (IDRC - Canada); French Development Agency (AfD), Swedish International Development Cooperation Agency (Sida), Department for International Development (DFID - UK), and Norwegian Agency for Development Cooperation (NORAD).
- Private Foundations and charities. These generally concentrate on specific areas, frequently in health and medicine. They include the Wellcome Trust, the Gatsby Foundation and the International Foundation for Science.
- National academies. National academies that support science and scientists in Africa, particularly early career scientists in include the Royal Society (UK), the National Academies of Science (US), the Japan Society for the Promotion of Science and most recently the Chinese Academy of Sciences.
- International scientific bodies. There are a number of international bodies that are representative of the global scientific community, and though they are not generally sources of major funding, they have valuable capacities that could readily be leveraged in

support of an African open science initiative if the Councils chose to promote it. They include:

- UNESCO is the United Nations agency with responsibility for scientific affairs and is the voice of science in international governance. It has considerable convening power that it uses to promote scientific developments that it deems to be of international significance, a standing that it would certainly recognize in a SGCI initiative on open science, an area that UNESCO is considering for a formal recommendation.

- The International Science Council (ISC) is the senior representative body for international science, encompassing the natural and social sciences. It promotes and sponsors major programmes for international science and collaborates widely with major international bodies such as the UN, UNESCO and the WMO. It is currently developing several programmes that would be of high relevance to open science in Africa: on data integration for interdisciplinary science, on global data governance and on open platforms, and which also supports the African Open Science Platform (AOSP). ISC members include many national academies and international disciplinary and interdisciplinary bodies, some of which currently support science in Africa.

- International data bodies: The ISC Committee on Data (CODATA) is a quasi-autonomous member organization that convenes great expertise in open data, the frontiers of data science, capacity building and data policies and practice. It is currently active in Africa, where it also supports AOSP. The ISC World Data System (WDS) is primarily concerned with the crucial issue of scientific databases, their creation, management and service operation. It allocates the CoreTrustSeal benchmark for databases. The Research Data Alliance (RDA) is an international consortium of individual members that focuses on the crucial issue of data integration for individual domains of science.

10.3 The Questionnaire to Science Granting Councils

10.3.1 The circulated questionnaire

Questionnaire: Open Science in Research and Development in Africa.

This questionnaire is part of a study commissioned by the African Technology Policy Studies Network (ATPS) working in partnership with The Scinnovent Centre under the Science Granting Councils Initiative (SGCI). The SGCs Initiative is jointly funded by the United Kingdom's Department for International Development (DFID), Canada's International Development Research Centre (IDRC), South Africa's National Research Foundation (NRF) and the Swedish International Cooperation (Sida) with a mandate to strengthen the capacities of Science Granting Councils in sub-Saharan Africa in order to support research and evidence- based policies that will contribute to economic and social development. In the era of the Internet, open science, open publishing and open data frame humanity's thinking about science and the potential it holds for development and innovation. The aim of this study is to develop an framework for operationalizing open science in the 15 SGCI member countries.

Purpose

The questionnaire seeks to collect expert views on the potential of an open science initiative as a means of enhancing the work of African Science Granting Councils in driving innovation and development. The results will be used to produce (1) a report that will inform the debate on Open Science in Research and Innovation for Development to be held at the annual forum of the Councils in Dar es Salaam on 11-15 November 2019, and policy briefs to inform further debates and operationalization of open science for development and innovation in SGCI member countries in Africa. As an expert in research and innovation in your country and beyond, we would appreciate your brief input into the report by responding to a number of questions. The time window for the report is very short, and we apologise for the short notice for this request, but we would be most grateful if you would collaborate. All responses will be anonymised in the analysis and project reports. You will however be included in the distribution lists for final project outputs in the first quarter of 2019.

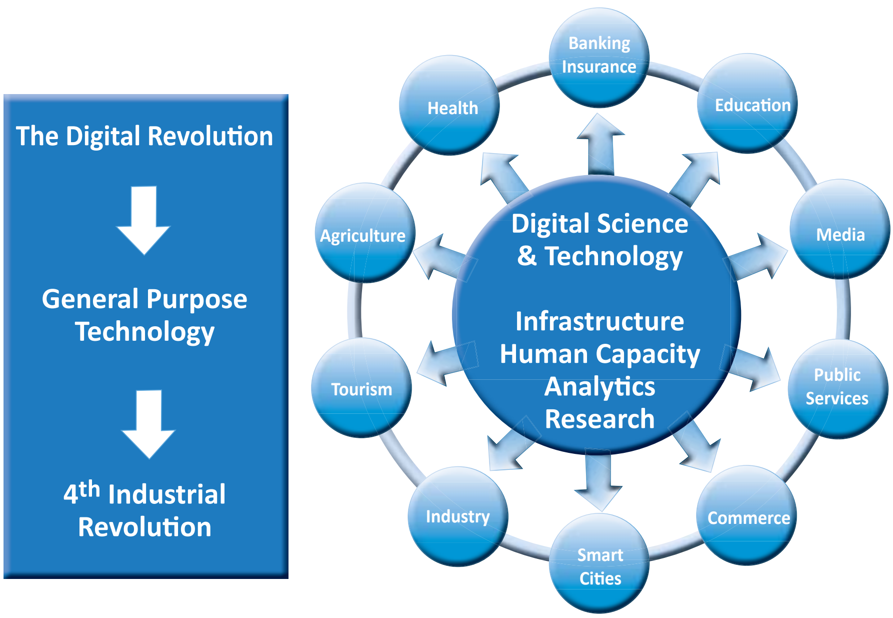
First however, we briefly describe the hypothesis that is being investigated by the report and then list the questions that we would like to you to respond to.

Hypothesis: the benefits of open science to Africa

a) The fourth industrial revolution is powered by the tools of the digital revolution.

These tools a general-purpose technology that continually transforms itself, progressively penetrating new domains, boosting productivity across all sectors and industries because of their cost effectiveness. They are globally pervasive, with profound economic and social implications that fundamentally disrupt pre-existing norms. They have created the so-called 4th industrial revolution, the impacts of which are shown in the diagram. Africa must adapt to exploit the potential of this revolution.

Revolutionising, economies, societies, lives



b) A collaborative "Open Science" area would be an efficient response to this challenge

Enhancing intra-African collaboration (STISA2024) through an open science initiative that stimulates and enables data-sharing, open access to scientific results and federated, shared infrastructure would be a powerful means of harnessing the technologies of the digital revolution to invigorate and release the potentials of African science, to stimulate innovation and creativity, and to dynamise economic and social development. It would create:

i. efficiencies of scale in planning, procurement and provision;
ii. scaling-up through collaboration and shared capacities;
iii. stimulating creativity through interaction of diverse groups;
iv. amplifying impact through common purpose and voice;
v. building consortia and collaborations with a greater critical mass;
vi. support from a shared capacity in cutting-edge data science.

The questions

a) Do you agree, in principle, with the above hypothesis?
b) Has your country experienced open science at a national level?
c) If yes, what has been your country experience?
d) What are the principle barriers to achieving a collaborative open science area ?
e) Are the barriers surmountable?
f) If yes, explain how?
g) Is there a mood, amongst politicians and science leaders, to consider and commit to intra-African collaboration on the scale required?
h) What role could the SGCs play in fostering a collaborative open science area among the 15 SGCI member countries?
i) What role would you play as an SGCI member country in a collaborative open science area?
j) What would be the pros and cons of a 15 SGCI collaborative open science area?

Priorities

What, in order of priority, are the key issues that would need to be prioritised in a collaborative open science area?

Use a scale of 1 – 10. 1 being the highest, 10 being the lowest

- Wide area networks,
- Open science policies
- Incentives for researchers
- Capacity building
- Cloud computing,
- High Performance Computing (HPC)
- Multi-national mission-led programmes (e.g. STISA2024),
- Institutional commitments (e.g. universities),
- Commitment of external funders etc].
- Collaboration among the 15 SGCI member countries

Respondents details Bio details

Name .... (Optional)
Organisation:
Position in organisation:
Years in organization:
Country:
Category:

Government/Policy maker	
R & D organisation	
Academic Institution	
Funding Institution	x
National Science Council	
Continental/Global Agency	
Other (Please specify)	

10.3.2 Respondents

Responses were sought from 15 and returned from 13 SGCI members and one from the Sudan Bank for Development. The respondents were: Kenya, Burundi, Uganda, Tanzania, Côte d'Ivoire, Botswana, Ghana, Zambia, Mozambique, Sudan, Sierra Leon, Malawi and Namibia.

10.3.3 Responses to questions

Table 10.1 below summarises our evaluation of the detailed responses in relation to four fundamental issues: barriers to open science; possible roles for SGCs; merits of an open science collaboration; demerits of an open science collaboration. The full responses are available in an XL file, and are available on request.

Table 10.1. Summary of responses to key questions

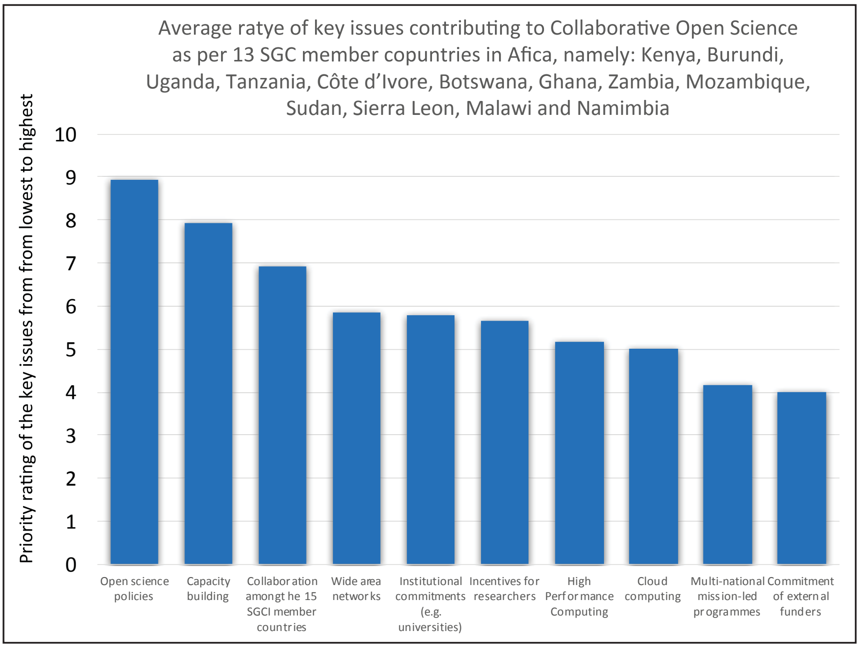


Figure 10.1. Average ranking of open science priorities.

Figure 10.2. Ranking of priorities by individual countries

Principle barriers to achieving collaborative open science in Africa:
• African states at different levels of development • Lack of political commitment among member states • Open science is a new paradigm and not yet fully understood • Researchers are anxious about their career prospects and how open science would affect this. They are concerned about the ownership of results, technologies generated, and the importance of prime authorship • Lack of adequate capacities both human and ICT infrastructure • Fear of loss of Intellectual Property • Funders, universities and research institutions pressure researchers to publish in high impact factor journals, which are often not open-access journals • Few research databases and journals based in Africa • Lack of policies at national and institutional level for coordinating research • Lack of awareness among policy makers of open science • Lack of open science culture among researchers that needs to be inculcated • Lack of enabling environment eg policies, regulations and infrastructure, human capacities • Absence of mechanisms to effectively drive/achieve collaborative open science • Country specific frameworks that are still traditional may potentially affect operationalization of an open science area
Roles suggested for SGCs in fostering collaborative open science:
• Facilitator • lobbying for policy reform • resource mobilization • creating awareness • facilitate discussions on restructuring and promotion of open science among researchers • coordinating implementation of open science initiatives • offer visibility of national science councils • supporting existing systems and policy environments to embrace open science • facilitating development of appropriate frameworks to drive open science

SGCIs identified the following roles they could play in support of open science in Africa:
• Advocacy for a national open science policy • Stimulate open science support amongst stakeholders • Fund both research projects involving pro-open science researchers • Support open science sensitization forums • Reform policies to accommodate the inevitable changes due to open science • Factor in principles of open science in national and co-funding grants initiatives • Conduct workshops that create awareness about open science benefits and IP • Stimulate and fund joint research and ensure that the results are published in open access to enable reach and access by all stakeholders • Coordinate and facilitate all open science activities • Lobby Government to embrace open science
Merits of a 15 SGCI collaborative open science area:
• Sharing of resources and experiences. • Higher probability of innovation • Wide dissemination • Speedy publication • Easy access to science research in developing countries • Free access to scientific knowledge • Enable the results of research and innovation to be disseminated more rapidly and widely thus contributing to knowledge economy. • Speaking with one voice in facilitating transformation towards open science e.g. developing common open science policies • Maximizing scientific output utilization • Providing a platform for cost effectiveness through resource, capacity and experience sharing and exchange

Demerits of a 15 SGCI collaborative open science area:

- Member countries are at different levels in terms of readiness, technological capacities and governing policies
- Lack of mutual trust
- Expensive for researchers
- Quality concerns
- Financial issue for journals
- Gaps in human, infrastructure and financial resources
- High costs of internet
- Poor last mile connectivity in most institutions
- Inadequate availability locally generated research results
- Different levels of NRENs capabilities
- Researchers seek to publish in journals with a high impact factor which open access journals do not have.
- There is potential disagreements on the principles and practices related to open science and access to information
- Lack of confidentiality

Some ATPS Research Paper Series

- New Approaches for Funding Research and Innovation in Africa, [Julius Mugwagwa, Geoffrey Banda, Nicholas Ozor, Maurice Bolo, and Ruth Oriama] ATPS Research Paper No. 30

- Towards Effective Public-Private Partnerships in Research and Innovation, A Perspective for African Science Granting Councils, [Banji Oyeleran-Oyenyinka, Bertha Vallejo, Shruti Vasudev] ATPS Research Paper No. 29

- A Review of the Kenyan Policy Environment for Off-grid Solar PY, [Kevin Urama, Nicholas Ozor and Edith Kirumba], ATPS RESEARCH PAPER No.28

- Design and Analysis of a 1MW Grid-Connected Solar PV System in Ghana, [Ebenezer Nyarko Kumi, Abeeku Brew-Hammond], ATPS RESEARCH PAPER No. 27

- Farmers' Response and their Adaptation Strategies to Climate Change in Mafeteng District, Lesotho, Farmers' Response and their Adaptation Strategies to Climate Change in Mafeteng District, Lesotho, [Sekaleli T.S.T, Sebusi K.] ATPS Research Paper No. 26

- The The Rationale and Capacity of Pastoral Community Innovative Adaptation to Climate Change in Ethiopia, [Tibebu Solomon], ATPS Research Paper No. 25

- Contribution of Information and Communication Technologies (ICTs) in Climate Change Awareness in Seke and Murewa Districts of Zimbabwe, (Shakespear Mudombi, Mammo Muchie), ATPS RESEARCH PAPER No. 24

- Vulnerability and adaptation strategies to climate variability and change of the Bos-taurus dairy genotypes under diverse production environments in Kenya. [Kiplangat Ngeno, Bockline O. Bebe], ATPS RESEARCH PAPER No. 23

- Indigenous Rain Water Harvesting Practices for Climate Adaptation and Food Security in Dry Areas: The Case of Bahi District, [Deusdedit Kibassa], ATPS Research Paper No. 22

- Influencing National Legislation, Policies, Strategies and Programmes to Ensure Appropriate Protection and Benefit Sharing of Traditional Herbal Medicinal Knowledge (THMK) with and by Traditional Herbalists in Uganda, [Wanakwakwa J., Munabi C., Lwanga H., Muhumuza J., Gateese T.], ATPS RESEARCH PAPER No. 21

- Analysis of Traditional Healers in Lesotho: Implications on Intellectual Property Systems, By Pitso Masupha, Lefa Thamae and Mofihli Phaqane. ATPS Research Paper Series No. 20

- Intellectual Property, Traditional Knowledge, Access Benefit Sharing Policy Environment in Eight Countries in Eastern and Southern Africa: Swaziland, Kenya, Lesotho, Mozambique, Malawi, Tanzania, Uganda and Ethiopia, By Joseph M. Wekundah. ATPS Research Paper Series No. 19

- Assessment of Possible Intellectual Property Protection Options of Traditional Knowledge System in Ethiopia: Special Reference in Herbal Medicine for Livestock, By Tibebu Solomon. ATPS Research Paper Series No. 18

- National Policies and Legal Frameworks Governing Traditional Knowledge and Effective Intellectual Property Systems in Southern and Eastern Africa: The Case of Traditional Healers in Tanzania, By Georges S. Shemdoe (PhD) and Loy Mhando. ATPS Research Paper Series No. 17

- Assessment of the impacts and Adaptive Capacity of the Machobane Farming System to Climate Change in Lesotho, By Sissay B. Mekbib, Adesola O. Olaleye, Motlatsi N. Mokhothu, Masia Johane, Spirit B. Tilai and Taddese Wondimu. ATPS Research Paper Series No. 16

- Farmers' Perceptions of Climate Change and Adaptation Strategies in Northern Nigeria: An Empirical Assessment, By Dr. Blessing Kaletapwa Farauta, Chukwudumebi Leticia Egbule, Dr. Yusuf Lawan Idrisa and Dr. Victoria, Chinwe Agu. ATPS Research Paper Series No. 15