FINAL SYNTHESIS REPORT

ASSESSING STI METRICS IN AFRICA

MARCH 2021

Disclaimer

This document is an output from a project funded by the East Africa Research Fund, supported by the UK Foreign Commonwealth and Development Office (FCDO) through the East Africa Research and Innovation Hub. However, the views expressed, and information contained in it is not necessarily those of, or endorsed by FCDO, which can accept no responsibility for such views or information or for any reliance placed on them.

The project was implemented by the African Centre for Technology Studies, the Science, Policy Research Unit at the University of Sussex and the Africa Research and Impact Network.

Joanes Atela[1, 2] j.atela@acts-net.org

Nora Ndege[1, 2] n.ndege@acts-net.org

Tommaso Ciarli[3] t.ciarli@sussex.ac.uk

Diego Chavarro[4] dchavarro@gmail.com

[1] African Centre for Technology Studies

[2] Africa Research and Impact Network

[3] Science Policy Research Unit, University of Sussex

[4] Independent Consultant

Executive Summary

Context

This report provides a synthesis of the process and outcomes of the 'Assessing Science and Technology Metrics in Africa' study, aimed at developing an integrated set of indicators (scoreboard) that can be applied to assess Science Technology and Innovation (STI) progress and performance in African countries. Developing STI indicators builds on the fact that most African countries are aspiring to transition to sustainable industrialized economies by the year 2030 through STI, thus the need to understand which investments can yield benefits that align to this development ambition and the broader Sustainable Development Goals (SDGs). The role of STI as a driver of a knowledge-based economy and achievement of the SDGs is articulated in the continent's Agenda 2063 and operationalized under the STI *Strategy for Africa 2024* STISA-2024).

Africa's STI landscape is characterized by mixed progress differentiated across countries. A fundamental milestone is the establishment of national-level policies and agencies to guide investments, monitoring, and reporting. Most countries have established the National Science Granting Councils (SGCs) as the key agencies coordinating national STI-related activities. However, majority of these countries still face challenges in translating these national plans into action, coupled with little evidence on what is going on in practice. Several ongoing innovative activities/initiatives within many countries are poorly documented, measured, or reported, thus lack of clarity on how interventions contribute to the countries' development and overall SDGs. The need to assess ongoing interventions as a way of prioritising new ones is central to Africa's ST&I progress. This study, therefore, aimed to first assess the various approaches/methodological frameworks that have been used to assess STI, and build on this to develop a suitable set of standards and key indicators for making a comparable and robust assessment of STI in Africa.

Methodological Approach

The study involved nine (9) main steps, with the first seven (7) steps focused on the development of a suitable scoreboard for African countries while the last two (2) steps involved uptake and decision support for countries. Guided by these steps, the study was anchored on co-production, where stakeholders were engaged in the design, review, and validation of the study activities and outcomes. The specific methods applied included: in-depth literature review; secondary data inventory including review of existing ST&I scoreboards (e.g., AIOIII, GII, OECD scoreboard, UNESCO science report, RICYT scoreboard, national scoreboards, among others); and academic literature (reviews, assessments, citing documents) in Scopus, Google Scholar, WoS, and Scielo. Specific key informant interviews (from a select number of stakeholders in Kenya, Nigeria, South Africa, Zambia, and Rwanda), and a non-representative survey of STI data users were interviewed in addition to exploratory surveys.

To guide the development of a suitable scoreboard, we first developed an STI framework that clearly describes and defines Africa's STI ecosystem, and measurements. The framework identified indicator categories in terms of input, enablers, linkages, and outputs, allowing for a more targeted and strategic assessment and stakeholder engagement in the scoreboard construction. Guided by the framework, the study then identified STI indicators available for African countries drawing on a review of fifteen (15) existing scoreboards which were assessed in terms of relevance, completeness, and appropriateness to both the country and regional contexts.

Key Findings

Assessment of existing initiatives

In terms of existing/ongoing initiatives that attempt to measure STI in Africa, the study identified fifteen (15) scoreboards that represent existing initiatives, some with international and others Africa-specific focus. These include initiatives by the United Nations Economic Commission for Africa (ECA), and Africa-specific ones such as the Research & Development (R&D) and Innovation surveys conducted under the African Science and Technology Innovation Indicators (ASTII) initiative supported by AUDA-NEPAD, among others. Currently, there are efforts to put in place a new initiative, Afri-look, which will continue to support R&D surveys and build the capacity of national experts to understand and interpret R&D/innovation survey findings.

Nonetheless, further inquiries through key informants and stakeholder dialogues, especially with the fifteen (15) SGCs from sub-Saharan Africa, revealed that despite the existing initiatives, many countries seem to be at very early stages of developing mechanisms for measuring STI. At the moment, STI measurements are limited to setting benchmarks aimed at certifying the quality of academic courses and research priorities. Most of these benchmarks are also largely qualitative, generally involving provisions such as aligning research to the development goals of a country, among others.

The main approaches applied in developing the existing scoreboards are the input-output framework and the national system of innovation. These have been extensively used to develop scoreboards and rank STI outlooks. Despite being dominant in framing the scoreboards in Africa, the approaches lack robust components for assessing the effectiveness of the scoreboards over time. As such, the assessment of these scoreboards remains relatively uncoordinated and unsystematic. A handful of studies have attempted to assess the STI scoreboards, such as the European Innovation Scoreboard (EIS) and the Global Innovation Index (GII) in terms of quality and highlighting challenges such as lack of a clear framework, arbitrary selection of indicators, among others.

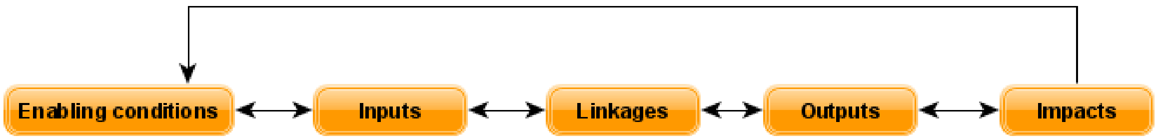
There are no studies that have assessed these scoreboards in relation to the relevance, completeness, and appropriateness of the indicators that they use. This is true for scoreboards measuring an African or any other economy. However, most scoreboards are populated by a large number of footnotes that address specific issues with the indicators. Other limitations associated with the existing scoreboards include lack of clear linkages to SDGs (i.e., 5% of indicators relate to SDGs and lack of gender considerations, with only 3% diverging on gender, which is again focused on personnel). Scoreboards are also not well linked to the STISA 2024, something that could be a result of the lack of systemic assessments to ensure indicators are relevant and aligned to prevailing policies.

Ultimately, the lack of systematic assessment impedes the adoption and use of the scoreboards for various users. The study revealed that policymakers remain the main users of these scoreboards even though others such as research organizations, national statistical offices, and educational institutions also use the scoreboards. While the use of the scoreboards for policy initiatives is a key step, evidence and discussions with stakeholders indicate that the intensity of usage by most policy bodies is relatively ad-hoc and based on specific needs rather than for planning support. The ad-hoc use is also fueled by ad-hoc data collection towards these scoreboards since the current R&D surveys performed by countries are yet to be fully institutionalized. This means that these scoreboards are yet to fully support countries in making investment decisions and prioritization, further limiting widescale usage especially by some important stakeholders such as the private sector.

Based on the literature and stakeholder consultation, multiple challenges impede the use of existing scoreboards. While literature tends to highlight broader conceptual and comparability challenges, ***stakeholders*** were more inclined towards contextual challenges mainly informed by their practical experiences. Stakeholders particularly reiterated the need to develop contextually relevant indicators that are co-produced with the users of these indicators based on a well-understood framework. It was emphasized that indicators need to go beyond narrow measurements (e.g., Gross Expenditure on Research and Development) and be more integrated and inclusive, especially paying attention to the informal sector. While the informal sector continues to play a key role in the continent, there is a general lack of proper understanding of how activities in this sector can be assessed. As such, there is a need to assess and strengthen the capabilities and capacities required to generate, uptake of the STI indicators, and use them.

Construction of Scoreboard

Building on the assessment of initiatives, the study developed an integrated framework to guide the construction of an integrated scoreboard, taking into account the gaps identified in the existing initiatives. The logical framework that was proposed allows the analysis of the relationship between STI inputs and outputs, adding categories that identify societal outcomes. The framework also allows for scoreboards that differentiate between inputs and enabling conditions, i.e., variables that are not considered to have a direct influence on outputs but are necessary for the STI system to work. The framework also includes categories for linkages, which are a special type of inputs emphasized by the National Innovation Systems approach which allows for feedback loops and minimizes linearity in the framework.



Building on the framework, the study identified indicators from fifteen (15) STI-related scoreboards based on the original data sources from which they obtain the data, even though data availability was a major constraint. A total of 263 indicators were retrived and organized based on the framework categories then assessed using the systematic quality assessment criteria of relevance, completeness, and appropriateness. In terms of completeness, there is a generally high percentage of missing data that affects coverage and timeliness of the information, further confirming the challenges identified in the initial review. For more than half (62%) of the countries or years (2010-19), the data is missing, and for inputs, outputs, and linkages, this share goes up to 80%. The high standard deviation suggests that there are large differences in the collection of information between different countries/years.

In terms of appropriateness, i.e., the number of consecutive years for which the data is available, allowing comparison over time results show that on average, countries report data for 37% of the years in the 10 year period, with South Africa and Egypt having a higher balance on the number of years with average data, with South Sudan and Somalia having the lowest. Again this suggests very large variations in the value of indicators across countries and over time.

In terms of relevance, measured through stakeholder consultations, decision making support was a key consideration on how relevant an indicator is. In this case, the Gross Expenditure in Research and Development (GERD) was prioritized by many users not only as one of the relatively easy indicators to use in STI policy processes, but because it also reflects the R&D intensity crucial for securing the intellectual property rights of innovators. The GERD indicator also supports the tracking of R&D funding and performance, allowing for disaggregation.

The resulting suggested scoreboard allows users to navigate the STI framework for each of the 54 African countries, ranked and classified by region and income group as defined by the World Bank. This allows the user to filter these indicators by any of the categories, providing an opportunity for a web-based decision-making tool for countries. In other words, a unique feature of our proposed scoreboard is that the user can filter indicators by their quality (i.e., relevance, appropriateness, and completeness) and monitor how this changes with the other assessment criteria.

The scoreboard also includes a column for Quality Ranking, which provides the mean ranking of each indicator for the criteria. Finally, the scoreboard provides a compilation of available data for all countries in the world, providing a basis for comparative analysis.

Uptake and decision support

The resulting scoreboard provides enabling features for uptake including decision making support and comparability across countries. This presents a paradigm shift in the way most scoreboards have been used, i.e., as a tool to compare the performance of countries. The purpose of this scoreboard however, is not to develop a ranking of countries, given that the data is very variable in terms of quality, and also because an emphasis on rankings promotes competition instead of collaboration. We, instead, try to provide a view towards decision making and collaboration. The scoreboard provides data that is contextually and globally relevant, thus can be used to derive strategies across Africa and beyond. For instance, it can be used to analyze a country like Kenya with respect to potential global partners. More practically, the scoreboard can be harnessed to develop a web-based decision support tool that is user-friendly for different stakeholders, alongside other possible uses such as research analysis, evidence-based convening, policy dialogues, and Communities of Practice, among others.

As a way of moving this forward, discussions around a web-based decision-making tool are already ongoing in collaboration with the AUDA-NEPAD. The AUDA-NEPAD has been hosting the African Science and Technology Innovation Indicators (ASTII). The platform brings together efforts to support countries in monitoring their ST&I progress through R&D surveys. The ASTII has now built into the Afri-look initiative, which converges efforts to build the capacity of national experts to understand and interpret R&D/innovation survey findings, as well sustain the production of internationally comparable R&D and innovation indicators. As such, as part of understanding the wider continental progress on STI measurement and laying the ground for uptake, the team integrated the Afri-look periodic dialogues. On 2nd to 4th March 2021, the project team engaged in a three-day Afri-look training session where we provided insights to the member states on STI measurement efforts based on the study insights. From these engagements, the project is currently engaging with the NEPAD-AUDA to build on the Afri-look efforts and jointly establish a web-based decision-making tool from the scoreboard.

Conclusion and Recommendations

This study provides one of the most comprehensive assessments of scoreboards in Africa, and uses the same to build an integrated, contextually relevant, and comparable scoreboard for the continent. The insights from this study do not only support a solid basis for understanding the quality of data for STI in Africa, but provides a set of relevant indicators conceptually organized around a more flexible framework. The strengths of the scoreboard mainly lie in its ease of use and contextual applicability that could easily inform decision making, collaboration, and learning. This provides a significant milestone in the STI pursuit and an opportunity to link STI to investments decisions. The proposed web-based decision-making tool is a valuable step towards this ambition, and should be developed further.

To be able to use the scoreboard in these ways, it is fundamental to:

- Support the necessary infrastructure, design, and development for making it available through a web platform.
- A cross-cutting need is to provide a repository and resources for updating the data, as weIl as to include new indicators in the platform to achieve better coverage of the STI framework developed in this project.
- For research, countries must increase support to the training of STI data scientists so that the data can feed into relevant research questions, and a community of active researchers can be developed around them.
- Open repositories, forums, hackathons, academic events, are needed to improve the uptake of this data.
- Support research to continue developing and identifying ways to cover all dimensions of the STI framework so that Africa can count on a more complete and relevant source of information.

Overal, we hope that we have contributed a solid basis towards understanding the quality of data for STI in Africa and providing a set of relevant indicators that are conceptually organized around a framework. It is therefore worth noting that the scoreboard is an exploratory collection of indicators and an evaluation of the quality of STI assessment in Africa, and not a mechanism to address the challenges faced by the current assessment processes. Further work will be required to identify ways in which some of the identified challenges can be addressed to ensure a robust ST&I in Africa. We look forward to co-leading further development of the scoreboard and related research on STI policy, together with the Foreign, Commonwealth and Developmnet Office (FCDO) and partner countries and agencies.

Table of Contents

List of Figures

List of Tables

1. Introduction

This report presents a synthesis of the '***Assessment of Science, Technology and Innovation Metrics in Africa'*** study. The study aimed to understand the approaches/methodological frameworks that have been applied in Africa to assess ST&I and to develop a suitable set of standards, key indicators for making a comparable and robust assessment of science, technology, and innovation in Africa.

The study involved an in-depth literature review, secondary data inventory including review of existing ST&I scoreboards (e.g., AIOIII, GII, OECD scoreboard, UNESCO science report, RICYT scoreboard, national scoreboards, among others), and academic literature (reviews, assessments, citing documents) in Scopus, Google Scholar, WoS, and Scielo.

The study was also strongly built on co-production where various stakeholders including experts in the field, Science Granting Councils, and non-academic stakeholders were engaged through workshops and focused group discussions to design, review and validate the study. Specific key informant interviews (from a select number of stakeholders in Kenya, Nigeria, South Africa, Zambia, and Rwanda), and a non-representative survey of STI data users were interviewed, to provide answers to the TOR research questions.

Based on the initial literature review and stakeholder consultation and lessons from the inception phase of the project, the study developed an STI framework that clearly describes Africa's STI ecosystem and indicators-based measurements. A well-defined framework allows for strategic engagement of stakeholders and better assessment of the quality of indicators, and their relationships within the overall STI system.

We then proceeded to identify the STI indicators available for African countries, which covered different components of the framework. We have compiled a set of 263 indicators into an STI scoreboard that is: i) reproducible, as it relies on open access data; ii) relevant, as it draws on the literature on STI indicators, the most influential STI scoreboards, and consultation with Key Informants (KI) and wider STI stakeholder base in Africa, defined in the stakeholder map during the inception phase; and iii) informed by a quality assessment, providing a set of evaluations at the indicator level, allowing users to understand the advantages and disadvantages of each indicator in measuring STI. We assessed STI indicators based on the following established criteria: **relevance**, **appropriateness**, and **completeness**.

In producing the STI scoreboard, we went beyond the request set in the terms of reference (TOR) to develop "metrics that are relevant and comparable in Africa". Our scoreboard includes all African countries, as well as all other countries in the world, and provides time-series data. These two features make our scoreboard relevant for country, regional, and international comparative analysis of STI.

In the next section of this report (section 2), we present the methodological approach. In the third section, we summarise the literature and stakeholder perspectives on STI measurement as well as an assessment of the main scoreboard, with a focus on relevance, appropriateness, and completeness of their indicators. In section four we present our proposed STI framework, which guides the classification of indicators for the scoreboard. Section five explains the detailed construction of the scoreboard. Section six provides the analysis of the scoreboard indicators. Finally, section seven discusses the use and uptake of the scoreboard.

2. Study Methodology

2.1. Methodological Framework

The study methodology is strongly built on co-production, where various stakeholders including experts in the STI field, policymakers (e.g., the Science Granting Councils), and non-academic stakeholders were engaged through workshops and focused group discussions to design, review and validate the study. Stakeholder engagement was embedded in all the steps taken in the study (Figure 1) and as illustrated in the overall methodological framework (Figure 2).

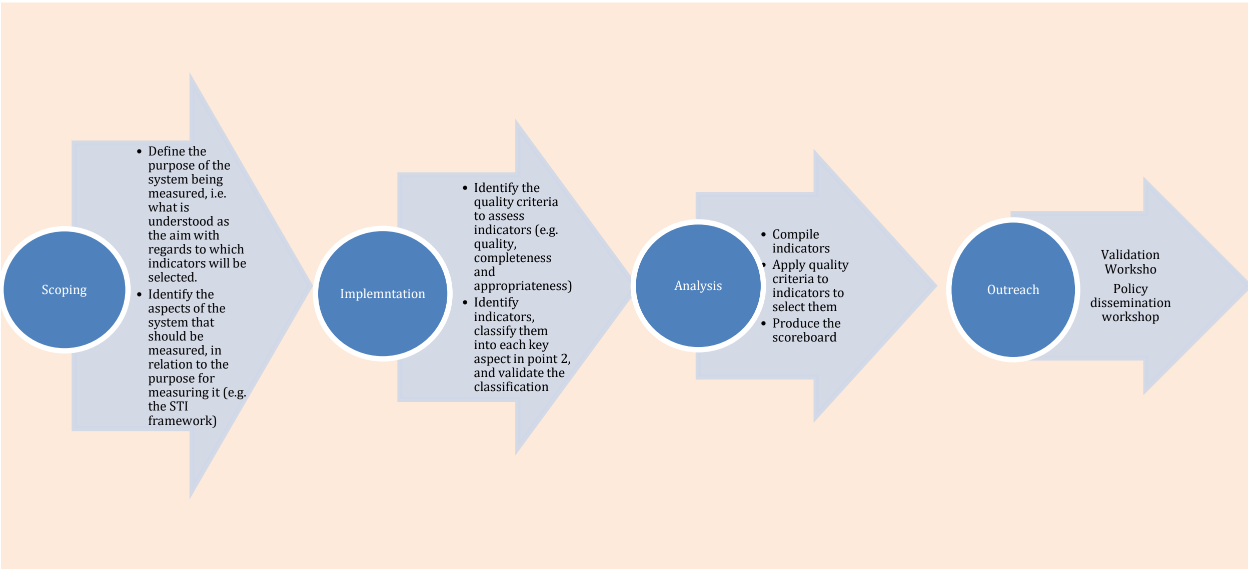


Figure 1: Steps applied in the development of the scoreboard

Figure 2 shows a visual representation of these steps. It is followed by a more detailed explanation of the literature review and stakeholder engagements. The main results from these methods will be detailed in section three, and the remaining sections explain the scorenboard construction process. We briefly discuss the different methods in turn. More detail is provided in the Annexes and the remaining sections.

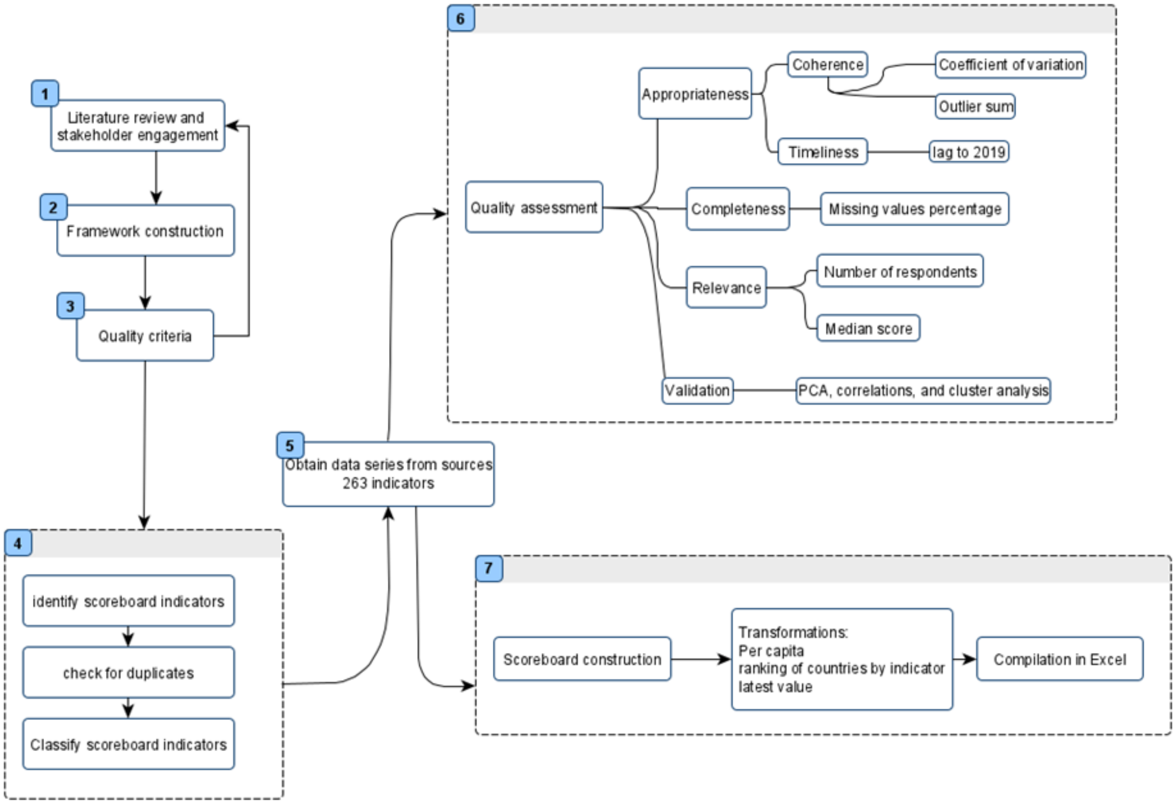


Figure 2: Methodology framework

2.1. Data collection

The main data collection methods employed in this study included: in-depth literature review; secondary data inventory including review of existing ST&I scoreboards (e.g., AIOIII, GII, OECD scoreboard, UNESCO science report, RICYT scoreboard, national scoreboards, among others); and academic literature (reviews, assessments, citing documents) in Scopus, Google Scholar, WoS, and Scielo. Specific key informant interviews (from a select number of stakeholders in Kenya, Nigeria, South Africa, Zambia, and Rwanda), and a non-representative survey of STI data users were interviewed in addition to exploratory surveys.

Table 1: Data collection

TOR Question	Method		
	LR	FG	KI
RQ1.1 What approaches/methodological frameworks have been used to assess the science, technology, and innovation sector in Africa and other parts of the world?	x	x	x
RQ1.2 To what extent do the existing approaches address the quality, completeness, and appropriateness of tracking science, technology, and innovation in Africa?	x	x	x
RQ1.3 What challenges have these initiatives/platforms faced? (data challenges, capacity, etc.). The assessment will review existing indices on Africa such as ASTII among others.	x	x	x
RQ2.1 Who are the main beneficiaries or users of the indices?	x	x	x
RQ 2.2 How have these monitoring initiatives impacted the development of science, technology, and innovation ecosystems in Africa? How have they been used to inform or influence policy or practice? To what extent does the measurement of science, technology, and innovation link with development goals?		x	x
RQ 2.3 What lessons can be learned to improve and strengthen these initiatives and the monitoring of science, technology, and innovation in Africa overall?	x	x	x
RQ3.1 What is the most suitable framework and design of a scoreboard/dashboard to effectively track science, technology, and innovation metrics, ensuring timeliness and accessibility to a wide range of users (easy to navigate, user-friendly)?	x	x	x
RQ3.2 What is the most suitable set of standards, key indicators for making a comparable and robust assessment of science, technology, and innovation in Africa (including tracking the impact of science, technology, and innovation)?	x	x	x
RQ4.1 To what extent does the measurement of science, technology, and innovation link with development goals?	x	x	x
RQ5.1 Which indicators has the UK contributed to significantly and where is there potential for UK investments to drive impact? [Engagement/interviews]	x	x	x

2.1.1. Literature Review Rationale and Procedure

In this section, we briefly discuss the literature review approach. The detailed literature review procedure is discussed in Annex 1. The full list of references is provided in annex 2. We followed Hagen-Zanker and Mallet (2016) approach to perform systematic literature reviews: setting the focus of the literature search, selecting the sources and databases, defining search strings, specifying inclusion criteria, screening the documents, and classifying evidence for analysis.

The focus of our literature search was to gather relevant information to define a suitable STI framework and identify relevant and robust indicators for STI tracking in Africa and Globally. To build the STI framework, we focused the review on two information sources:

i. We reviewed fifteen (15) STI-related scoreboards: this includes the well-known scoreboards listed in our original proposal, plus scoreboards that were identified during the inception phase via experts' suggestions, references, and earlier related reviews by Hollanders and Janz (2013) and Pouris (2016). The additional scoreboards identified during the inception phase allowed us to expand the original list especially in relation to integrating additional data, e.g., on agriculture, and societal outcomes (e.g., gender and human development), and to enhance the relevance of the identified indicators to Low-to Middle Income Countries (LMIC) contexts. Reviewing all these STI scoreboards provides an understanding of the different perspectives used by STI-related organizations across a variety of countries in measuring STI (and related aspects).

ii. We reviewed the academic literature on STI systems, STI measurement, and STI indicators. To reduce the well-known biases in the coverage of the literature from the global South and in languages different from English (Ràfols et al, 2019), we searched a diversity of repositories: Scopus, Google Scholar, Scielo South Africa, and Open Doar . This approach provided us with an academic perspective in addressing the outlined research questions, STI measurement, conceptual frameworks, and challenges in the measurement of STI and use of indicators.

2.2.2. Stakeholder engagement/Focused Group Discussions (FDGs)

Stakeholder engagement aimed to critically support the review and selection of indicators, critically discuss intermediate outputs, and lay ground to design the strategy for uptake and application of an STI scoreboard. We employed five engagement methods outlined below.

i. Stakeholder engagement prioritization

To prioritize the stakeholders to engage, stakeholder mapping and re-prioritization were conducted at various stages during the project lifetime. We focused on sampling a wide variety of users to broaden the variety of views and perspectives on the STI indicators and the underlying frameworks. The stakeholders sampled included: a wide research community (about 70 researchers from 21 African countries under the Africa Research and Impact Network (ARIN)) working in different fields and disciplines; policymakers from different sectors, beyond STI; and several African networks[1] and institutions in STI.

ii. Surveys

We conducted two exploratory, non-representative surveys: a pilot survey to gather views on key data priorities (involving 30 stakeholders who also participated in the inception workshop); and a pilot survey to appraise the relevance of the STI indicators that are available across most African countries (about 50 respondents, mainly from government agencies mandated with STI, researchers, and development

[1] UN economic Commision of Africa, Council for the Development of Social Science Research in Africa (CODESRIA), Organisation for Social Science Research in Eastern and Southern Africa (OSSREA), Society for the Advancement of Sciences in Africa (SASA), Human Sciences Research Council (HSRC), National Research Foundation (NRF), various African chambers of commerce, Network of African Sciences Academises (NASAC), African Academy of Sciences (AAS) and Non-Governmental Organisations

agencies). Respondents were asked to select the five most relevant indicators in each of the STI framework components. The definition of relevance was based on previous stakeholder consultations.

iii. Focus group discussions

We organized five (5) Focus Group Discussions (FGD) in the priority countries to examine the monitoring of STI initiatives and their impact on the development of the STI ecosystem, their influence on policy processes, and the relevant criteria to assess STI indicators. Each focus group was composed of six members, who had been sampled from the following criteria: one country representative in the African innovation outlook; ministers/government officials in charge of Higher Education, Science and Technology; two representatives from academia, technical institutes, universities, or research institutes; one representation of the national statistics office; and non-state actors/representation from the technology and innovation hubs selected purposively in each country.

iv. Dialogue sessions

The research dialogues were convened to discuss the following: STI measurement, monitoring, and financing to promote resilience to pandemics such as COVID-19, as well as addressing wider sustainability agendas such as the SDGs (dialogue a), validate project's findings at various stages (dialogue b), and to discuss strategies for uptake and impact pathways for the scoreboard (dialogue c). These were targeted to involve various research networks, development agencies, international organization mandated with the development of indicators, as well as policymakers within the Science Granting Council Initiative (SGCI). We outline the dialogues below:

a. We organized a webinar in collaboration with the ARIN on the 25th of September 2020. The webinar focused on harnessing STI for managing COVID-19, incorporating STI monitoring, and evaluation needs to serve both strategic and development ends. The webinar was attended by over 60 participants drawn from diverse African stakeholders who include think tanks, researchers, policymakers, innovators, incubators, the productive sector both the formal and informal sector players, and STI experts. These stakeholders discussed opportunities for building STI indicators now and post COVID-19.
b. A webinar organized on the 29th of October 2020 discussed the project's early findings, validated an STI framework developed to guide the selection of indicators, discussed the assessment criteria of indicators, and prioritized relevant STI indicators. About 45 stakeholders with backgrounds in STI policy and decision making discussed their understanding of relevance, validity, and coherence, which are some of the assessment criteria for indicators (see full report **Annex 3).** The webinar provided an avenue for constructive dialogue and discussion, including future/areas of development in supporting relevant African STI indicators. Following a consensus on the definition

of relevance, stakeholders selected the indicators individually and discussed the reasons behibd their selections.

c. Another dialogue session held on the 20th of November 2020 discussed how to support the various continental initiatives that have supported capacity-building initiatives, including impact and uptake pathways of STI indicators for decision making. This dialogue was conducted through a panel and explored: evidence procurement for decision making and its impact on the development of STI ecosystems in Africa and individual countries; and impact pathways/channels that can support uptake and use of STI indicators collected and developed in the scoreboard. The dialogue was attended by about 30 stakeholders drawn from the continental STI agencies such as AUDA-NEPAD, international agencies like UNESCO, research institutions, innovation agencies/hubs, SGCs as well as ARIN fellows.

2.1.3. Key informant interviews

We aimed at discussing strategic perspectives on STI measurements across African countries. We interviewed 29 key informants from the following sectors: high-level officials from the national agency responsible for STI policymaking in the priority countries (n=5); representatives of the national statistics office (n=5); high-level officials of a non-governmental organization with an STI related mandate (n=5); previous persons involved in STI or related commissions particularly ASTII country focal points (n=5); non-profit organizations with previous mandates on the development of STI indicators addressing gender issues in Africa (n=2); and officials from the UK government departments/development agencies (n=7).

2.1.4. Integration with existing assessment platforms – Afri-Look

Building on the assessment of the existing initiative, the African Science and Technology Innovation Indicators (ASTII) initiative supported by AUDA-NEPAD was identified as a key platform bringing efforts together and supporting countries to monitor their ST&I progress through R&D surveys. The ASTII has now built into the Afri-look initiative which converges efforts to build the capacity of national experts to understand and interpret R&D/innovation survey findings, as well sustain the production of internationally comparable R&D and innovation indicators. As such, as part of understanding the wider continental progress on STI measurement and laying the ground for uptake, the team integrated with the Afri-look periodic dialogues. On 2nd to 4th March 2021, the project team engaged in a 3-day Afri-look training session, where we provided insights to the member states on STI measurement efforts based on the study insights. From these engagements, the project is currently engaging with the NEPAD-AUDA to build on the Afri-look efforts and jointly establish a web-based decision-making tool from the scoreboard.

3. What the literature and KI perspectives tell us about STI scoreboards, indicators, and their assessment.

This section report results from part 1 of our methodology. We summarize key insights from the literature review and stakeholder engagements in relation to the existing STI scoreboard and frameworks, and in relation to the research questions asked by the TOR. A more detailed literature review was presented in our earlier report.

3.1. Overview of the STI Landscape in Africa

Understanding Africa's general STI context is useful to guide the process of indicator assessment and the policy perspectives that frame these indicators. Currently, the STI landscape in Africa is driven by country-specific policies anchored on the regional policy agenda, the African Union's STI *Strategy for Africa 2024* STISA-2024). The STISA 2024 is a pre-requisite for achieving *Agenda 2063* 'The Africa We Want', and aims to guide member states in their efforts to strengthen their STI systems with new policies, institutional arrangements, and investments (Frost *et al.*, 2019).

For most African countries, the STI policy is anchored on the ambition to speed up the transition to middle and developed economies through an innovation-led, knowledge-based economy. Many African countries have either established their STI policies/strategies to guide investments, monitoring, and reporting. Most countries have established the National SGCs, as the key agencies coordinating national STI-related activities within countries.

Within these SGCs, countries are pushing for science and innovation systems with a strong orientation towards science systems and innovation that facilitate mainly knowledge production but, still relatively weak on knowledge use. For instance, the KSI study revealed that East African countries have invested in Knowledge production driven by both states, especially public Universities and non-state actors such as Think Tanks and Private Firms from both formal and informal sectors. Countries have also established systems to promote innovation and research use aided by National Research and Innovation Funds (e.g., in Kenya and Rwanda) supported by government budgets and external funders.

In terms of progress monitoring, most countries undertake periodic R&D surveys that are carried out under the auspices of the African Science Technology and Innovation Indicators (ASTII), operating under the auspices of the AUDA-NEPAD. The number and frequency of these surveys however vary from one country to another (Figure 3).

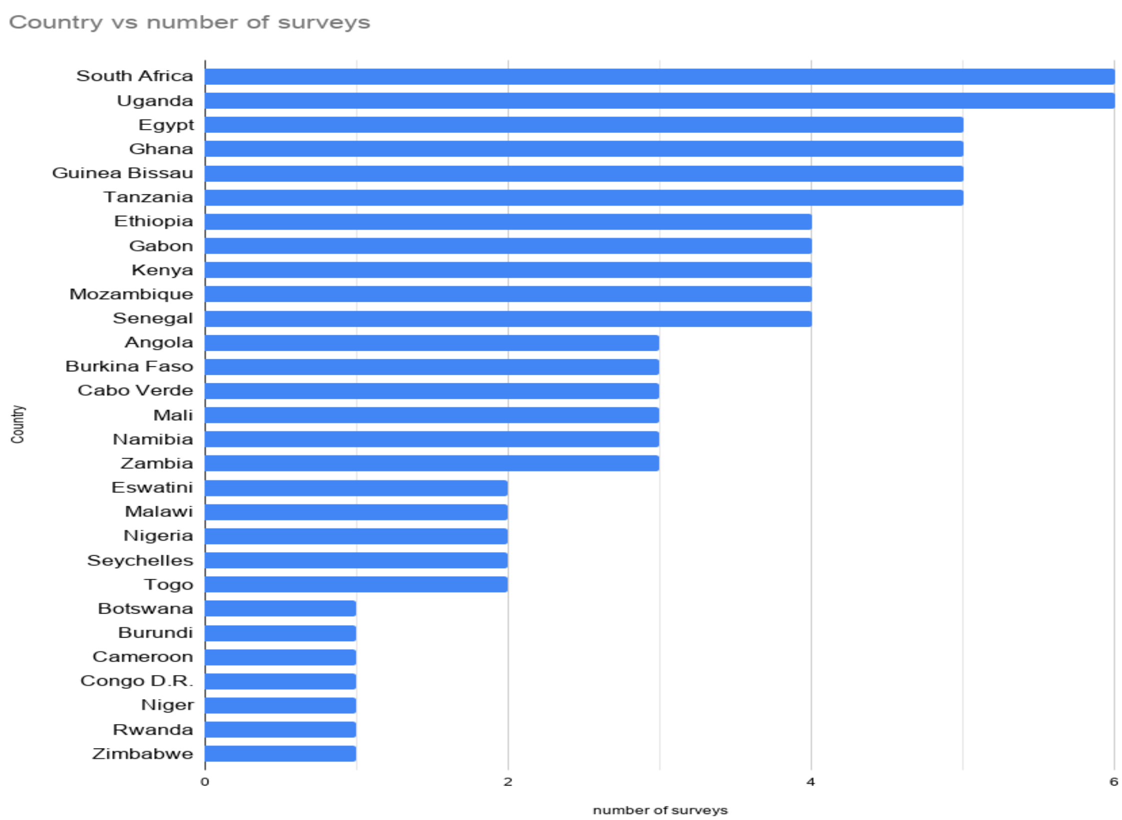


Figure 3: Number of R&D surveys undertaken by various African countries

Despite the growing STI ambition among African Countries, within the policy realm, most African countries still struggle with many many policy and practical challenges mainly owing to the weak investment in ST&I. Most African countries have historically lagged in STI investment. For the last decade, Africa has contributed only 0.6% of the world's gross expenditure on research and development (GERD), as compared to the figures for Asia and Europe, 30.5% and 27.2% respectively, according to the 2020 R&D Global Funding Forecast.

Additionally, most STI action and planning is still largely at the national level, with countries still grappling with what effective STI systems look like in practice, and what can work for SSA.[2] This is also exacerbated by concerns around the mismatch of STI theories developed elsewhere, and the practical African context and realities. This ultimately has led to little clarity on the practical engagements with various systems, e.g., non-state actors, informal systems, and strategic interventions relevant to the STI.

[2] https://www.nri.org/?option=com_fileman&view=file&routed=1&name=KSI_%20Blog%20Post%20final%202.pdf&folder=Development%20programmes&container=fileman-files

A recent FCDO supported study on Knowledge Systems in East Africa (Frost et al, 2019; Frost et al.,2020) highlighted multiple challenges to Africa's STI systems which include: i) lack of practical means and evidence on what works to promote STI in practice; ii) institutional weaknesses that limit the creation and use of knowledge; and iii) weak capacity to interpret, design and monitor STI interventions. These challenges are largely exacerbated by the lack of adequate indicators to assess investments and associated impacts.

Overall, Africa's STI landscape is characterized by mixed progress differentiated across countries. A fundamental milestone is the establishment of national-level policies and agencies. However, most countries still face challenges in translating plans into action, coupled with little evidence on what is going on in practice. Several ongoing innovative activities/initiatives within countries are poorly documented, measured, or reported, thus lack of clarity on how interventions contribute to countries' development and overall SDGs. The need to assess ongoing interventions as a way of prioritizing new ones is central to Africa's ST&I progress.

3.2. Existing/ongoing initiatives that attempt to measure STI in Africa

Understanding the ongoing initiatives on STI measurements provides answers to RQ 2.2, i.e., *How have these monitoring initiatives impacted the development of science, technology, and innovation ecosystems in Africa? How have they been used to inform or influence policy or practice?*. In this, we first sought to understand ongoing/existing initiatives that attempt to measure STI in Africa. We then analyzed how these initiatives have influenced policy and practice.

At the continental level, stakeholders including experts and organizations with long histories working in the ST&I field identified many initiatives including initiatives by the United Nations Economic Commission for Africa (ECA), R&D and Innovation surveys conducted under the African Science and Technology Innovation Indicators (ASTII) initiative supported by AUDA-NEPAD, among others. Additionally, there are initiatives such as the African Observatory of Science, Technology, and Innovation (AOSTI) which is the continental repository for STI statistics and a source of policy analysis in support of evidence-based policymaking in Africa. Currently, there are efforts to put in place a new initiative, Afri-look, which is a follow-up of the ASTII project that ended in 2019. This initiative builds on the 13-year successful legacy of the ASTII project, which has been building the capacity of national experts to understand and interpret R&D/innovation survey findings as well sustain the production of internationally comparable R&D and innovation indicators. The AU-NEPAD has also supported the SGCI phase I in monitoring and evaluating their STI policies.

The quick scan survey with the fifteen (15) SGCs from sub-Saharan Africa (mainly associated with the SGCI initiative), revealed that many countries seem to be at very early stages of developing mechanisms for measuring STI. At the moment, STI measurements are limited to setting benchmarks aimed at certifying the quality of courses and research priorities. Most of these benchmarks are also largely qualitative and general, involving provisions such as aligning research to the development goals of a country, among others.

The countries still lack clear guidelines and indicators on measuring STI interventions. Nonetheless, various countries are now receiving technical support towards understanding the measurement of certain (not holistic STI spectrum) elements of STI. For instance, the Kenya National Commission (KNATCOM) for UNESCO is supporting the measurement of STEM and Gender Advancement (SAGA) by collecting gender-segregated data around gender and Science, Technology, Engineering and Math (STEM). In Nigeria, initiatives such as Tertiary Education Trust Fund (TETFUND), an agency supporting tertiary education, are setting up programs to support STI measurement. The East Africa Science and Technology Commission (EASTECO) is also supporting regional member states to understand ways in which STI can be measured. These efforts are nonetheless at preliminary levels, targeting more monitoring and evaluation rather than specific indicator-based measurements.

Monitoring initiatives are also developed at the sectoral level with two priority sectors (agriculture and health) in most African countries. In the agricultural sector, the development of the Agriculture Science Technology and Innovations (ASTI) indicators have supported the evaluation of the performance of agricultural research systems by assessing how research investments relate to the quantity and quality of research outputs. These have supported more investments, targeting gender-specific types of investments both at the policy level and at an institutional level. As a result, they have supported the development of national/country data and indicators in agriculture. Such indicators (used at the institutional level and/or country level) support trend analysis of countries and thus influence decisions in modeling their R&D. The ASTII indicators have further widened the scope and as such, many agricultural organizations in Africa embrace the idea of gender equity while working with national agricultural institutions to collect data. In the health sector, indicators have supported research capacity development, including capacity building for outputs such as publications and the development of infrastructure for capacity development.

3.3. Existing approaches and frameworks to measuring STI

Several scholarships have referred to STI in relation to societies and development in a variety of explicit ways. These include the literature on: i) the sociology of science (Merton 1943 [1973]); ii) economic growth (e.g. Guellec and van Pottelsberghe de la Potterie, 2004; Bloom et al. 2020); iii) the Input-Output framework (OECD 1963; OECD 1992); iv) the National System of Innovation (Lundvall 1992; Freeman 1995); v) the Mode 2 knowledge production (Gibbons et al. 1994); vi) social impact and social value (Martin 2007; Molas-Gallart et al. 2002); vii) open science (Wouters et al. 2019, Bartling and Friesike, 2014); viii) the quadruple helix (Hasche 2019); ix) responsible research and innovation (Von Schomberg 2013); and x) novel frameworks for research evaluation (Ofir et al. 2016).

All this literature presents an encompassing view that conceives science as part of a broader system that includes technology and innovation and is deeply connected with other aspects of society. Two categories of literature/frameworks have been extensively used to develop scoreboards and rank STI outlooks: the input-output framework and the national system of innovation. These two frameworks inform most of the scoreboards that we reviewed, and which are described in the Oslo and Frascati manuals and other scoreboards that were developed to measure STI in African countries. For instance, the African Innovation Outlook III (AIOIII), relies on the Oslo manual, and the Consultative Group for International Agricultural Research (CGIAR) relies on the Frascati manual. From our systematic review of existing scoreboards and literature, we conclude that existing approaches to STI measurement in Africa are based on the input-output and the national innovation systems (NIS) frameworks (this is discussed in greater detail in Section 4). We refer to this as the standard STI framework.

3.4. Existing approaches to assessing STI: relevance, completeness, and appropriateness

From the literature review, we found that the literature has not invested substantially in assessing the existing scoreboard. The existing scoreboards that we reviewed mainly discuss features of quality, completeness, or appropriateness of the STI indicators that they use. This applies to the Global scoreboards as well as to those focused on African countries, or other LMIC.

Beyond the scoreboards, Molas-Gallart et al. (2002) assesses third mission indicators, and Wouters et al. (2019) assesses open science indicators. However, they do not explain in any detail the methodology used to evaluate the indicators, i.e., who participated in the evaluation of the indicators and what rubriques were used. Wickson and Carew (2014) used stakeholder engagement through organizing World-Cafe discussion where they discussed quality criteria, and how to measure them.

A handful of studies assess the STI scoreboards, such as the European Innovation Scoreboard (EIS) and the Global Innovation Index (GII). Adam (2014) offers a thorough assessment of the quality of the EIS. He finds many issues in relation to, among others: lack of a clear framework, arbitrary selection of indicators, the dependence of the scores on the availability of indicators, the large variability of the indicators, abrupt changes in the values of indicators between editions of the scoreboard, and comparability. His findings expand on earlier studies (Schibany & Streicher, 2008; Grupp & Schubert, 2010; Archibugi, Denni & Filippetti, 2009).

We could not find studies assessing the scoreboards we reviewed in relation to the relevance, completeness, and appropriateness of the indicators that they use. This is true for scoreboards measuring African economies, or any other. However, most scoreboards are populated by a large number of footnotes that address specific issues with the indicators. Table 2 compiles some of the issues annotated through careful and systematic inspection of footnotes and annexes of the scoreboards. These findings should be considered as an exploratory exercise because they are based on details in footnotes, which may be difficult to spot in some of the reports. They do confirm the need for a better and more systematic analysis of the quality of the data in the scoreboard, for instance in relation to their relevance, completeness, and appropriateness.

*Table 2: Issues discussed in some of the scoreboards: frequency**

Issue	Number of Mentions	Percentage
Missing data	27	42%
Methodology change	22	43%
Different rears of reference	9	14%
Comparibility issues	5	8%
outliers	1	2%

* Missing data refers to empty values; methodology change refers to notes that warn about different collection strategies either for countries or for different years in the same country; Different years of reference means that the data collection time for one or more indicators is not the same across countries; comparability issues cover notes that do not identify explicitly the cause for these issues but suggests not to use in comparisons. Outliers refer to notes warning about the presence of extreme values that may affect a ranking. The percentage column is calculated for each row over the sum of all mentions.

Source: own elaboration based on the footnotes of African Regional Integration Index, African Innovation Outlook III, European Innovation Scoreboard, Global Innovation Index, Innovation Output Indicator, OECD Main science indicators, South Africa Science, Technology, and Innovation Indicators, UK Competitiveness Index, UNESCO World Science Indicators.

Based on our literature review of the scoreboard's evaluation of STI, and evaluation of indicators, we conclude that relevance, appropriateness, and completeness of STI indicators and scoreboards have not been adequately addressed in the literature and the existing scoreboards. This lack of attention to data quality in the literature severely undermines the usability of the data and scoreboard for comparative analysis (Jerven, 2013). We suggest that it is crucial to assess the data gathered in a scoreboard, and weight analysis against potential differences in quality across variables and countries.

Stakeholder perspectives

Gathering detailed views from stakeholders about these scoreboards was a bit of a challenge due to their (stakeholders) lack of adequate information about the scoreboards, and inability to clearly understand all the three (3) assessment criteria (relevance, completeness, and appropriateness). It was particularly difficult to get stakeholders' views on completeness and appropriateness. They were more interested in discussing the relevance of indicators. Because it is not possible to define relevance from the data analysis, we opted to assess relevance through stakeholder engagement, and completeness and appropriateness through data analysis (Sections 5 and 6).

We first aimed at defining STI indicators relevant to users. The assessment of relevance was mainly subject to the country's needs, investment priorities, and users, which often varied within between different countries. Stakeholders' definitions of relevance were closely influenced by: policy, legislative and institutional environment in supporting STI; the contribution of STI to the socio-economic development of countries; and reporting frameworks in the STI observatories and national statistical systems. Some of the stakeholders' perceptions of relevance are presented in Box 1.

Box 1:Stakeholder's perspectives/understanding of the concepts of relevance.

As emphasized previously and in literature, the understanding and interpretation of relevance of STI indicators largely depends on the user and the context the scoreboard/ indicators are used. The below broad definitions were provided.

Users emphasized that a relevant STI indicator measures and tracks impacts of STI.

A relevant indicator will also meet the needs of the users of the indicators and the scoreboard. These users need to be defined and go beyond of policy makers but more broadly users in different spheres.

A relevant indicator will help prioritize how societal needs are addressed in different countries or subnational contexts. However, contexts evolve, are dynamic and always keep changing. As such both political contexts and policy contexts should be discussed when addressing the issue of the relevance of the indicators. Political contexts could imply how various governments allocate their national priorities at different times.

A relevant indicator contributes to priorities how the SDGS are addressed in different countries.
Gender considerations are necessary and are outlined as part of relevance of the indicators. Most of the indicators collected will provide elements of the participation of various gender categories or disaggregate the information with regards to gender, however this may not be enough. A relevant indicator may need to provide insights into cultural barriers associated with inequality, the determinants of various genders participation and progress in various research skills as part of the broader considerations of complete, appropriate, and relevant indicators.

Most of the scoreboards are reflecting on the above definitions differently. For instance, scoreboards such as the SAGA by UNESCO are already providing gender aspects and piloting some of the gender considerations as key aspects of a relevant indicator. We reckon that these are useful ideas for various scoreboards to consider. While analysis of the relevance of indicators was considered as defined, the scoreboard produced by this research did not include these gender aspects, but proposes that individual countries support/choose the indicators and customize them to fit their gender considerations.

While the relevance of the indicators is key for STI measurement in Africa, stakeholders also suggested considering the legitimacy of the indicators in future work. Legitimacy was defined in relation to who has the mandate to collect data on what indicator within a national system. This also suggests the design indicators within the country/continent at the outset, in a way that they can be integrated into international standards.

3.5. Adoption of existing frameworks

The findings presented here respond to RQ 2.1, i.e., *who are the main beneficiaries or users of the indices?* From the ***literature review***, we found that the main audience is policymakers, at different regional, national, and international levels. Users targeted by scoreboard reports are also researchers, who mainly use the scoreboards for analysis. Other users include national statistical offices as well as education institutions (including teachers and students) where they are used for academic projects.

To have a better understanding of who uses these STI scoreboards, we searched for the titles of the scoreboards in Google Scholar to have a picture of who cites them in the academic and grey literature. An overview of the citation map shows that most of the citing instruments use scoreboards as data sources for quantitative analyses. Additionally, we searched the database to identify the countries and organizations that use these scoreboards. We found that universities are the most active users of scoreboards, although other actors are making use of the data. For instance, in the case of the AIO, we found organizations such as the African Population and Health Center, the Asian Development Bank, the Human Science Research Council of South Africa, and the Council for Scientific and Industrial Research in Ghana (see annex 4 for a more complete analysis).

Specifically, for Africa, we found four papers that analyze the national systems of Kenya, Ghana, and South Africa using national R&D and innovation surveys specifically designed for these countries, and available scoreboards (Bertels, Koria, & Adriano 2016; Abrahms & Pogue 2012; Bartels & Koria 2014; Kahn 2006). These papers confirm that indicators produced on the STI systems in Africa are being used to produce knowledge in academic documents even though the extent to which this happens remains unclear.

Discussion ***with stakeholders*** confirms that policymakers (governments and governmental agencies, public sector agencies, and technology centers are the main users of the frameworks. However, most stakeholders felt that the intensity of usage by most policy bodies is relatively ad-hoc based need at a particular time rather than planning support for countries. The ad-hoc use is also fueled by ad-hoc data collection because the current R&D surveys performed by countries are yet to be fully institutionalized.

We observed that the National Science Commissions for science and technology and the SGCs in the various countries are keen to continue supporting the harmonization of the institutions collecting and using STI data. Other users of STI scoreboards are the boards of management of productive enterprises heavily reliant on R&D, parliamentary groups concerned with STI, and the SGCs who use the indicators to procure evidence for decision making in terms of funding needs.

Moving forward, stakeholders suggested that other non-policy users, e.g., the private sector, Technical and Vocational Education and Training (TVETS) should be made more familiar with STI scoreboards through training and continuous dialogues. While Universities and Government Training Institutes could support the training of the formal sector on the STI frameworks and their utility, TVET institutions and incubation hubs could support training in the informal sector.

3.6. Challenges and limitations of existing frameworks

From the ***existing literature**,* clear challenges emerged from the discussion in the literature on the need to assess the completeness, relevance, and appropriateness of indicators and scoreboards. Beyond assessment, the literature points out further challenges to be taken into account in the construction, communication, and use of STI indicators and scoreboards. We first discuss general challenges and then move to challenges specific to the global South and African countries (Box 2).

Box 2: Challenges identified from the literature.

- Need to stay up to date with new conceptual developments, and advances in the frameworks that define STI systems (Gault 2013a; Bordt 2007; Gault 2018; Ciarli 2018).
- Challenges in interpreting the meaning of indicators (Gault 2013a; 2013b).
- Obtaining comparable micro-data, across different countries (Bartels 2014).
- Developing standards for the collection and classification of data in universities and NGOs, among others (Gault 2013a).

Specific challenges for African countries include:

- Need to enhance capabilities in STI indicators and policies (Iuzuka, Mawoko, & Gault 2015; Kruss 2018).
- Increasing data gathering and availability (Klingebiel & Stadler 2015).
- Incorporating sustainability and STISA priorities in the scoreboards to increase their relevance (Kruss 2018).
- Obtaining internationally comparable indicators that are also relevant for developing countries that are diverse in terms of their systems of innovation (UNESCO 2010).
- Incorporating alternative forms of knowledge creation such as
    - Indigenous and informal knowledge (Gault 2018; Kruss 2018).
    - Incremental innovations (Janger et al. 2017).
    - Open science (Wouters et al. 2019).
    - Inclusive structural change (Ciarli et al. 2020).
- Fostering the demand for indicators from policy makers (UNESCO 2010)
- Differentiating between private and public universities (UNESCO 2010)

While literature tended to highlight conceptual challenges more generally, ***stakeholders*** were more inclined towards contextual challenges, perhaps informed by their practical experiences. Stakeholders particularly reiterated the need to develop contextually relevant indicators that are co-produced with the users of these indicators based on a well-understood framework, but this will require strategic technical support to be offered to these stakeholders. Such frameworks should go beyond the traditional input-output indicators' approach and allow for the generation of evidence for policymaking from a much broader perspective. Already new developments like the GOSPIN-methodology developed by UNESCO are including new methodological frameworks that account for "national and international contextual factors, geopolitics, financial resources, fluxes of goods, information" etc. This methodology expands Science Engineering Technology and Innovation assessment not only to consider country-specific contexts, but also emerging knowledge of technological advances that contribute to sustainable development.

While it is time to develop and collect the indicators, ***stakeholders*** cautioned the need to assess capabilities across African countries and capacities required to support the uptake of the STI indicators and use. This is because many countries have weak institutional structures for statistical measurement, including STI measurement. This leads to low comparability even if similar methodologies are applied across the continent**.** Other barriers include the poor understanding of what constitutes the informal sector.

Prioritizing Africa's informal sector and its contribution to development needs urgent attention. It was emphasized that indicators need to go beyond narrow measurements (e.g., Gross Expenditure on Research and Development) and focus on the informal sector, for example, in conducting surveys. The focus on the formal sector is also detrimental to outcomes such as inequality. Social innovation indicators (Box 3) also need to be included as these innovations have the potential to reduce inequalities.

Box 3: Social innovation indicators

- Develop more contextual indicators.
- Inclusive indicators set covering both formal and informal sectors
- Assess the capacities among the producers and users to collect and use the indicators developed.
- Low funding and budgetary allocations for STI initiatives
- Lack of support and involvement of most of the private sectors players in the development of STI indicators.

4. Proposed integrated framework for the development of an African STI scoreboard

This section corresponds to step 2 of our methodology. Building on the above analysis of the literature, complemented by stakeholders' views, we define an STI framework to select and assess STI indicators that can be used to effectively track science, technology, and innovation metrics, ensuring timeliness and accessibility to a wide range of users (RQ3.1).

The selection of indicators that compose the scoreboards is based on the explicit or implicit STI framework adopted (Godin, 2007). Frameworks can be defined as analytical models of an STI system through its components and relationships (Godin, 2015).

Our review of STI frameworks has shown that there is no unique view about the components that are considered to be part of systems of innovation, nor the indicator against which such components should be measured. From the review of scoreboards and of the academic literature, we have identified the following three dimensions against which we define an STI framework to build a scoreboard:

1) **Domains**: we distinguish three main domains of indicators:
    a. science and technology.
    b. innovation,
    c. social, economic, and environmental conditions, and demands.
2) **NSI components**: We follow the categories of actors, linkages, and activities used in the NSI literature (see section 3.2).
3) **Logical framework categories**: We distinguish among the categories used in the Input-Output framework and NSI literature (see section 3.2): enablers, inputs, linkages, outputs, and impacts.

We explain these three dimensions in more detail below.

4.1. Domains

Science and technology*:* For the scoreboard, we refer to science and technology as an umbrella term to identify formal activities aimed at producing new knowledge and its application for practical purposes. We understand that this definition does not fit exactly with the Frascati Manual (OECD 2015), which is focused only on R&D. Our definition combines different inputs from the literature and the different scoreboards.

Innovation: For the scoreboard, we refer to innovation as “a new or improved product or process (or a combination thereof) that differs significantly from the unit’s previous products or processes, and that has been made available to potential users (product) or brought into use by the unit (process)” (OECD 2018).

Within this domain, we classify activities, resources, and relationships considered by the literature as directly related to the production of such innovations.

Social, economic, and environmental conditions and demands: These are the contextual factors that facilitate or hinder knowledge flows, access to innovation opportunities, the performance of STI activities, or constitute domains of expected impacts of the STI system. These conditions include state of the financial system, governance and institutional development, human and sustainable development, ICT readiness, infrastructure, entrepreneurship environment, job and product market, regional integration, social mobilization, trade, consumption, and entrepreneurship.

4.2. NSI Components

Actors*:* An STI system is composed of several actors that are considered relevant to produce, use, and exchange STI. These actors contribute to generating STI in various forms. These actors include: brokers and suppliers of technology and knowledge; education and training organizations such as teaching universities or technical education organizations; financial organizations and venture capitalists; firms in the formal sector; innovating actors in other sectors of the economy, e.g., organisations and individuals such as the self-employed or farmers, including the informal sector, non-profit institutions, R&D organizations (mainly research universities, laboratories, and their associated R&D personnel), and state institutions, which include a wide range of public organizations.

Activities: The literature considers many activities that the above actors undertake and which contribute to generating STI in various forms. Activities include education and training; expenditure on R&D; expenditure on science and technology; innovation activities, innovation outputs, knowledge flows such as scientific international collaboration, technology transfer, among others; research activities represented in time devoted to research by R&D personnel; science and technology outputs such as radical innovations, basic science discoveries, patents, and research papers, creative outputs such as trademarks, films, and art.

4.3. Logical framework categories

The logical framework that we propose allows us to analyze the relation between STI inputs and outputs, as explained by Godin (2007), adding categories to identify societal outcomes (Figure 4). We follow scoreboards that differentiate between inputs and enabling conditions, which are variables that are not considered to have a direct influence on outputs but are necessary for the STI system to work. Additionally, we include categories for linkages, which are a special type of inputs emphasized by the National Innovation Systems approach (see 3.2). Finally, we include impacts, which consider societal outcomes (influenced by innovation outputs), which are related to enabling conditions in the following periods, and which are considered as a separate category in some of the scoreboards such as the EIS.

To escape the non-realistic sense of linearity that is implied by an input-output framework, we acknowledge that the different categories of the logical framework have feedback loops, which mean that any category can influence another. For instance, enabling conditions can affect linkages, and linkages can affect enabling conditions, etc. For the sake of visualization, we only show the feedback loop between contiguous categories and between impacts and enabling conditions, to show that the improvements in the environment, society, and economy can become enablers for STI. Figure 4 illustrates our logical framework categories.

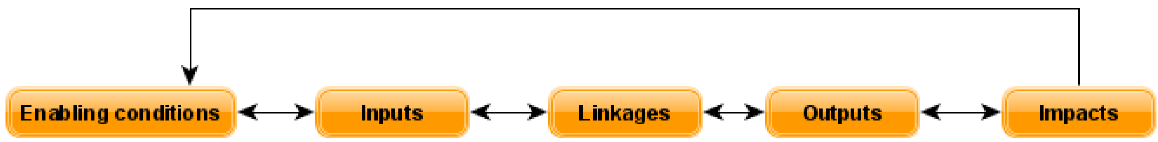


Figure 4:STI Logical framework categories. Source: Own elaboration based on scoreboards, literature review, and interaction with stakeholders

Figure 5 plots a visual representation of the framework, combining the NSI components and the logical framework. This enables us to explain how they are related. The rows of the figure show the logical framework categories, and the nodes and arrows across the NSI components. For each component, Figure 5 plots, the percentage of STI indicators identified in the 15 scoreboards that pertain to that component or logical framework over the total number of indicators (1,113). The bigger circles surrounded by a green ribbon show the percentage of indicators for social, environmental conditions and demands, actors, and activities.

The small circles show a disaggregation of the percentages for specific NIS components. The classification of indicators into actors and activities overlap, as every indicator was assigned to an actor and activity. For this reason, the sum of social, economic, and environmental conditions and actors is 100%, and the sum of social, economic, and environmental conditions and activities is also 100%.

The percentages show where indicators from the available scoreboards are concentrated. 42% of the indicators refer to social, economic, and environmental conditions, which are points of departure and arrival (enablers and impact variables). In terms of actors, the indicators are concentrated in innovative firms in the formal economy, and R&D organizations of personnel. Other important actors such as innovators in other sectors of the economy, financial organizations, education and training organizations, non-profit institutions serving households, and brokers and suppliers have very few indicators or do not have any. From an actor's perspective, most indicators concentrate on expenditure and activities performed by R&D personnel. Fewer indicators address linkages and STI outputs. These results show the areas of STI systems that can be better studied using the available scoreboards, because of the higher number of indicators available, and point out the areas where more research is needed. Results confirm the findings from the more critical literature that suggest an expansion from stakeholder engagement.

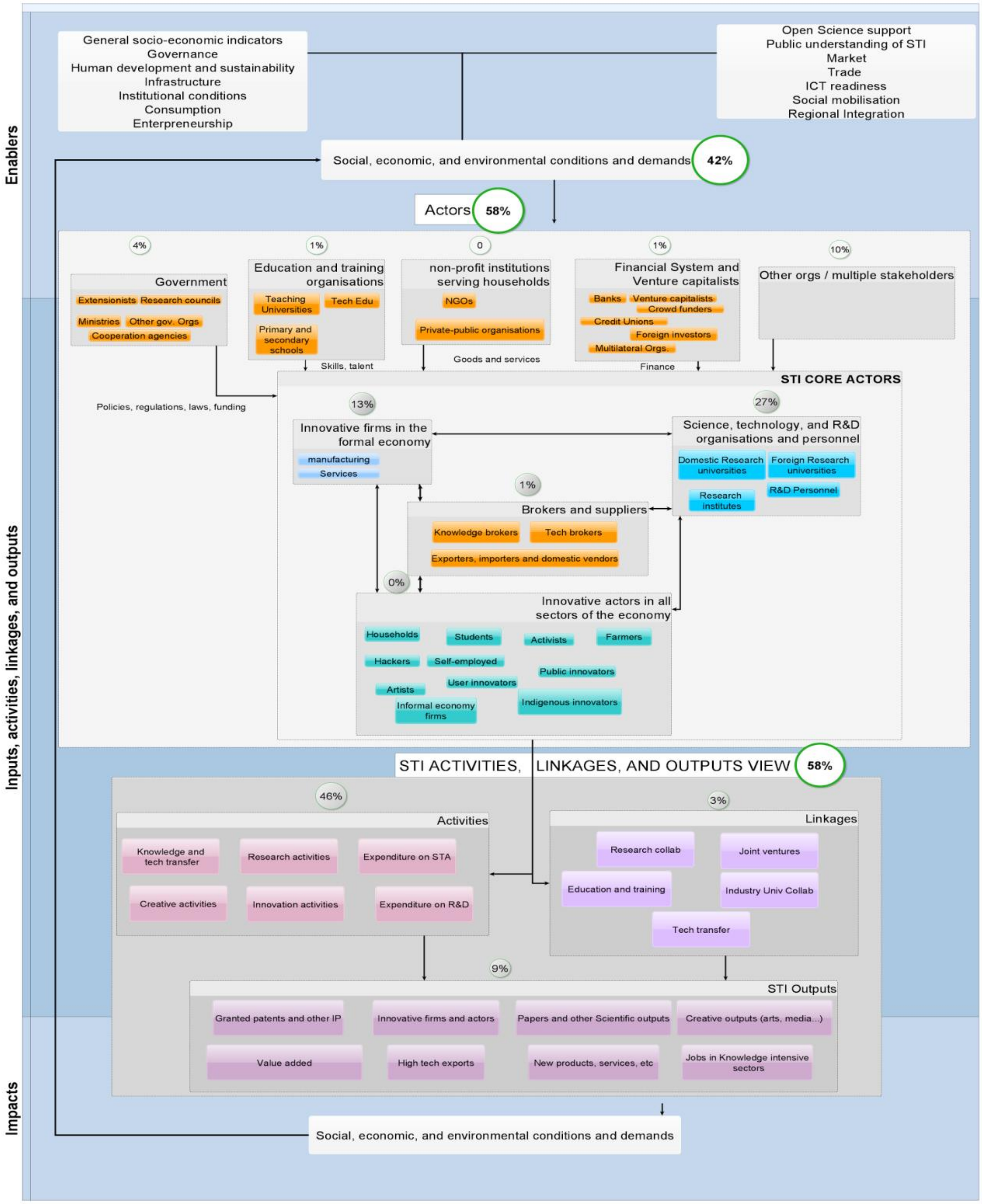


Figure 4: STI framework

5. Quality criteria definition

This section refers to part 3 of our methodology. The literature and stakeholder perspectives (see section 3.5) have shown that completeness, appropriateness, and relevance are not agreed upon any specific concepts, and we have not found a comprehensive evaluation of STI scoreboards using these three criteria. Based on our review and engagements, we present our definition for this research.

5.1. Completeness

We define completeness as the extent to which an indicator (or a set of indicators) captures a specific aspect of STI. This includes identifying missing data (*coverage*), understanding how the indicator is calculated (*composition*), and which actors it covers (*actors*).

Among these variables, *coverage* is amenable to an evaluation measurement because it is defined within minimum (min) and maximum (max) values. The max is 100% coverage of the African countries and we can assess the extent to which each indicator fulfills this. Composition and actors are not bound. In terms of composition, there is usually no one best way to calculate indicators, as this depends on their meaning and use. In the case of actors, there is no agreement on how many actors should be covered by each indicator, and again there can be trade-offs. Therefore, although we compute those two variables, we do not include them as part of the quantitative assessment of indicators.

Operationally, we calculated coverage as the percentage of data points missing in the series for 10 years across 54 African countries. The total data points expected are always 540. The higher the percentage of missing values, the less complete the indicator.

5.2. Appropriateness

We defined appropriateness in relation to the validity, timeliness, comparability, and coherence of an indicator. *Validity* is the degree to which an indicator represents its concept - whether it measures what it is supposed to. *Timeliness* indicates whether the information is reported at a suitable time. *Comparability* is the extent to which indicators allow to find similarities and dissimilarities between units of analysis. *Coherence* indicates the relatedness of indicators to specific aspects of STI.

A thorough assessment of *validity* would require knowing the methodology used to collect and measure each indicator in all countries and years. Although this could be done for some indicators by collecting the footnotes of scoreboards and their technical annexes, it would take a considerable time that exceeds the timeframe and resources of the project.

Timeliness is also problematic. Reporting time may differ from collection time. For instance, GERD can be reported in 2020 in a given country, but the collection time may differ. Also, there may not be information on when the data series are inputted, since it is not available each year. We provide a measure of timeliness as an approximation to this concept, by computing the mean number of years missing between the year of the latest data available (LDA) and 2019.

The following two assessment criteria were computed to address coherence. *Comparability* was computed as the number of countries with observations available for that indicator along with the time series. C*oherence* is partially assessed using the coefficient of variation and outlier identification. The coefficient of variation shows the dispersion of data points in the data series for each STI indicator. The number of outliers shows potential anomalous values for countries in each STI indicator, through the calculation of the Interquartile range (IQR). Data points that fall beyond or below 1.5 times this range are considered outliers.

5.3. Relevance

We defined relevance as a subjective measure, which refers to the extent to which the indicators contain the information needed to describe a given aspect of ST&I. We consider the construction of our STI framework as an initial assessment of the relevance of the indicators, and the perceptions on the relevance of indicators through a non-representative survey to KI described in section 3.5. As the number of respondents is not representative, we do not include the result of the relevant survey in the calculation of indicator quality.

Annex 6 provides the detailed operational definition of each criterion as they were applied to the STI scoreboard.

6. Scoreboard Construction

The previous sections have explained steps 1 to 3 in our methodology. This section builds on those previous steps (1-3) to present the scoreboard. The section constitutes steps 4-7 (identification, classification, and compilation of indicators) of our methodology as illustrated in Figure 1.

6.1. Identification

A key criterion for the construction of an STI scoreboard for African countries is that it can be reproduced with open access data from reliable sources. To ensure this, we reviewed all indicators used by the 15 STI-related scoreboards and identified the original data sources from which they obtain the data. We compiled a list of 1,013 scoreboard-indicator records. Annex 7 shows a description of each scoreboard and provides the list of indicators for the 15 scoreboards selected.

We then identified duplicate indicators by performing a match between the names of the indicators, and looking manually at their semantic proximity. In the first case, indicators that had the same names were identified as one indicator. In the second case, indicators were identified as a group of indicators. For instance, all indicators related to GERD, such as GERD funded by Higher Education and GERD funded by Government, were clustered into a group of GERD indicators. These procedures yielded 364 different indicators.

6.2. Classification and validation

Based on our framework, we classified each indicator into Dimension, Actors, Activities/Subjects, and Logical Framework. This classification was performed by the consultant team, based on the original classifications of the scoreboards when available, the literature reviewed, and our expert knowledge of the STI evaluation field.

We acknowledge that other classifications of the indicators are possible. For this reason, we considered it important to ensure that our classification is coherent with the data. We ran a principal component analysis (PCA) of the logical framework to study if the indicators are classified among different components, each providing coherent information on a part of the STI framework. We find that the indicators within each component (which were classified in each aspect of the logical framework) are closely related between them and are loosely related to indicators in other components.

We also produced a correlation matrix to see if different components were strongly associated with them. Very strong associations mean that the components are giving redundant information, pointing out potential overlaps in the classification. We did not find very strong correlations between components (see annex 8 section 2), which means that each component is contributing additional information.

Besides, we performed a cluster analysis of the indicators to study if an unsupervised classification of the indicators, based on their similarity, resembled our classification. Using Hierarchical Clustering on Principal Components (HCPC) (Husson, Josse, & Pages 2010), we find that an unsupervised clustering produces a similar grouping of indicators as inputs, outputs, enablers, linkages, and impacts. Annex 8 section 3 presents the results of this exercise. The three findings show that although other classifications are possible, ours offers one robust possibility.

6.3. Compilation

Once the original data sources were identified, we obtained the data through web downloads. We were able to find open access data for 263 indicators that were measured, at least for one African country. These indicators were compiled into a common format for analysis. The most important data producers are the World Bank and the United Nations. It should be noted, however, that key indicators such as the number of scientific publications and internet use indicators are not openly available. This poses limitations to the reproducibility of indicators found in well-known STI scoreboards such as the GII.

The following table shows the data sources used for the scoreboard and the number of indicators identified:

Table 3: Data sources and number of indicators for the STI scoreboard

Source	Number of indicators
World Development Indicators (WDI)	76
UN - Research and experimental development indicators	51
Global Competitiveness Index	43
UN - Innovation indicators	30
UN - SDGs indicators	29
UN - National monitoring indicators	15
World Economic Forum (WEF Enterprise Survey)	5
World Governance Indicators (WGI)	4
UN - Other indicators	2
UN - Demographic indicators	2
Penn World Tables 9.1 2020	2
Global Innovation Index (GII)	1
UN - films indicators	1
UN- Cultural Trade	1
UN - external data from ITU	1
Total	**263**

6.4. Quality criteria application

The quality criteria were applied to the full dataset of 54 African countries. To provide a synthetic measure of quality, we ranked each STI indicator according to their performance in the quality criteria defined above. We computed the mean ranking for each STI indicator in coverage, timeliness, coefficient of variation, and outliers, which accounts for the completeness and appropriateness aspects of quality. As mentioned, we did not consider the relevance of indicators in this calculation, given that the relevance survey is not representative, and we cannot safely assume that indicators that were not chosen by respondents are not relevant for other stakeholders. Annex 9, worksheet quality ranking, allows indicators to be ordered by quality measures (a lower number means a better-quality ranking).

6.5. Production of the scoreboard

The production of the scoreboard included selecting the data for the scoreboard, compiling the data, applying transformations to some of the indicators for comparability, and organizing the data into a spreadsheet.

6.5.1. Selection of indicators

Our selection of key indicators for the scoreboard was based on our STI framework presented in section 4. This selection was constrained by data availability from the main sources, as explained in steps 4 and 5. We compiled a set of 263 indicators. Our set of indicators is organized according to the following framework classifications:

- *STI domain*: Science and Technology; Innovation; Social, economic, and environmental context.
- *Logical framework*: Enablers; Inputs; Outputs; Linkages; and impacts.
- *Actors*: Education and training organizations; Financial organizations and venture capitalists; Brokers and suppliers; Education and training organizations; Financial organizations and venture capitalists; Firms in the formal sector; Multiple actors; non-profit institutions serving households; Science, technology, and R&D organizations and personnel; State Institutions; and Users/consumers.
- *STI activities/subjects*: Research activities; Innovation activities; Education and training; Expenditure on R&D; Expenditure on STA; Science and technology outputs; Innovation outputs; Creative outputs; Financial system; Governance and institutional development; Human and sustainable development; ICT readiness; Infrastructure; innovation determinants; Job and product market; Knowledge flows; and Regional integration.

6.5.2. Transformation of indicators and compilation

Some transformations were applied to specific indicators to allow comparability among indicators, domains, framework, and actors.

First, we normalized all indicators that were expressed in absolute terms by population, transforming them into per capita measures. These were indicators expressing currency, number of outputs such as patents, and number of people such as researchers.

Second, we produced rankings of countries based on the values of the STI indicators to allow the comparison of countries across different magnitudes and units of measurement (e.g., to compare between different activities measured as dollars per head or number of patents per head). We also calculated inverse rankings for those indicators that indicate an undesirable outcome, such as homicide rate.

Third, we used the latest available year for indicators that had missing values in most recent years. In the cases where the value of an STI indicator was not measured in the period 2010 to 2019, the value of the indicator is empty. The scoreboard uses 2019 as the reference year.

Indicators were compiled into a spreadsheet for easy manipulation by users. This is a common format used by scoreboards such as the GII, the GCI, and others. Annex 10 is the scoreboard for 54 African countries.

7. STI scoreboard for African countries

The steps presented in section 5 yielded a scoreboard for African countries composed of 263 STI indicators. Below we present the composition of the scoreboard and an assessment of the scoreboard indicators for African countries.

7.1. Proposed scoreboard

As discussed in section 5, each indicator was allocated to one logical framework category, actor, and/or activity. Table 4 reports the number of indicators allocated to each category, actor, or activity.

Table 4: Number of indicators by framework classification

Logical framework category	Enabler (128), Input (67), Linkages (14), Output (21), Impact (33)
Actors	State Institutions (84), Firms in the formal sector (50), Science, technology, and R&D organizations and personnel (34), Education and training organizations (33), Multiple actors (29), Brokers and suppliers (13), Financial organizations and venture capitalists (11), Users/consumers (8), Non-profit institutions serving households (1)
Activities / Subjects	Human and sustainable development (32), Education and training (32), Research activities (29), Job and product market (26), Governance and institutional development (25), Expenditure on R&D (23), Innovation outputs (23), Knowledge flows (15), Innovation activities (11), Expenditure on STA (10), ICT readiness (10), Science and technology outputs (9), Financial system (8), Infrastructure (6), Innovation determinants (3), Creative outputs (1)

Note: number of indicators in brackets

The scoreboard in annex 10, allows users to navigate the STI framework for each of the 54 African countries. We ranked countries for each indicator and provided the indicator value. The indicators have also been classified by region and income group as defined by the World Bank. The user can filter information on any of these categories by using the checkboxes at the top of the worksheets.

A unique feature of our proposed scoreboard is that the user can also filter indicators by their quality. This allows users to observe countries' STI only based on indicators with high quality, and assess how this compares when all indicators are taken into account. We have provided a column named Quality Ranking, which provides the mean ranking of each indicator for the criteria timeliness, coefficient of variation, outliers, and coverage.

Finally, we provide a compilation of data for all available countries in the world, and the 263 indicators as a basis for further comparative analysis. The data is provided in annex 11.

7.2. Assessment of scoreboard indicators- based on relevance, completeness and appropriateness

As explained in section 5, we applied the quality criteria to the 263 indicators for African countries. Overall, we see a worrying percentage of missing data that affects coverage and timeliness of the information. This confirms our findings from the literature review, which finds that missing data problems and other issues are usually hidden in footnotes, and are rarely taken into account in the production of rankings (Godin 2004; Schibany & Streicher 2008; Adam 2014). This is especially important for African countries that are usually ranked globally as if they had the same data availability as European countries. Below we discuss results on each quality criterion. Annex 12 provides the full tables.

7.2.1 Completeness

Table 5 reports the average percentage of missing values from 2010 to 2019 per logical framework category, and overall, for all indicators. The results show an alarming percentage of missing values for African countries across all framework categories, except for impact indicators. The percentages suggest that for more than half (62%) of the countries or years (2010-19) the data is missing. For inputs, outputs, and linkages, this share goes up to 80%. The high standard deviation suggests that there are large differences in the collection of information between different countries/years.

Table 5: Missing values

Logical framework	Mean of missing values percentage	Standard deviation
Enabler	52.0%	30.8%
Input	85.9%	12.5%
Linkages	78.3%	34.4%
Output	82.5%	26.1%
Impact	33.8%	28.7%
Overall	62.2%	32.5%

The differences between countries are seen in the number of indicators ever reported, and the percentage of values reported for those indicators in the 10 years. For instance, Ethiopia has the highest number of indicators reported (237 out of 263), but its average missing values percentage for those indicators in the 10 years is 48%. South Africa has an average of 23% missing values over the 10 years, but it has 199 indicators that were ever reported. The average of indicators ever reported for all countries is 165, with a standard deviation of 49. This again shows large differences in the collection of data across African countries. Egypt and South Africa provide the highest balance in these two variables, as measured by their average rankings. The lowest balance is achieved by Eritrea and South Sudan. Annex 12, worksheets “logical framework completeness” and “country completeness” provide the full tables.

7.2.2. Appropriateness

Appropriateness was measured by timeliness and coherence. This last variable was measured with the number of outliers and with the coefficient of variation. Table 6 shows the average time lag in years for the logical framework categories and the whole dataset. This is the number of consecutive years for which the data is available, allowing comparison over time. The mean lag of indicators is 5.1 years in the 10-year window, and the standard deviation 2.9. Impact indicators show the lowest lag in the sample. On average, countries report data for 37% of the years in the 10 years.

Table 6: Average time lag in years by logical framework category

Logical framework	Average lag in years	Standard deviation
Enabler	4.1	2.7
Input	6.8	1.2
Linkages	7.6	3.4
Output	7.6	2.9
Impact	2.7	2.2
Overall	5.1	2.9

The countries with a higher balance on the number of years with data and average time lag are South Africa and Egypt. The countries with a lower balance are South Sudan and Somalia. Worksheets “logical framework timeliness” and “country timeliness” in annex 12 show the full details on a table.

On Coherence, Table 7 reports the total number of outliers. We do not find a significant number of outliers in the data series at the indicator level. Still, the effect of these outliers on calculations has to be checked for each specific use case.

Table 7: Number of outliers by the logical framework

Logical Framework	Percentage of outliers
enabler	4.0%
input	5.4%
linkages	5.7%
output	6.1%

impact	4.4%
Overall	4.4%

Note: percentages are calculated as outliers over the total number of data points available for all indicators in each logical framework category.

The countries with more data points and fewer outliers are South Africa and Egypt, and the countries with fewer data points and more outliers are South Sudan and Somalia. Worksheets “outliers logical framework” and “country outliers” in annex 12 provide more details on tables.

On coherence, Table 8 reports the coefficient of variation (CV). Coefficient of variation varies considerably and can take negative values, as well as values above one. This makes it difficult to interpret the data when grouped by the logical framework, but we provide here the values for high-income countries for comparison. In any case, the higher the values, the higher the dispersion of the data. The following table shows a high dispersion, which in the cases above 1, means that the standard deviation is higher than the mean. This suggests that, despite the fact that we find few outliers, there is a very large variation in the value of indicators across countries and over time. Indeed, this may also suggest that the number of outliers is relatively low, because of the high variation in the data. CV and standard deviations for Africa do not differ much from high-income countries.

Table 8: Coefficient of variation (CV) by logical framework category.

Logical framework	**Africa**		**High-Income Countries**	
	Mean CV	**SD**	**Mean CV**	**SD**
enabler	1.3	2.7	1.6	3.8
input	1.2	1.1	1.0	1.1
linkages	0.8	0.4	1.1	1.5
output	0.8	1.0	0.7	0.6
impact	1.1	1.1	1.0	1.0
Overall	1.2	2.0	1.3	2.8

The highest coefficients of variation are seen in Libya and the Gambia, and the lowest in Gabon and Tunisia. Worksheets “coefficient of variation LF” and “country coefficient of variation” in annex 12 show the tables.

7.2.3. Relevance

As explained, the number of responses (50) does not allow us to analyze the relevance of indicators in a meaningful way. For this reason, the results on relevance are only exploratory and illustrate what some of the stakeholders consider most relevant in measuring STI, across framework categories. They provide some insights into the perceptions of users mainly from governmental and academic organizations.

To calculate the highest-ranked indicators in relation to relevance, we averaged the respondents' rankings in the variables *number of respondents* and *median score*. The following table shows the three most relevant indicators or groups of indicators for respondents by the logical framework.

Table 9: Indicators identified as most relevant by respondents

Enablers	*1) availability of research and training services; 2)* indicators related to health; 3) Indicators related to education attainment
Inputs	1) *Percentage of innovation-active firms in manufacturing*. 2) *availability of scientists and engineers.* 3) *Number of technicians*
Linkages	1) *Percentage of manufacturing firms that cooperated with universities or other higher education institutions.* 2) Indicators from national innovation surveys on sources of information for firms 3) *Total official development assistance for technical cooperation*
Output	1) indicators related to the percentage of innovative firms. 2) capacity for innovation. 3) firms that engaged in in-house R&D.
Impact	1) *Proportion of medium and high-tech industry value added in total value added (%)*. 2) Indicators related to health. 3) Indicators related to exports of goods and services and exports of design and creative services

We further provide an analysis of specific user category, various UK agencies, to not only examine which indicators they considered relevant, but also explore the indicators the UK has contributed to. Table 9 in annex 11 provides indicators prioritized by the UK agencies. As previously highlighted, indicators are prioritized by various users, and as such, these agencies prioritized indicators based on their research/funding mandate. Nevertheless, the UK government is keen on the enabling environment that supports focus around trade and business as well as strengthening UK investments in Africa.

The UK government through its agencies is keen on influencing policy, partnerships through research, and collaborations, and thus its big focus is on equitable partnerships. These agencies are also keen to interface science and policy, including extending this interaction/linkages to the private sector. Therefore, developing more indicators to demonstrate this aspect (public-private interface indicators), which the UK supports, would support its measurement. While important for the UK investments in STI, gender equality as an indicator needs to be considered holistically in the broader ODI, and included when developing African STI scoreboard. Nevertheless, this as discussed previously would require a second phase mapping with specific countries to entrench gender aspects into the indicators.

The UK also supports a huge focus around trade and business, and UK investments and aspects around new technologies are key. Nevertheless, even with the indicators and prioritization, the UK has had its fair share of challenges in funding and achieving the aspirations of the UK government. This is because the indicators currently used have different levels of volatility, where various changes occur much more slowly and as such, require lots of effort in time and resources to track the changes.

8. Stakeholder's views on the indicators considered most relevant

This section delves deep into specific indicators useful for decision making as outlined by stakeholders. Gross Expenditure in Research and Development was prioritized by many users not only as one of the relatively easy indicators to use in STI policy processes ,but because it also reflects the R&D intensity and has four components (BERD, GOVRD, HERD, and PNPERD) crucial for securing the intellectual property rights of innovators, including the reduction of production costs and the emergence of new products in the marketplace. The indicator also supports the tracking of R&D funding in the R&D performing countries, and allows for disaggregation. Additionally, it was argued that this indicator determines whether R&D is a worthwhile activity, and supports further development of other indicators such as researchers supported by R&D personnel and level of public expenditure on tertiary education, which determines the quantitative and qualitative changes in the system.

In other words, such an indicator is a precursor to many decision-making areas. These as identified in our framework as the enabler indicators. See for example, government effectiveness critical in the process of technological innovation, driver of the overall STI ecosystem, creates the right environment including basic law & order and prevailing peace, although they can play a disabling role. On the contrary, lack of this enabling condition can hamper business and innovation.

Human resources were established at the core of STI indicators commonly used. R&D personnel in both headcount and FTE (based on profession, qualification, the field of R&D, and age) are vital for the calculation of R&D expenditures through current costs (labor costs). The indicator, as used, supports the tracking of a portion of SDG 9.2 (increase number of researches in FTE per one million inhabitants). Other indicators such as statistical information on intellectual property helps policymakers to develop evidence-based policies as well as encourage innovation and creativity, including tracking of how innovative the firms are.

The analysis provides a snapshot of what stakeholders consider relevant, and demonstrates the subjectivity of the indicators to contexts, users, user needs, etc. This assessment provides an opportunity to engage with the available indicators but requires the need for data while engaging further with the different users.

Indicator quality ranking

As explained in section 5, we produced a ranking of indicators by quality criteria. We do not attempt to select indicators based on this criterion, and instead, provide all the 263 indicators in the scoreboard along with their quality assessments, to allow the users to build their own set of indicators. We present here the ranking only as one possible way to identify indicators that perform well on most quality criteria. We exclude from this exercise the relevance criteria. Table 10 shows the top 5 indicators by quality ranking by logical framework category.

Table 10: Indicators with the highest average quality (top 5)

Logical framework category	List of indicators
Enablers	Organized crime, social protection, Political stability -non-violence, Rule of law, Red List Index
Input	Availability of scientists and engineers, Government procurement of advanced tech products, Company spending on Research & Development, Tertiary education, academic staff (% female), Percentage of graduates from Science, Technology, Engineering and Mathematics programs in tertiary education who are female (%),
Linkages	FDI and technology transfer, University-industry collaboration in Research & Development, Total official development assistance for technical cooperation, Total official development assistance to medical research and basic health sectors, Percentage of manufacturing firms for which suppliers of equipment, materials, components or software were a highly important source of information
Outputs	GCI 12th pillar Innovation value, Capacity for innovation, Global Competitiveness Index rank, Global Innovation Index rank, Percent of firms whose new product/service is also new to the main market
Impact	Labor force participation rate, Employment to population ratio, Employment in services, Self-employed population, Employment in agriculture

9. Uptake strategy

The scoreboard provided here can be used in different ways. The most common use of scoreboards is as a ranking tool to compare the performance of countries. This is the approach taken by the GCI, GII, and other scoreboards analyzed in this project. To provide this standard view for the users of our STI scoreboard, we produced a global scoreboard that includes 216 countries (see annex 11). The global scoreboard is based on the latest available data for each country between 2010 and 2019.

Our purpose, however, is not to develop a ranking of countries, given that the data is very variable in terms of quality and also because an emphasis on rankings promotes competition instead of collaboration. We, instead, try to provide a view towards collaboration, which is in line with the African Union guidelines for research grants, aimed at "*Enhance(ing) intra-regional scientific collaboration and research that contributes to Africa's sustainable development"* (African Union 2021).

To provide a way to understand countries' similarities, Figure 6 shows clusters of African countries identified through their performance on enablers, inputs, outputs, linkages, and impacts. Beyond an interpretation based on performance, we interpret the clusters as indicators of opportunities to build research collaboration and cooperation strategies, for instance, by promoting Intra and inter-cluster research mobility.

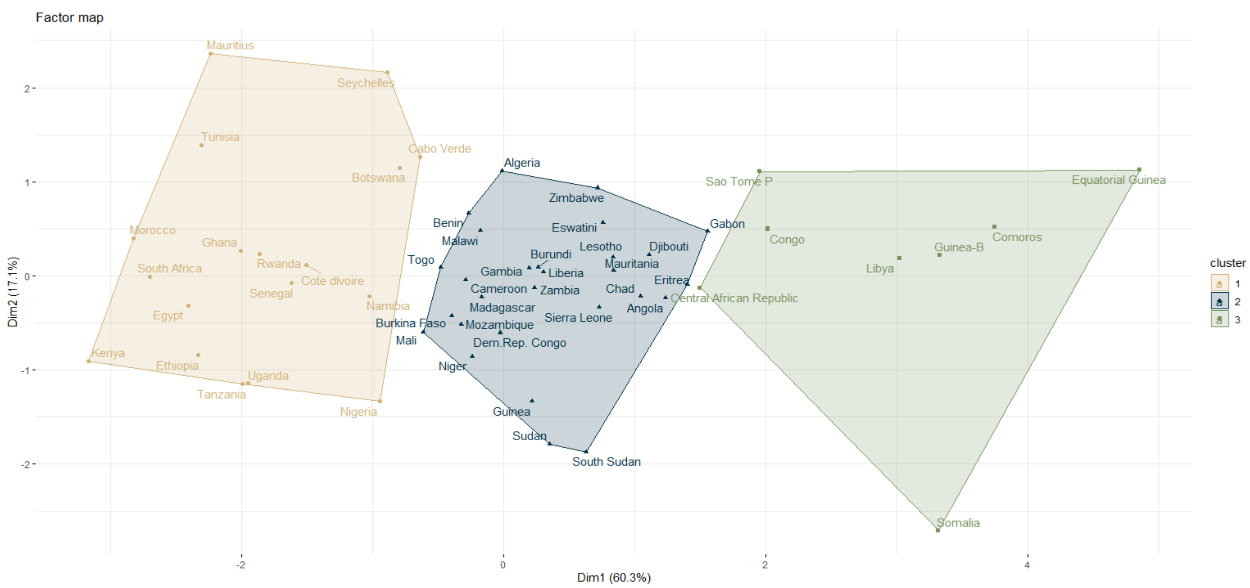


Figure 6: African countries by clusters, according to performance on logical framework categories.

The X axis represents the first component, and Y axis represents the second component. The percentage of variance explained by each component is identified within brackets.

Given that the scoreboard provides data globally, it can be used to derive strategies beyond Africa. For instance, it can be used to analyze Kenya with respect to potential global partners. Figure 7 shows a cluster of countries at the global level, with which Kenya has some similarities. Countries can use the scoreboards to prepare briefs when approaching global partners.

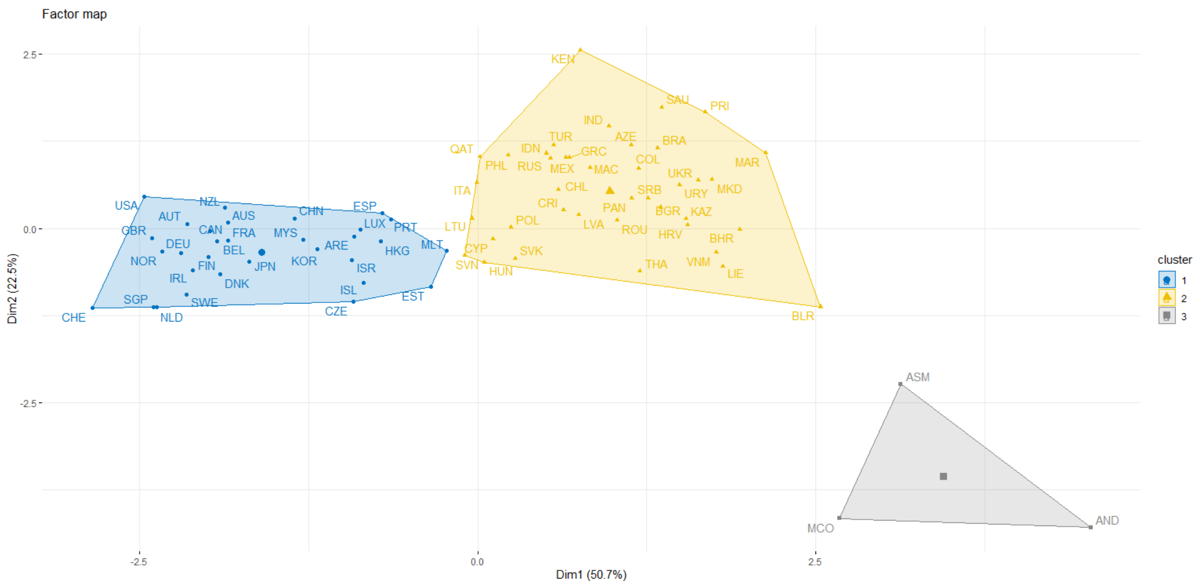


Figure 7: Clusters of countries, with a focus on Kenya.

The X axis represents the first component, and Y axis represents the second component. The percentage of variance explained by each component is identified within brackets.

Note: Three-letter ISO codes are provided for space constraints. Kenya's cluster is composed of Qatar, Italy, Lituania, Cyprus, Hungary, Belarus, Puerto Rico, Saudi Arabia, India, Turkey, Philippines, Russia, Mexico, Chile, Costa Rica, St. Lucia, Slovak Republic, Thailand, Vietnam, Liechtenstein, Bulgaria, Serbia, Ukraine, Macao - China, Greece, Indonesia, Panama, Croatia, and Slovenia.

As the cluster for Kenya may be too broad, a zoom into Kenya's cluster can show finer clusters, which the data allows. Figure 8 shows that Kenya has more similarities with some South American, European, and Asian countries than with other countries in Africa from this global perspective.

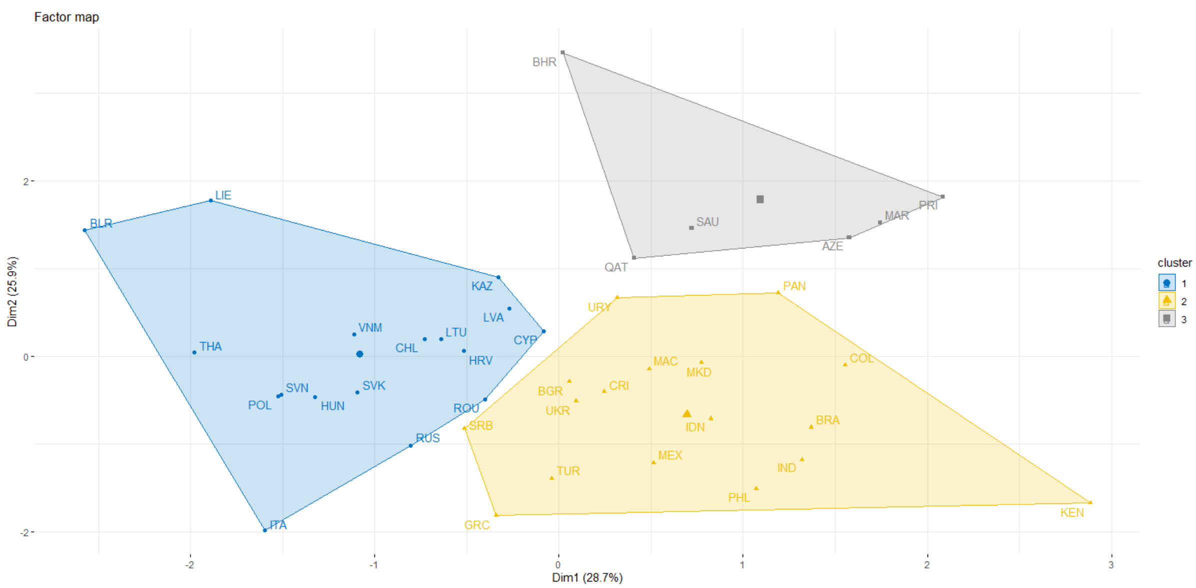


Figure 8: A zoom into Kenya's cluster

The X axis represents the first component, and Y axis represents the second component. The percentage of variance explained by each component is identified within brackets.

Note: Three-letter ISO codes are provided for space constraints. Kenya's cluster is composed of: Serbia, Bulgaria, Ukraine, Greece, Costa Rica, Indonesia, Mexico, Turkey, Greece, Philippines, India, Brasil, Colombia, Madagascar, and Macao-China

The above is but one way in which the scoreboard can be used. Beyond this, we see that there is potential to use the scoreboard for policy-relevant research. Examples of these uses include:

1. ***Interactive web services***: the potential of these data can be leveraged through a web-based, user-friendly platform in which citizens, policymakers, entrepreneurs, and academics can tailor their collection of indicators and obtain relevant information for their needs. We do not propose here to offer rankings as part of this interactive platform because of data quality.
2. ***Research***: The compiled data allows studying different research questions where the quality of data is acceptable. This includes determinants of STI variables of interest, the relationship between quality of data and performance of countries, identification of strengths and weaknesses, identification of information needs, critical analysis of rankings.
3. ***Policy discussions***: the data can be used to plan policy and funding agendas, as well as measuring African Union's progress towards quality in data production.
4. If the data is offered in open access, this can help users innovate in the use of these indicators.
5. The framework developed in this project hints at a research agenda on the construction of more relevant indicators, not only on improving the quality of the existing ones, but also on finding new ways to measure unaddressed dimensions of innovation.

To be able to use the scoreboard in those ways, it is fundamental to:

1. Support the necessary infrastructure, design, and development for making it available through a web platform.
2. A cross-cutting need is to provide a repository and resources for updating the data, as well as to include new indicators in the platform to achieve better coverage of the STI framework developed in this project.
3. For research, countries must increase support to the training of STI data scientists so that the data can feed into relevant research questions and a community of active researchers can be developed around them.
4. Open repositories, forums, hackathons, academic events, are needed to improve the uptake of this data.
5. Support research to continue developing and identifying ways to cover all dimensions of the STI framework so that Africa can count on a more complete and relevant source of information.

We hope we have contributed, with a solid basis, to understanding the quality of data on STI in Africa, and provided a set of relevant indicators conceptually organized around a framework. It is therefore worth noting that the scoreboard is an exploratory collection of indicators and an evaluation of quality of STI assessment in Africa, and not a mechanism to address the challenges faced by the current assessment processes. Further work will be required to identify ways in which some of the identified challenges can be addressed to ensure a robust ST&I in Africa. We look forward to co-leading the further development of the scoreboard and related research on STI policy together with the FCDO and partner countries and agencies.

10. References

Abrahams, L., & Pogue, T. E. (2012). South Africa's national system of innovation and knowledge economy evolution: thinking about 'less-favored regions'. *International Journal of Technological Learning, Innovation, and Development*, *5*(1-2), 58-82.

Adam, F. (2014). *Measuring national innovation performance: the Innovation Union Scoreboard revisited*. Slovenia, Ljubliana: Springer.

African Union (2021). Science and Technology division. https://au.int/en/st-division

Archibugi, D., Denni, M., & Filippetti, A. (2009). The technological capabilities of nations: The state of the art of synthetic indicators. *Technological Forecasting and Social Change*, *76*(7), 917-931.

Bartels, F. L., & Koria, R. (2014). Mapping, measuring and managing African national systems of innovation for policy and development: the case of the Ghana national system of innovation. *African Journal of Science, Technology, Innovation and Development*, *6*(5), 383-400.

Bartling, Sönke, & Sascha Friesike, eds. (2014). Opening Science. Cham: Springer International Publishing, 2014.https://doi.org/10.1007/978-3-319-00026-8

Binz, C., & Truffer, B. (2017). Global Innovation Systems—A conceptual framework for innovation dynamics in transnational contexts. *Research Policy*, 46(7), 1284-1298.

Bloom, N., Jones, C.I., van Reenen, J., Webb, M. (2020). Are ideas getting harder to find?†. Am. Econ. Rev. 110, 1104–1144. doi:10.1257/aer.20180338

Bordt, M., Rosa, J. M., & Boivin, J. (2007). Science, Technology, and Innovation for Sustainable Development: Towards a Conceptual Statistical Framework. In: OECD (2007). *Science, Technology and Innovation Indicators in a Changing World: Responding to Policy Needs*, 251-268.

Bornmann, L. (2013). What is the societal impact of research and how can it be assessed? A literature survey. *Journal of the American Society for information science and technology*, *64*(2), 217-233.

Charmes, J., Gault, F., & Wunsch-Vincent, S. (2016). Formulating an Agenda for the Measurement of Innovation in the Informal Economy. *The Informal Economy in Developing Nations: Hidden Engine of Innovation*, 536-549.

Ciarli, T., Savona, M., Thorpe, J., & Ayele, S. (2018). Innovation for inclusive structural change. A Framework and Research Agenda. *A Framework and Research Agenda (January 23, 2018). SWPS*, *4*.

Freeman, C. (1995). The 'national system of innovation' in historical perspective. *Cambridge Journal of Economics*, 19, 5-24

Frost, A., Hall, A., Marshall, F., Atela, J., Ndege, N., Ciarli, T., Shee, A., Confraria, H., Dolley, J., & Chataway, J. (2019). Understanding knowledge systems and what works to promote science technology and innovation in Kenya, Tanzania and Rwanda. *Global Development Network – Knowledge for Sustainable Development: The Research Policy Nexus, October 2019.*

Frost, A. et al. (2020) Understanding Knowledge Systems and What Works to Promote Science Technology and Innovation in Kenya, Rwanda and Tanzania, Unpublished report from the Knowledge Systems Innovation project to the UK Foreign, Commonwealth and Development Office (FCDO), OJEU Publication Reference Number: 2017/S 184-377250

Gault, F. (2007). Science, technology, and innovation indicators: opportunities for Africa. *African Statistical Journal*, *6*(May), 141-161.

Gault, F. (2013a). Innovation indicators and measurement: an overview. In Gault (ed) 2013. *Handbook of innovation indicators and measurement*. Edward Elgar Publishing.

Gault, F. (2013b). Innovation indicators and measurement: challenges. In Gault (ed) 2013. *Handbook of innovation indicators and measurement*. Edward Elgar Publishing.

Gault, F. (2018). Defining and measuring innovation in all sectors of the economy. *Research Policy*, 47(3), 617-622

Godin, B. (2007). Science, accounting and statistics: the input-output framework. *Research Policy*, 36(9), 1388-1403.

Godin, B. (2015). Models of innovation: Why models of innovation are models, or what work is being done in calling them, models?. *Social Studies of Science*, 45(4), 570-596.

Godin, B., & Doré, C. (2005). Measuring the impacts of science: Beyond the economic dimension. *INRS Urbanisation, Culture et Société*, 44.

Grupp, H., & Schubert, T. (2010). Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy*, *39*(1), 67-78.

Guellec, D., Van Pottelsberghe De La Potterie, B. (2004). From R&D to productivity growth: Do the institutional settings and the source of funds of R&D matter? *Oxf. Bull. Econ. Stat*. , 66, 353–378. doi:10.1111/j.1468-0084.2004.00083.x

Hagen-Zanker, J., & Mallett, R. (2016). How to do a rigorous, evidence-focused literature review in international development. A guidance note.

Hasche, N., Höglund, L., & Linton, G. (2019). Quadruple helix as a network of relationships: Creating value within a Swedish regional innovation system. Journal of Small Business & Entrepreneurship, 1-22.https://doi.org/10.1080/08276331.2019.1643134

Hollanders, H., & Janz, N. (2013). Scoreboards and indicator reports. In *Handbook of innovation indicators and measurement*. Edward Elgar Publishing.

Husson, F., Josse, J., & Pages, J. (2010). Principal component methods – hierarchical clustering - partitional clustering: why would we need to choose for visualizing data?. http://www.sthda.com/english/upload/hcpc_husson_josse.pdf

Iizuka, M., Mawoko, P., & Gault, F. (2015). Innovation for development in Southern & Eastern Africa: Challenges for promoting ST&I policy. UNU-MERIT Policy Brief, 1, 1-8.

Janger, J., Schubert, T., Andries, P., Rammer, C., & Hoskens, M. (2017). The EU 2020 innovation indicator: A step forward in measuring innovation outputs and outcomes?. *Research Policy*, 46(1), 30-42

Jerven, M. (2013). *Poor numbers: how we are misled by African development statistics and what to do about it*. Cornell University Press.

Kahn, M., & Blankley, W. (2005). The changing face of South Africa's national system of innovation, 1991–2001. *Industry and Higher Education*, *19*(2), 121-130.

Klingebiel, R., & Stadler, C. (2015). Opportunities and challenges for empirical strategy research in Africa. *Africa Journal of Management*, 1(2), 194-200.

Kohl, H., Cap, J. P., Blaich, E., Von Raesfeld, A., Dameri, R. P., Garelli, R., ... & Beltrametti, L. (2015). The Innovation Network Scoreboard. Towards Key Performance Indicators for the Assessment of Innovation Networks. The University of Genoa (Hg.): Proceedings of ECIE, 375-383.

Kruss, G. (2018). Towards an agenda for measuring innovation in emerging African economies: what can we learn from the case of South Africa?. *International Journal of Technological Learning, Innovation, and Development*, *10*(3-4), 347-365

Lundvall, B., (1992). National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning. Pinter, London.

Martin, B.R. (2007). Assessing the impact of basic research on society and the economy. Paper presented at the Rethinking the impact of basic research on society and the economy (WF-EST International Conference, 11 May 2007), Vienna, Austria.

Merton, R. (1943)[1973]. The normative structure of science. In N. Storer (Ed.), The sociology of science(pp. 267–278). Chicago and London: The University of Chicago Press.

Molas-Gallart, J., Salter, A., Patel, P., Scott, A., & Duran, X. (2002).Measuring third stream activities. Final report to the Russell Group of universities. Brighton, United Kingdom: Science and Technology PolicyResearch Unit, University of Sussex.

OECD (2018). Guidelines for collecting, reporting, and using data on innovation. The Measurement of Scientific, Technological, and Innovation Activities. OECD. https://www.oecd.org/science/oslo-manual-2018-9789264304604-en.htm

OECD. (1963). Frascati Manual, first edition. OECD.

OECD. (1992). Oslo Manual, first edition. OECD.

OECD. (2015). Frascati manual 2015: Guidelines for collecting and reporting data on research and experimental development. OECD Publishing.

Ofir, Z., Schwandt, T., Duggan, C., & McLean, R. (2016). Research Quality Plus (RQ+): a holistic approach to evaluating research.

Pouris, A. (2016). South African innovation scorecard framework. http://www.naci.org.za/wp-content/uploads/2016/10/Composite-Innovation-Indicators-for-South-Africa-discussion-document.pdf

Ràfols, I. (2019). S&T indicators in the wild: Contextualization and participation for responsible metrics. *Research Evaluation*, *28*(1), 7-22.

Schibany, A., & Streicher, G. (2008). The European innovation scoreboard: Drowning by numbers?. *Science and Public Policy*, *35*(10), 717-732

Solow, R.M. (1957). Technical Change and the Aggregate Production Function. Rev. Econ. Stat. 39, 312. doi:10.2307/1926047

Tahamtan, I., & Bornmann, L. (2020). Altmetrics and societal impact measurements: Match or mismatch? A literature review. *El profesional de la información (EPI)*, *29*(1).

Unesco (2010). Measuring R&D. Challenges faced by developing countries. Unesco. http://uis.unesco.org/sites/default/files/documents/measuring-rd-challenges-faced-by-developing-countries-2010-en.pdf

Von Schomberg, R. (2013). A vision of responsible research and innovation. In Responsible Innovation: Managing the responsible emergence of science and innovation in society, 51-74.

Wickson, F., & Carew, A. L. (2014). Quality criteria and indicators for responsible research and innovation: Learning from transdisciplinarity. *Journal of Responsible Innovation*, *1*(3), 254-273.

Wouters, P., Rafols, I., Oancea, A., Lynn-Kamerlin, S. C., Britt Holbrooke, J., & Jacob, M. (2019). Indicator Frameworks for Fostering Open Knowledge Practices in Science and Scholarship. European Commission. DOI: 10.2777/445286

11. Annexes

Annex 1. Literature review procedure

Annex 2. A detailed list of references

Annex 3. Relevance, validity, coherence workshop report

Annex 4. Overview of citations to scoreboards

Annex 5. Sustainability and gender analysis

Annex 6. Quality criteria detailed definition

Annex 7. STI indicator list

Annex 8. Validation of indicator classifications

Annex 9. Quality assessment

Annex 10. Scoreboard for 54 African countries

Annex 11. Global scoreboard for 216 countries

Annex 12. Quality assessment analysis (tables)

Annex 13: Table 11 providing relevant indicators as selected by UK agencies.