Identifying Institutional Structures for Data Policy and Governance Frameworks: Case for The Education Sector in Kenya

Eldah Onsomu

Gideon Nyakundi

Japheth Kathenge

Working Paper 045-2026

Bringing Rigour and Evidence to Economic Policy Making in Africa

CONSORTIUM POUR LA RECHERCHE ÉCONOMIQUE EN AFRIQUE

IDENTIFYING INSTITUTIONAL STRUCTURES FOR DATA POLICY AND GOVERNANCE FRAMEWORKS: CASE FOR THE EDUCATION SECTOR IN KENYA

BY

Eldah Onsomu

Gideon Nyakundi

Japheth Kathenge

Kenya Institute for Public Policy Research and Analysis (KIPPRA)

AERC Research Paper 045-2026
African Economic Research Consortium, Nairobi
February 2026

Published by: The African Economic Research Consortium

P.O. Box 62882 - City Square
Nairobi 00200, Kenya

Abstract

The population aged 35 years and below constitutes 75% of the population in Kenya, with 36.1% of the population aged between 15 and 35 years, presenting a significant development opportunity. If properly educated and skilled, this demographic could drive innovation, productivity, and sustainable development. However, education systems across the sub-Saharan region, Kenya included, often suffer from inefficiencies that stem from weak data coordination and management. This study aims to critically assess the state of data governance in Kenya's education sector and the effects on education outcomes. Five foundational pillars of data governance—Roles and Responsibilities, Privacy Standards, Policies, Tools and Practices, and Processes and Procedures—were assessed using six key operational dimensions.

The overall Data Governance Index (DGI) shows a clear and positive relationship with the Human Development Index (HDI) across Kenya's counties, confirming that counties with stronger data governance frameworks tend to achieve higher levels of education outcomes and development. Counties with well-structured data systems—characterized by high data quality, accessibility, timeliness, security, metadata documentation, and stewardship—demonstrate stronger planning capabilities, more efficient service delivery, and better outcomes in the education sector.

However, main constraints affecting data governance in the education sector include: Undefined roles and responsibilities in counties; Lack of dedicated data teams and committees; Inconsistent funding for data leadership and capacity building; Incomplete or outdated metadata in counties; Lack of standard documentation guidelines and tools and Limited staff awareness and technical capacity.

Interventions towards improving data governance in the education sector include: implementing mandatory data quality checks and validation protocols; automating collection processes to reduce manual errors; and providing training on data accuracy and integrity. Enforce encryption, access controls, and multi-factor authentication; Provide regular staff training on cybersecurity; Monitor and report compliance with data protection laws. Develop open-access platforms for education data; Standardize accessibility policies across counties; Create user feedback channels for improvement; Enforce routine data submission timelines; Invest in mobile data collection tools and dashboards; Schedule periodic data audits; Institutionalize stewardship roles and guidelines; Form cross-sectoral education data committees; Allocate dedicated budget lines for data leadership activities; Adopt national metadata standards; Train staff on metadata use and documentation; Establish centralized metadata repositories per county.

1. Introduction

Education is universally recognized as a key driver of economic development, social equity, and personal empowerment. In the 21st century, data has become an indispensable tool in transforming education systems worldwide. From planning and resource allocation to learner tracking and policy evaluation, data empowers governments to respond to challenges and craft evidence-based solutions. As the OECD notes, data governance refers to the arrangements—technical, regulatory, policy, and institutional—that manage the data lifecycle, from creation to deletion, across national and organizational boundaries (OECD, 2022).

In advanced economies, the integration of data governance into education systems has enabled more targeted interventions, improved student performance monitoring, and strengthened accountability. Countries like Estonia, New Zealand, and Canada have implemented robust data governance frameworks, ensuring interoperability of education systems and protecting student privacy rights while maximizing data value. These frameworks are not only technical blueprints but also reflect a culture of transparency, ethical standards, and data stewardship.

In Sub-Saharan Africa, the urgency for effective data governance in education is amplified by demographic and socioeconomic realities. The region is home to the youngest population in the world, with over 60% under the age of 25 (World Economic Forum, 2023). This "youth bulge" presents both a development opportunity and a formidable challenge. If properly educated and skilled, this demographic could drive innovation, productivity, and sustainable development. However, as the African Union's Continental Education Strategy for Africa (CESA 2016–2025) acknowledges, education systems across the region often suffer from inefficiencies that stem from poor data management.

Despite progress, systemic weaknesses persist:

i) **Data Fragmentation**: Education data is often collected by various actors—ministries, regional authorities, counties, schools, NGOs—using incompatible systems. This results in siloed information that cannot be aggregated or compared across institutions or jurisdictions. For example, A UNESCO IIEP comparative analysis of six Sub-Saharan countries (Gambia, Ghana, Guinea, Namibia, Senegal, Zambia) found that learning assessment and EMIS data were frequently underutilized in national planning processes, due to weak coordination, limited dissemination, and capacity gaps (Raudonyte, & Foimapafisi, 2022).

ii) **Data Quality Issues**: Outdated or incomplete data continues to hamper strategic planning. Materu (2007) highlighted how inaccuracies in school enrollment figures and teacher deployment in several African countries led to misaligned budgeting and resource allocation.

iii) **Limited Accessibility**: Disparities in digital infrastructure and institutional capacity mean that many regions lack the tools or expertise to access and analyze education data, leading to inequities in planning and response strategies (Tchamyou et al., 2019).

iv) **Underutilization of Data**: Perhaps most critically, data—when available—is often not used effectively in decision-making. This is partly due to weak data infrastructure, lack of analytical capacity, and low levels of data literacy among education administrators. As a result, many education policies and interventions are not aligned with real needs on the ground.

Without a coherent governance framework to ensure quality, secure, and ethical data practices, the potential for education data to drive inclusive and responsive development in Sub-Saharan Africa remains largely untapped.

Kenya has made commendable progress in expanding access to education. Based on the o the Kenya National Bureau of Statistics (KNBS), there were 49,033 registered learning institutions as of 2023, serving approximately 15 million learners from pre-primary to tertiary levels (KNBS, 2024). This growth has been accompanied by digitization efforts such as the National Education Management Information System (NEMIS), biometric student attendance tracking, and digital curriculum delivery platforms.

However, significant data governance gaps remain:

i) **Fragmentation Across Agencies**: The Ministry of Education (MoE), Teachers Service Commission (TSC), Kenya National Examinations Council (KNEC), Kenya Institute of Curriculum Development (KICD), and county departments all collect education data, often using different formats and with varying frequency. This leads to duplicated efforts, inefficiencies, and inconsistencies in national reporting.

ii) **Inconsistent Data Quality**: Reports from the 2024 National School Census Pilot indicate that disparities in school registration, teacher qualifications, and infrastructure data persist, especially in marginalized counties such as Turkana and Mandera.

iii) **Data Protection and Privacy Gaps**: With increasing use of biometric data and student tracking technologies, concerns have emerged regarding compliance with the **Data Protection Act (2019)**. While Kenya has made strides by establishing the Office of the Data Protection Commissioner (ODPC), most educational institutions have yet to operationalize internal policies for data security, access control, and informed consent.

iv) **Low Data Use in Policy Making**: Low Data Use in Policy Making: Despite ongoing investments in Education Management Information Systems (EMIS), evidence suggests that data remains underutilized in informing policy and planning decisions at the county level. Wamutoro et al. (2022) observed that education officers in Uasin Gishu County rarely used EMIS data for strategic decision-making, often due to limited technical capacity and insufficient integration of data use into institutional processes, resulting in policies that may not adequately reflect ground-level needs or promote equity.

These challenges mirror those observed in other sectors, such as agriculture, where the Ministry of Agriculture, Livestock, Fisheries, and Cooperatives (MoALFC) implemented a national data governance framework to address fragmentation and ensure consistency. The success of that framework offers a valuable model for the education sector.

By adopting a comprehensive data governance approach grounded in international best practices and adapted to local realities, Kenya can transform its education system into a more responsive, accountable, and inclusive engine for national development. Such a framework would not only support the realization of Vision 2030 and the Sustainable Development Goals (SDGs) but also ensure that every learner in Kenya is counted, protected, and empowered through the responsible use of data.

Human capital—the stock of knowledge, skills, and health embodied in individuals—plays a pivotal role in national development. Among its key determinants, education stands out as the most direct and measurable contributor. According to Human Capital Theory, education enhances individuals' productivity, earning capacity, and capacity for innovation, thereby generating positive externalities for society and economic systems at large (Becker, 2009; Schultz, 1961). In Kenya, where over 70% of the population is under 35 years of age, education is both a strategic investment and a critical foundation for long-term socioeconomic transformation.

Yet the effectiveness of education as a tool for human capital development is increasingly mediated by how well education data is governed. Effective data governance ensures that

policies, resource allocation, and reforms are evidence-based, equitable, and responsive. Without reliable, timely, and secure data, education systems operate in the dark—unable to accurately diagnose learning challenges, monitor outcomes, or evaluate policy interventions. Data governance in education, therefore, becomes not a peripheral administrative task, but a central pillar of national human capital strategy.

In this context, investing in robust education data systems—anchored in sound governance frameworks—is imperative. Kenya's efforts to digitize education data through platforms like the National Education Management Information System (NEMIS) represent progress. However, persistent challenges related to data fragmentation, quality control, limited metadata documentation, and weak data stewardship continue to constrain the value of such systems. These issues undermine Kenya's potential to improve its Human Capital Index (HCI), particularly with regard to learning-adjusted years of schooling.

Theoretical and empirical evidence strongly support the case for improved data governance. Becker (2009) underscores that education is not merely a consumption good but a productive investment with measurable returns. Schultz (1961) reinforces this by framing education as a public good with transformative potential for nations. Heckman (2008) highlights early childhood education as a particularly high-yield investment in breaking cycles of poverty and improving long-term health and labor outcomes. Chetty et al. (2014) demonstrate that the accurate assessment of teacher effectiveness through value-added models can predict long-term student earnings—affirming the role of micro-level data governance in macro-level economic outcomes.

At the macro level, Hanushek and Woessmann (2022) present compelling evidence that education quality—not merely years of schooling—is the principal driver of long-term economic growth. Their concept of "knowledge capital" emphasizes the importance of not just education quantity, but the effectiveness with which learning is delivered and assessed. Similarly, the World Bank (2018) stresses that learning outcomes, not enrolment rates alone, determine the real return on education investment. These findings underscore the urgency of moving from enrolment-driven to learning-driven education systems—a transition that depends on high-quality data governance. Moreover, the link between education data governance and policy effectiveness is well established. Barrera-Osorio and Linden (2009) find that countries investing in education data systems and using them systematically report improved outcomes in education quality, access, and equity. In Kenya, leveraging such insights can enhance decision-making in areas such as curriculum development, teacher deployment, infrastructure planning, and equity targeting.

In sum, education data governance serves as both an enabler and an amplifier of human capital development. It ensures that investments in education are strategic, impactful, and inclusive. For Kenya to realize the full potential of its education sector in advancing human capital, it must institutionalize data governance as a core element of policy design and implementation. This includes operationalizing a national education data governance framework with clear roles, standards, and safeguards—supported by training, funding, and institutional accountability.

In light of the challenges and opportunities outlined, this study aims to critically assess the state of data governance in Kenya's education sector. It seeks to understand current practices, identify existing gaps, and explore the implications of improved data governance on education quality and human capital development. Specifically, the objectives include: assessing the institutional frameworks and policies guiding education data management; evaluating the extent to which data is used for decision-making at various administrative levels; identifying challenges related to data quality, accessibility, and privacy; and proposing a comprehensive framework that aligns with Kenya's Data Protection Act (2019) and international best

practices. Through this analysis, the study contributes to ongoing efforts to enhance transparency, accountability, and equity in the education system through responsible and effective data use.

2. Methodology

To explore the current state of data governance in Kenya's education sector, this study employed a qualitative methodology with three primary approaches:

Related desk review

A comprehensive desktop review was conducted to examine national and international literature on data governance policies, strategies, and frameworks. Key documents reviewed included the Data Protection Act (2019), the Basic Education Act (2013), the National Education Sector Strategic Plan (NESSP 2018–2022), and sector-specific documents from ministries and agencies such as the Ministry of Education (MoE), Teachers Service Commission (TSC), and Kenya National Examinations Council (KNEC). In addition, the study drew insights from global and regional frameworks such as the OECD Data Governance Framework, the Continental Education Strategy for Africa (CESA), and the African Union Digital Transformation Strategy. These resources provided context on data governance principles, institutional roles, privacy and security standards, and emerging best practices.

The study identified and analysed key stakeholders involved in the generation, processing, and use of education data in Kenya. This included actors from the public sector, private education technology providers, development partners, school-level administrators, and regulatory bodies. The mapping exercise focused on understanding institutional mandates, inter-agency coordination mechanisms, data flows, and the current use of data in decision-making processes. This helped establish a baseline for assessing governance gaps and alignment with national legal requirements.

Key informant interviews (KIIs)

Structured interviews were conducted with selected informants across the education sector to gather in-depth perspectives on current data governance practices, challenges, and opportunities. Participants included Ministry of Education officials, county education directors, quality assurance officers, data protection officers, NEMIS administrators, and education data specialists. The participants were selected from both the national government and county government institutions. The interviews provided insights into how data is managed, shared, and secured in practice, highlighting both the strengths and the barriers faced at operational and policy levels. These insights were triangulated with findings from the desktop review and stakeholder analysis to ensure a comprehensive understanding of the context. A Data Governance Index (DGI) was computed based on the data and information collected from the various stakeholders across counties to give an indication of the state of data governance in Kenya. The data that was utilized to compute the national and county-level indices were calculated on the basis of assessments by the Institute in 2024.

The five foundational pillars of data governance—Roles and Responsibilities, Privacy Standards, Policies, Tools and Practices, and Processes and Procedures—were assessed using six key operational dimensions. These dimensions provide a measurable lens through which the maturity, gaps, and strengths of existing governance structures can be understood. Sample statements used to measure each dimension are given in Table 1.

Table 1: Data Governance Dimension

Data Governance Dimension	Sample Statements by stakeholder
Data Quality	Regular training is conducted to maintain and improve data quality standards

Data Security	Measures (e.g. Data encryption and access control) are in place to ensure the confidentiality of sensitive data.
Data Access	Data in the sector is readily available to authorized users
Data Timeliness	Data in the sector is collected and reported promptly
Data Stewardship/ Leadership	Data roles and responsibilities are well-defined and understood.
Metadata Documentation and organization	Metadata documentation is comprehensive and up-to-date

This was computed as follows:

Scaled values for the indicators were computed using the formula

$$S_i = \frac{X_i - Min(x)}{Max\ (x) - Min\ (x)}$$

Where Si = Scaled value for the indicator and Xi = Data value of the indicator.

A composite indicator was computed using the formula:

$$Weighted\ Index\ \ = \frac{\sum W_i\ .S_{\ i}}{\sum W_i}$$

Where Wi represents the unique weights associated with the scaled value Si.

All indicators included in the study were assumed to have equal importance in contributing to data governance in education.

3. Status of Data Governance in Kenya

Effective data governance is a critical enabler of digital transformation, ensuring accountability and fostering inclusive development. In an increasingly data-driven world, the ability of a country to manage, protect, and leverage its data assets determines the effectiveness of its policies, the trust of its citizens, and the competitiveness of its economy. While global frameworks offer guiding principles, each country faces unique legal, institutional, technological, and socio-political contexts that shape how data governance should be designed and implemented. Tailoring governance systems to national realities ensures that data policies are not only compliant with international norms but also responsive to local needs. For developing countries like Kenya, strengthening data governance is especially vital to enhance public service delivery, monitor development outcomes, safeguard privacy, and attract data-driven investment.

To determine the status of data governance in Kenya, this analysis adopted the OECD Framework on Data Governance for the Public Sector (OECD, 2022), a globally recognized model developed to assess and strengthen how governments manage data as a strategic asset. The framework evaluates data governance maturity across six thematic pillars: strategic, regulatory, responsible, structural, participatory, and international dimensions. Each pillar comprises specific indicators—such as national data strategies, personal data protection laws, ethical frameworks, institutional bodies, public consultations, and global commitments—designed to capture the extent to which a country ensures the availability, accessibility, quality, protection, and ethical use of public sector data. The OECD framework provides a structured lens for diagnosing strengths and gaps in data governance, allowing countries like Kenya to benchmark their practices and align with international best practices for digital transformation and policy innovation.

Table 2: Status of Data Governance in Kenya

Data Governance Attributes	Indicator	Status	Score	Weight	Weighted Score
Strategic (1/6)	National Data Strategy	No evidence	0	1/24	-
	Public Administration Data Strategy	No evidence	0	1/24	-
	AI Strategy	Draft (2025-2030)	1	1/24	0.04
	Strategy for Data in Emerging Digital Ecosystems	Data & Digital Transformation Strategy 2024–2028	1	1/24	0.04
Regulatory(1/6)	Personal Data Protection Law	Data Protection Act, 2019	1	1/30	0.03
	Open Data Law for the proactive release of government information	Kenya's Access to Information Act of 2016	1	1/30	0.03
	Freedom of Information Law	Article 35 of the Kenyan Constitution which	1	1/30	0.03

		guarantees the right to access information.			
	Right to be protected from Automated Decision-Making	Section 35 of the Kenya Data Protection Act of 2019	1	1/30	0.03
	Right to Data Portability	Data Protection Act, 2019	1	1/30	0.03
Responsible(1/6)	Data Charter	No	0	1/30	-
	Public Sector Data Ethics Framework	No	0	1/30	-
	Responsible AI Initiatives	No	0	1/30	-
	Trust Framework for Digital Identity Management	No- (Kenya Information Communications Act (1998, section 83E))	0	1/30	-
	Guidelines for Nongovernmental Data Sharing	Guidance Note on the Data Sharing Code (ODPC)	1	1/30	0.03
Structural(1/6)	Personal Data Protection Body	ODPC	1	1/24	0.04
	Open Data Portal	Kenya Open Data Initiative (KODI)	1	1/24	0.04
	Open Data Coordinating Body	No evidence	0	1/24	-
	Public Sector Data Governance Body	ODPC for personal data	1	1/24	0.04
Participatory(1/6)	Public Consultation on Data	Through the ODPC	1	1/18	0.06
	Government Response to Consultation	No evidence	0	1/18	-
	Multistakeholder Advisory Body	No evidence	0	1/18	-
International(1/6)	Convention 108+	Not a member	0	1/30	-
	Open Government Partnership	Yes- joined in 2011	1	1/30	0.03
	OECD AI Principles	Yes- adopted	1	1/30	0.03
	Binding Trade Agreements on Cross-Border Data Flows	No evidence	0	1/30	-
	Budapest Convention	No evidence	0	1/30	-
Total				1	0.53

Kenya's data governance landscape demonstrates significant progress in regulatory and structural dimensions, supported by the enactment of the Data Protection Act (2019) and the establishment of the Office of the Data Protection Commissioner (ODPC) (Table 2). These frameworks provide a strong legal foundation for personal data protection, data subject rights, and institutional oversight. Additionally, platforms like the Kenya Open Data Initiative (KODI) contribute to promoting transparency and public access to data. However, coordination across government entities remains fragmented, with no centralized body to oversee open data governance or enforce consistent data standards across public institutions.

In the strategic and responsible pillars, Kenya's development is more nascent. The country lacks a comprehensive National Data Strategy and Public Administration Data Strategy, though initiatives like the draft AI Strategy (2025–2030) and Digital Transformation Strategy (2024–2028) show emerging momentum. On the responsible front, there are major gaps in ethical governance, with no current Data Charter, Public Sector Data Ethics Framework, or formal Trust Framework for Digital Identity Management. While the ODPC has issued a Guidance Note on Data Sharing, broader policies that ensure transparency, accountability, and ethical data use are still absent, especially for emerging digital ecosystems.

Kenya's participatory and international data governance attributes reflect mixed results. The country has mechanisms for public consultation, particularly through ODPC-led engagements, yet there is no formal process for government response to public feedback or an established multistakeholder advisory body. Internationally, Kenya is engaged through membership in the Open Government Partnership and its adoption of the OECD AI Principles, but it has not yet ratified key frameworks such as Convention 108+ or the Budapest Convention, nor entered into binding agreements on cross-border data flows. Advancing in these areas will require deliberate national efforts to institutionalize open, inclusive, and globally aligned data governance systems.

3.1. Data Governance in Education

Kenya's education system collects a diverse array of data at the institutional, county, and national levels. These datasets are broadly categorized as follows:

i) **Learner Data**: Includes demographic information (age, gender, disability status), enrolment, attendance, performance metrics, and progression records. Biometric data (e.g., fingerprints for attendance) is also being used in select schools.

ii) **Teacher and Workforce Data**: Captured by the Teachers Service Commission (TSC), this includes teacher registration, qualifications, deployment, performance appraisals, and professional development records.

iii) **School Infrastructure and Resource Data**: Managed within the National Education Management Information System (NEMIS), this data covers school locations, facilities (classrooms, laboratories, water, and sanitation), student-teacher ratios, and ICT access levels.

iv) **Finance and Budget Data**: Captures information on capitation disbursement, school fees, development funding, and utilization—integrated into public financial systems at the Ministry level.

v) **Assessment and Examination Data**: Managed by the Kenya National Examinations Council (KNEC), this includes national exam performance, school-based assessments under CBC, and placement records.

Despite the richness of data types, fragmentation remains a challenge. Institutions use different platforms and standards, making integration and harmonization difficult.

3.2. Regulatory Environment for Education Data Governance

Kenya's data governance landscape is underpinned by the Data Protection Act (2019), which applies across all sectors, including education. Key principles from this Act relevant to education data include:

i) Lawfulness, Fairness, and Transparency: Educational institutions must collect and process learner and staff data transparently and with appropriate consent.

ii) Purpose, Limitation, and Data Minimization: Only data necessary for educational objectives should be collected and retained.

iii) Accuracy and Storage Limitation: Institutions must maintain accurate and up-to-date records and ensure that data is not stored longer than necessary.

iv) Accountability and Security Safeguards: Schools, education boards, and data processors must implement both organizational and technical measures to secure personal data.

However, the operationalization of these provisions within the education sector is not consistent. Most schools do not have trained data protection officers (DPOs), nor do they conduct data protection impact assessments (DPIAs). Additionally, compliance audits by the Office of the Data Protection Commissioner (ODPC) reveal low levels of awareness and adoption of data privacy protocols in public learning institutions.

3.3. Existing Policies and Guidelines on Education Data

Several frameworks actively support strong data governance in education by promoting ethical, secure, and evidence-based data practices. At the global level, OECD Data Governance Framework encourages countries to adopt integrated data strategies that ensure privacy, quality, and accessibility of public sector data, including education (Table 3). Others include the UNESCO Education Data Strategy that emphasizes the improvement of education data availability and quality to support policy-making and monitoring progress toward Sustainable Development Goal 4, offering technical support and standards for education management systems. The World Bank EdStats and Open Data Initiative advances transparency and data-driven decisions through open access to global education indicators and technical support for national systems, while the UN SDG Indicator Framework (Goal 4) requires member states to systematically manage and report education data to enable global tracking and comparability, thereby pushing for more robust national data systems. Similarly, the Global Partnership for Education strengthens education data governance in low-income countries by funding EMIS development and promoting data equity and usability. These efforts collectively aim to build interoperable, secure, and policy-relevant education data ecosystems worldwide.

Table 3: Global data frameworks

Policy / Strategy / Framework	**Provisions Supporting Data Governance in Education**	**Description**
OECD Data Governance Framework	Promotes cross-sectoral data strategies, including privacy, quality, and accessibility standards for public and education data.	OECD encourages countries to adopt coherent data governance approaches that include education data, emphasizing ethical, secure, and interoperable systems.

UNESCO Education Data Strategy	Focuses on improving the availability, quality, and use of education data to support SDG 4 monitoring and policy making.	UNESCO supports countries in developing robust EMIS systems and provides technical guidance on education data standards and governance.
World Bank EdStats and Open Data Initiative	Promotes transparency and public access to global education data, encouraging data-driven decision-making.	The World Bank provides education indicators and technical support for country-level data systems that align with global data governance principles.
UN SDG Indicator Framework (Goal 4)	Mandates member countries to regularly collect, manage, and report education data for global comparability and development tracking.	Drives national governments to build robust data systems for education monitoring, promoting alignment with international data governance practices.
Global Partnership for Education (GPE) Data Strategy	Supports the development of national EMIS frameworks and emphasizes data quality, equity, and usability in education.	GPE provides funding and technical assistance to strengthen national education data governance, especially in low-income countries.

In the region, various regional strategies and initiatives have outlined how African and international institutions are working to strengthen data governance in education across the continent, with a focus on digital integration, harmonized policies, and technical capacity building (Table 4).

Table 4: Regional data frameworks

Policy / Strategy / Framework	**Provisions Supporting Data Governance in Education**	**Description**
African Union's Digital Transformation Strategy (2020–2030)	Emphasizes secure data sharing, digital ID systems, and regional data integration to support digital learning and research.	Guides member states, including Kenya, in adopting harmonized digital infrastructure and data governance policies to advance inclusive education and development.
Continental Education Strategy for Africa (CESA 2016–2025)	Calls for strengthened EMIS systems, data-driven planning, and monitoring to improve educational outcomes and equity.	AU's blueprint for transforming education systems through better use of data for policy, accountability, and performance tracking.
East African Community (EAC) Data Governance Policy Framework (Draft)	Proposes harmonized data protection and governance policies across EAC member states for cross-border education and research collaboration.	Aims to establish common standards for data governance to facilitate digital innovation and educational development across the region.
Smart Africa Alliance – Data Governance Framework	Supports national data strategies and regional interoperability for digital services, including education platforms.	Promotes digital transformation through coordinated data governance, enabling secure and inclusive education services across African nations.

UNESCO CapED Programme in Africa	Assists countries in building capacity for education planning and data governance using evidence-based tools.	UNESCO's Capacity Development for Education (CapED) initiative provides technical support to improve national education data systems in Sub-Saharan Africa.

Kenya has developed a comprehensive suite of national policies, legal instruments, and strategic plans to support data governance within the education sector. These frameworks collectively define the legal, institutional, and operational landscape for how education data is collected, protected, utilized, and shared. They span foundational legislation such as the Data Protection Act (2019) and Basic Education Act (2013), strategic documents like the National Education Sector Strategic Plan and the Digital Economy Blueprint, and operational guidelines including the Kenya School Data Management Guidelines and the draft Education Management Information System Policy (Table 5). Together, these policies establish standards for data security, interoperability, digital inclusion, and evidence-based decision-making, ensuring that education data serves both accountability and developmental objectives within Kenya's broader digital transformation agenda.

Table 5: National (Kenya) data frameworks

Policy / Strategy / Framework	**Provisions Supporting Data Governance in Education**	**Description**
Data Protection Act (2019)	Establishes rights of data subjects, obligations for data controllers and processors, and mandates compliance in handling personal data.	Kenya's foundational legal framework governing all personal data, including learners' and educators' data, is enforced by the Office of the Data Protection Commissioner (ODPC).
Basic Education Act (2013)	Empowers the Ministry of Education to collect and maintain accurate records and data on learners and institutions.	Provides the legal basis for mandatory data collection and reporting in schools; supports monitoring, planning, and allocation of resources.
National Education Sector Strategic Plan (NESSP) 2018–2022	Emphasizes the strengthening of the Education Management Information System (EMIS) and the use of data for evidence-based policy making.	Strategic roadmap guiding education sector reforms, including improved data quality, use, and system integration.
Digital Economy Blueprint (2019)	Advocates for enhanced data infrastructure, interoperability, and trusted digital services across sectors, including education.	Government strategy for a digital transformation agenda calls for secure, open, and ethical use of data, highlighting education as a key sector.
National ICT Policy (2019)	Encourages digital inclusion, data security, and the development of e-government platforms, including EMIS and digital learning systems.	Guides ICT use in public service delivery, providing a foundation for digital tools and secure data handling in the education sector.
Kenya School Data Management Guidelines (MoE, 2018)	Provides standardized procedures for school-level data collection, reporting, and validation through NEMIS.	Operational guide for school heads and administrators on managing learner data securely and accurately within the national education system.

National EMIS Policy (Draft)	Sets standards for education data generation, storage, dissemination, and access across education institutions and levels.	Aims to formalize EMIS operations and ensure consistent, timely, and accurate data collection to support planning and monitoring.
Competency-Based Curriculum (CBC) Implementation Guidelines	Includes data tracking mechanisms for learner progress and curriculum delivery outcomes.	Promotes data use for personalized learning and performance monitoring, requiring secure and real-time data collection from schools.
National Digital Master Plan (2022–2032)	Supports digital infrastructure in education, with emphasis on data systems, security, and interoperability.	The national 10-year plan integrating ICT in all sectors, including education, prioritizes smart learning environments and secure data ecosystems.
Public Finance Management Act (2012)	Mandates the use of reliable data for planning, budgeting, and accountability in public institutions, including schools.	Although primarily financial, it reinforces the importance of credible data in education resource planning and reporting.

3.4. Stakeholder Mapping

The participation of these stakeholders is considered essential in the discussions for the validation and implementation of the data governance framework for education data. Many of them have already developed policies, processes, and best practices around education data sharing that can be used as a reference point for the Education Sector Data Governance Framework (Table 6).

Table 6: Education Data Stakeholders

Stakeholder	**Description**
Ministry of Education (MoE)	Primary policymaker for the education sector; responsible for overall data governance policies, national education strategy, and implementation of EMIS.
State Department for University Education and Research	Oversees universities and research institutions; responsible for policies and data oversight related to higher education and innovation.
State Department for Technical and Vocational Education and Training (TVET)	Manages data policies, standards, and monitoring for TVET institutions; integrates sector-specific datasets into national planning and budgeting.
Commission for University Education (CUE)	Regulates quality assurance, accreditation, and data reporting for universities; manages institutional performance audits and research data.
Teachers Service Commission (TSC)	Responsible for collecting, managing, and utilizing data related to teacher recruitment, deployment, performance, and professional development.
Kenya National Examinations Council (KNEC)	Manages examination and assessment data for basic and tertiary education levels, ensuring data accuracy, integrity, and secure transmission of results.
Kenya Institute of Curriculum Development (KICD)	Develops and reviews curriculum and uses learner performance data to inform curriculum reforms and digital content development.
Kenya Universities and Colleges Central Placement Service (KUCCPS)	Manages student placement into TVETs and universities; collects and analyses enrolment and placement data.
Technical and Vocational Education and Training Authority (TVETA)	Regulates, licenses, and monitors TVET institutions. Collects compliance, performance, and institutional data to improve training quality.
TVET Institutions (Public and Private)	Collect student enrolment, skills training, completion rates, industry linkages, and employment outcome data for national skills development tracking.
Universities (Public and Private)	Generate data on student demographics, research, academic performance, admissions, and institutional quality; report to CUE and MoE.
County Education Offices	Implement education policy at the county level; manage school-level data collection, monitor performance, and support EMIS roll-out.

County Governments	Oversee Early Childhood Development (ECDE), community education, and school infrastructure at the county level. Collect, validate, and use data for localized planning, budgeting, and performance monitoring. Coordinate with MoE on EMIS and public participation in education.
Office of the Data Protection Commissioner (ODPC)	Regulatory authority for personal data in Kenya; oversees compliance with the Data Protection Act within education institutions.
Public and Private Schools	Generate and report student, teacher, and infrastructure data to MoE and TSC through EMIS and other school-based systems.
Development Partners (e.g., UNICEF, UNESCO, World Bank)	Provide funding and technical assistance for education data systems, EMIS strengthening, and policy development.
Educational Technology (EdTech) Companies	Develop digital learning platforms and analytics tools; may collect or process learner and instructional data.
Parents and Students	Primary data subjects whose personal and academic data is collected and used; they also increasingly demand transparency and privacy safeguards.

4. Data Governance Framework for Education

In an increasingly data-driven education ecosystem, Kenya's ability to deliver inclusive, equitable, and quality education is contingent upon how well education data is governed. The exponential growth in data—spanning learners' records, teacher deployment, assessments, digital learning platforms, and institutional performance—necessitates a formalized approach to ensure data is secure, interoperable, ethically used, and policy-relevant. Despite strides in digitization through systems like NEMIS and automated examination portals, the absence of a unified governance structure has led to fragmentation, duplication, and concerns over data privacy and accuracy.

Data Governance Framework

A data governance framework defines how and when decisions are made regarding the collection, management, and use of data. It establishes the structures, roles, and processes necessary to ensure data is handled responsibly and effectively across an organization.

A well-defined data governance framework serves several key purposes, including:

i) Clarifying roles and responsibilities
Clearly defines who is accountable for data-related decisions and processes.

ii) Establishing rules for data use
Sets guidelines for data collection, access, sharing, and retention.

iii) Minimizing risk
Reduces the risks associated with collecting, storing, and using data, such as breaches, misuse, or non-compliance.

iv) Ensuring regulatory compliance
Supports adherence to legal and policy requirements related to data protection and privacy.

v) Enhancing decision-making
Enables better business decisions through trusted, high-quality data.

vi) Improving communication
Promotes transparency and collaboration among data stakeholders.

vii) Increasing the value of data
Helps organizations leverage data as a strategic asset by ensuring it is accurate, accessible, and usable.

4.1. The Five Pillars of Data Governance

Applying the concept of a data governance framework to education data reveals five central pillars that form the foundation of a well-functioning and trustworthy data ecosystem. These pillars ensure that education data is effectively managed, protected, and leveraged to improve learning outcomes and inform policy decisions. They include:

i. Roles and Responsibilities

The first pillar of education data governance in Kenya focuses on defining and institutionalizing roles and responsibilities across all actors involved in the data lifecycle. Key personnel include data stewards, data managers, data entry clerks, school administrators, and ICT officers at institutions, counties, and national bodies such as MoE, TSC, CUE, and KNEC. These professionals are the operational backbone of education data governance. Their responsibilities span the accurate collection, validation, storage, processing, and ethical use

of learner and institutional data. Ensuring that these roles are assigned and coordinated across systems such as NEMIS, KUCCPS, and institutional EMIS platforms is essential for building a trustworthy data ecosystem. Investment in digital literacy and data protection training for these actors is critical. Without a competent and accountable workforce, even the best-designed data governance frameworks will fail to protect data subjects or enable data-driven decisions.

ii. Privacy Standards

The second pillar is the regulatory environment and privacy standards surrounding education data governance. The legal foundation is anchored in Kenya's Data Protection Act (2019), which applies across all educational institutions—public, private, basic, tertiary, and TVET. Ensuring a lawful basis for collecting, processing, storing, and sharing personal education data is a prerequisite for ethical and compliant data practices. Key privacy safeguards include:

(i) Ensuring informed consent for the collection and use of student and staff data;

(ii) Embedding privacy-by-design into EMIS and digital learning systems;

(iii) Respecting the rights of data subjects, such as access, rectification, and erasure;

(iv) Establishing clear data breach notification protocols;

(v) Appointing trained Data Protection Officers at institutional levels;

(vi) Conducting Data Protection Impact Assessments for new digital education systems.

These safeguards help ensure trust and protect learners—especially minors—whose data is increasingly being collected through biometric systems, learning apps, and centralized databases.

iii. Processes and Procedures

The third pillar outlines the standard operating procedures (SOPs) that govern how education data is collected, processed, secured, and shared. These SOPs establish the expected workflows, responsibilities, and checkpoints across the entire data lifecycle. Key areas include:

i) Mapping of data flows between schools, county offices, and national databases;

ii) Definitions of data quality standards (accuracy, timeliness, completeness, consistency);

iii) Procedures for data validation, storage, and backup;

iv) Mechanisms for data access control, version control, and audit trails;

v) Continuous monitoring and improvement protocols to enhance data reliability.

Instituting these procedures ensures operational efficiency, enhances data quality, and strengthens institutional accountability.

iv. Policies and Guidelines

The fourth pillar involves creating and harmonizing policies and guidelines that govern how education data is managed, accessed, and shared. These policies establish the operational framework within which public institutions, private schools, TVETs, EdTech firms, and universities interact with education data. They also communicate the government's commitment to treating education data as a strategic national asset. Relevant policy tools include:

(i) An internal data governance policy for ministries and institutions;

(ii) An external (open) education data sharing policy, specifying conditions for third-party access;

(iii) Data protection and information security guidelines, including minimum cybersecurity requirements for school IT systems;

(iv) A data retention and disposal policy to manage the lifecycle of learner records.

While some of these elements exist in draft or fragmented form, a consolidated framework under the Ministry of Education is needed to standardize compliance and foster data use in planning, budgeting, and quality assurance.

v. Tools and Practices

This pillar focuses on the technical instruments that support the implementation of the policies in Pillar 4. These tools enable institutions and data custodians to apply consistent data governance practices. They are grouped into two categories:

(a) Tools for Data Protection:

i) Templates for anonymization and pseudonymization of sensitive student data;
ii) Guidance on data minimization for surveys and assessments;
iii) Standardized informed consent forms for parents and adult learners;
iv) A framework for conducting institutional Data Protection Impact Assessments.

(b) Tools for Data Sharing and Access:

i) A standardized code of conduct for data-sharing between ministries, research organizations, and EdTech firms;
ii) Templates for data sharing agreements between institutions and third parties;
iii) Tools for managing data inventories and access logs;
iv) A data license registry outlining terms of use for external stakeholders accessing public education datasets.

These tools promote transparency, accountability, and structured data sharing while reducing legal and ethical risks.

4.2. Operationalizing the Five Pillars

To evaluate the current state of data governance in Kenya's education sector, the five foundational pillars—Roles and Responsibilities, Privacy Standards, Policies, Tools and Practices, and Processes and Procedures—were assessed using six key operational dimensions. These dimensions provide a measurable lens through which the maturity, gaps, and strengths of existing governance structures can be understood:

i) **Data Quality**: Evaluates the accuracy, completeness, consistency, and reliability of education datasets generated at the school, county, and national levels. Poor data quality undermines the effectiveness of policy decisions and planning.

ii) **Data Security**: Measures the presence and adequacy of technical and organizational safeguards to protect personal and institutional data from unauthorized access, loss, or breaches. This includes encryption, secure storage, and access controls aligned with the Data Protection Act.

iii) **Data Accessibility**: Assesses how easily stakeholders—including school heads, county officers, researchers, and policymakers—can access relevant data when

needed. It also considers issues of equity in access across urban and marginalized regions.

iv) **Data Timeliness**: Examines the frequency, punctuality, and update cycles of data reporting. Timely data is essential for real-time decision-making, budget planning, and emergency response.

v) **Data Stewardship**: Evaluates whether roles such as data stewards, custodians, and protection officers are designated and active at institutional and ministerial levels. This also includes clarity in accountability and oversight structures.

vi) **Metadata Documentation and Organization**: Focuses on how well datasets are described, catalogued, and managed. Strong metadata practices enable better understanding, interoperability, and re-use of education data across systems.

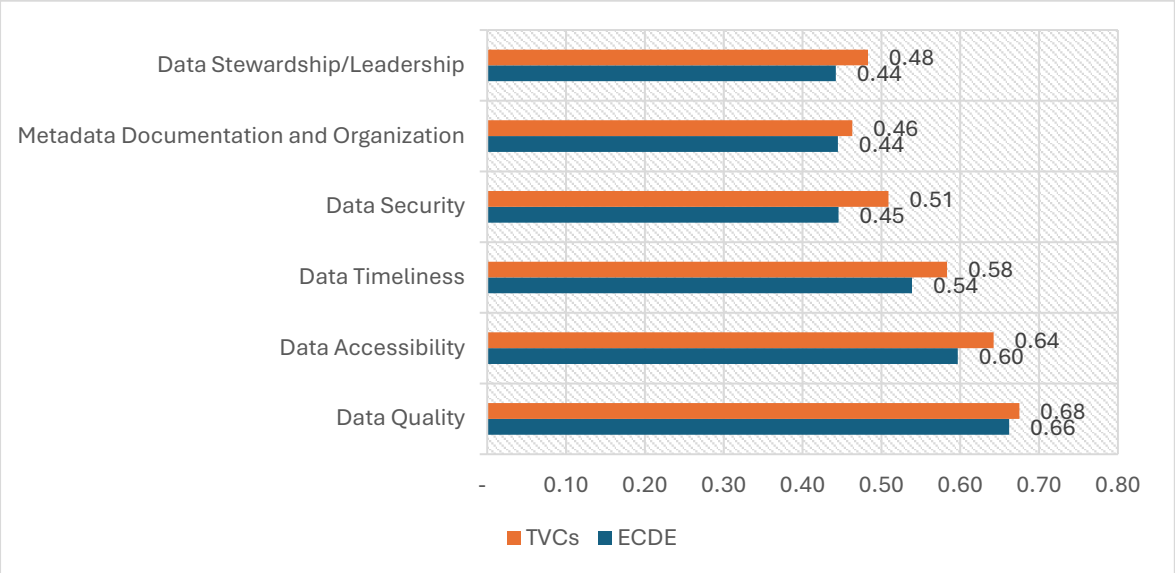
Each of these assessment criteria aligns directly with the objectives of the five pillars and offers a structured approach to benchmarking the current state of education data governance. Through this framework, it becomes possible to identify both systemic strengths and critical weaknesses.

4.3. Data Governance Index (DGI) for the Education Sector in Kenya

County level

The Data Governance Index scores indicate that county-managed Technical and Vocational Centers (TVCs) in Kenya outperform Early Childhood Development and Education (ECDE) programs across all six dimensions, though the margins are narrow (Figure 1). Data quality and accessibility are the strongest areas for both, largely due to counties adopting digital data collection tools such as NEMIS (National Education Management Information System), especially in TVCs. In counties like Makueni and Kakamega, structured reporting templates and digital records have improved the consistency and usability of data. Timeliness is moderately better in TVCs, where regular term-based assessments and county audits ensure more current data. In contrast, many ECDE programs rely on manual records, with limited schedules for updates, as seen in counties like Turkana or Marsabit.

Figure 1: Data Governance Index at County Level



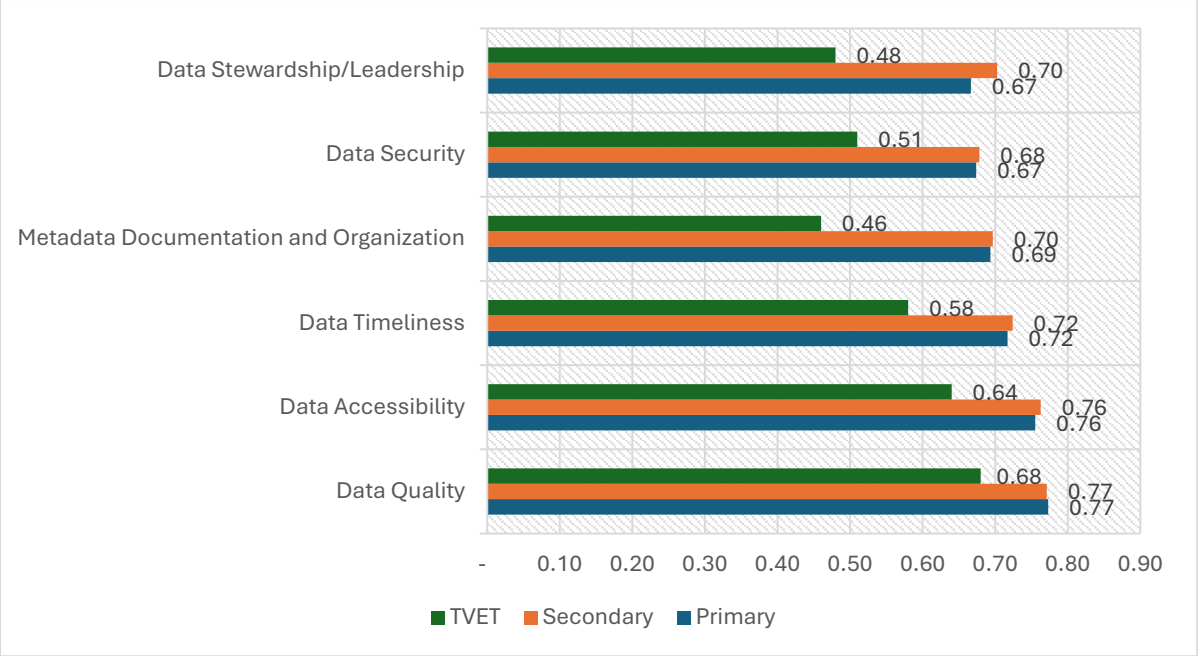
Both sectors score poorly in metadata documentation, data security, and stewardship. Counties vary significantly in how they define and document datasets, leading to inconsistencies. For example, counties like Kiambu and Kisumu have begun standardizing metadata through county data portals, while others lack any formal classification systems. On security, most counties lack clear data protection policies or access controls, exposing both ECDE and TVC data to risks. Leadership in data governance also remains weak, with only a few counties—such as Nyeri and Elgeyo Marakwet—having designated data officers or units within the education departments. These findings underscore the need for a national framework guiding county-level practices, focused on strengthening metadata systems, securing infrastructure, and institutionalizing data stewardship roles to ensure reliable, timely, and secure education data governance across Kenya.

National Level

At the national level, the analysis revealed a clear disparity in data governance performance between national government-managed primary and secondary education institutions and TVET (Technical and Vocational Education and Training) institutions (Figure 2). Both primary and secondary sectors score highly and almost identically across all six dimensions, indicating that national-level systems—particularly those supported through NEMIS—have matured significantly. The consistent scores of 0.77 in data quality and 0.76 in data accessibility show that data in these sectors is not only reliable but also readily available for planning and accountability. The strong performance in data timeliness (0.72) reflects regular updates and efficient reporting cycles enforced by national standards.

In contrast, TVET institutions managed by national agencies score significantly lower in all dimensions, with particularly wide gaps in metadata documentation (0.46) and data stewardship (0.48). These gaps suggest a lack of standardized metadata practices and unclear governance roles, which are likely consequences of fragmented systems and the absence of a unified data management platform. While NEMIS provides comprehensive support for basic education, TVET relies more on sector-specific databases or manual tracking systems, which may not be uniformly implemented or maintained across institutions. This affects data reliability, consistency, and the capacity to support policy decisions or funding allocation efficiently.

Figure 2a: Data Governance Index at the National Level

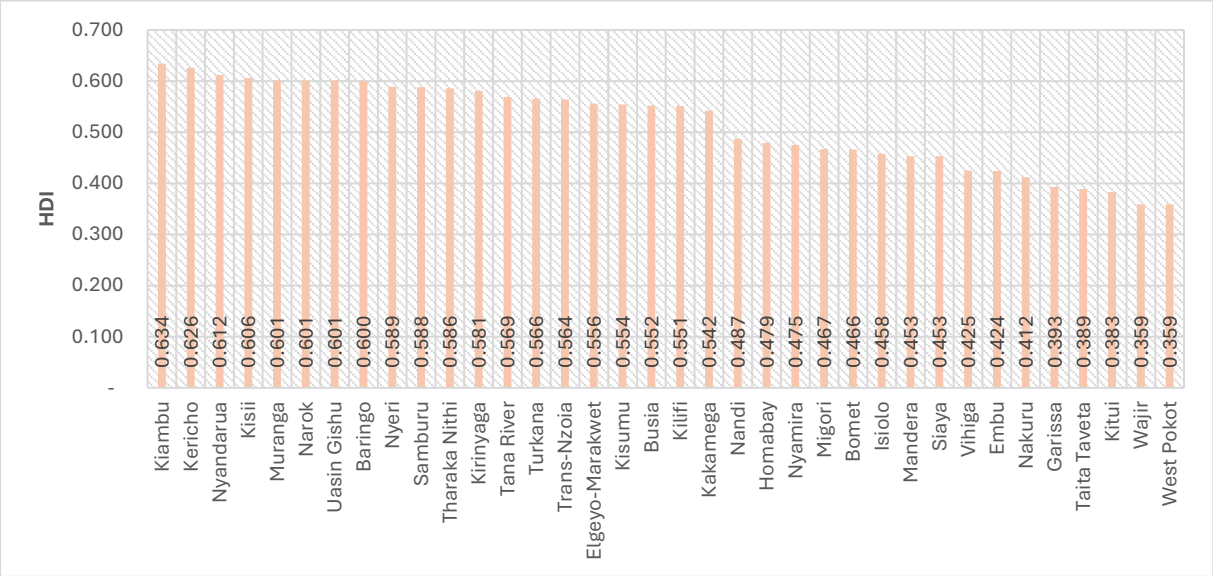


The lower data security score (0.51) for TVET institutions is also notable and raises concerns about vulnerabilities in how data is protected, especially in institutions with limited ICT capacity. Similarly, weak data stewardship signals that leadership structures for managing education data in the TVET sector remain underdeveloped. These findings emphasize the need for national government intervention through the TVET Authority (TVETA) and the Ministry of Education to invest in system integration, define clear governance roles, and develop a national data platform tailored to post-secondary training institutions. Aligning TVET data governance with the standards applied in basic education would significantly strengthen accountability, planning, and service delivery in Kenya's broader education system.

4.4. Human Development Index

The Human Development Index (HDI) for Kenya's counties was developed using a localized adaptation of the global HDI methodology pioneered by the United Nations Development Programme (UNDP). It incorporates county-level data to reflect sub-national disparities in human development more accurately. Specifically, it measures three core dimensions: long and healthy life, assessed through life expectancy at birth; knowledge, gauged using a composite of expected years of schooling for children and mean years of schooling for adults; and standard of living, represented by county-level Gross County Product (GCP) per capita, adjusted using inequality-sensitive methods. This county-focused HDI helps monitor progress toward both national development goals and global commitments such as the Sustainable Development Goals (SDGs). The HDI data in Figure 2b highlights regional disparities and provides a foundation for evidence-based policy interventions aimed at reducing inequality and enhancing human well-being at the local level.

Figure 2b: Human Development Index for Kenya



4.5. Data Quality and Human Development Index

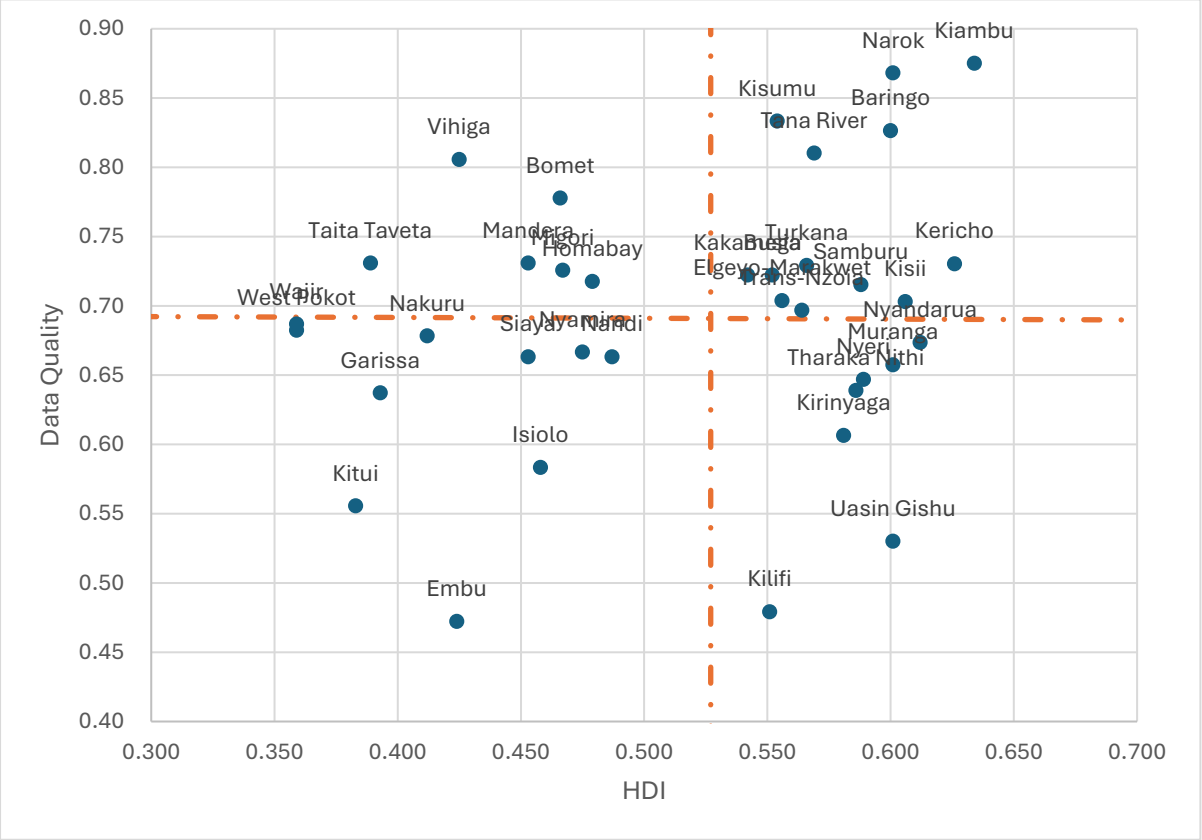
Data quality plays a vital role in advancing human development, as accurate, reliable, and timely data enables effective planning, resource allocation, and monitoring of services that directly impact HDI indicators such as education, health, and income. Figure 3 below shows a positive correlation between the Data Quality dimension and the Human Development Index (HDI) across Kenyan counties. Counties with high HDI values, such as Kiambu (HDI 0.634, Data Quality 0.88) and Narok (HDI 0.601, Data Quality 0.87), also report very high data quality scores. These counties have invested in stronger institutional frameworks and technological tools that enable accurate, consistent, and actionable data. Such data quality enhances planning for healthcare, education, and infrastructure, contributing to better living standards — the very pillars of HDI. Conversely, counties like Embu (HDI 0.424, Data Quality 0.47) and Kilifi (HDI 0.551, Data Quality 0.48) show lower data quality, suggesting weak systems that may hinder evidence-based service delivery and, in turn, limit development outcomes.

At the national level, Kenya's average HDI stands at 0.527 with an average data quality score of 0.69, reflecting moderate alignment between development and data systems. Counties like Kericho, Bomet, and Tana River perform above this average in both indicators, indicating that sub-national governments can indeed surpass national standards when supported by effective leadership and resource mobilization. Interestingly, counties such as Vihiga (HDI 0.425, Data Quality 0.81) show that strong data systems may exist even where overall development is lower — possibly due to targeted support from NGOs or donor-funded programs enhancing data management without yet translating into overall human development gains.

The variations also highlight key national and county-level policy implications. Nationally, tools like NEMIS and government efforts to standardize data practices in education have benefited counties with better digital access and stronger governance. However, counties with low HDI and high data quality (e.g., Vihiga and Taita Taveta) reveal that data investments alone are not enough — they must be coupled with capacity-building, infrastructure, and equitable public investment to yield broader development results. Therefore, enhancing data quality remains essential not just for monitoring development but also as a driver of HDI

improvements, especially when paired with responsive leadership and integrated development planning.

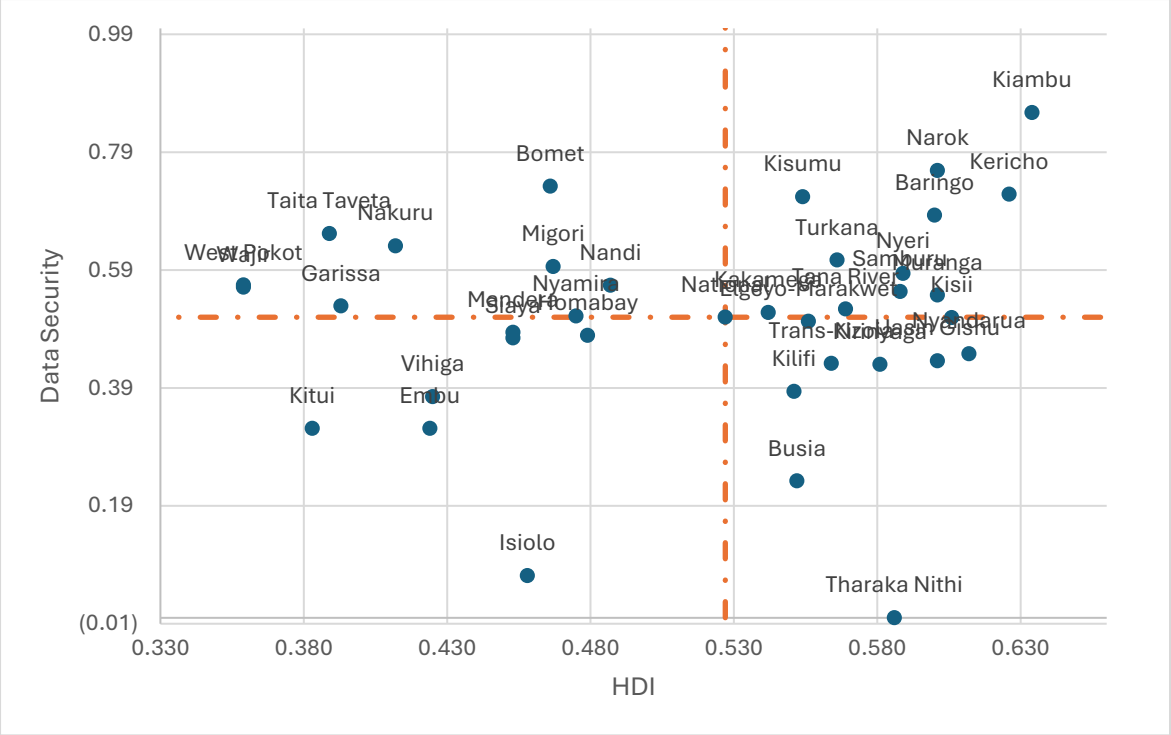
Figure 3: Data Quality and Human Development Index



Data Security and Human Development Index

Data security is a critical enabler of human development, as it ensures the integrity, confidentiality, and trustworthiness of information used for planning and service delivery. A secure data environment strengthens public confidence, supports compliance with legal frameworks like Kenya's Data Protection Act (2019), and directly influences the quality of education, health, and governance services that shape the Human Development Index (HDI). The findings indicate significant disparities in data security practices across counties in Kenya, with most counties performing below or just at the national average score of 0.51. High-performing counties like Kiambu (0.86), Narok (0.76), and Kericho (0.72) exhibit strong adherence to key data security standards. These counties are likely to have implemented encryption, access control, routine security audits, and backup systems, as well as compliance with the Data Protection Act of 2019. In counties such as Kiambu, enhanced investments in ICT and governance capacity may explain their robust training programs for staff, proactive incident response to breaches, and updated security protocols. Bomet (0.73) and Kisumu (0.71) similarly reflect counties where leadership prioritizes secure digital systems, likely due to external donor support or integrated county ICT strategies.

Figure 4: Data Security and Human Development Index

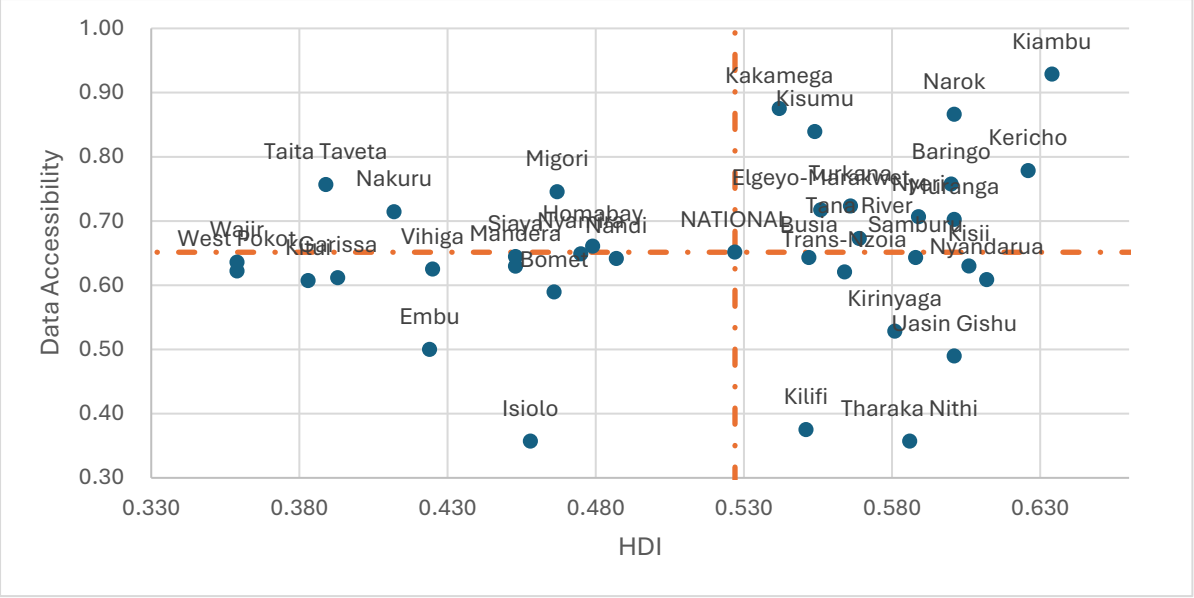


Conversely, counties such as Isiolo (0.07), Busia (0.23), and Embu (0.32) score extremely low, signaling major vulnerabilities. These counties likely lack formalized data security structures — including encryption, regular training, or breach mitigation processes — exposing sensitive information to risk. In places like Isiolo, this may be due to limited infrastructure, low staffing capacity, or inadequate awareness of data protection laws. Even counties with mid-level HDI scores like Kilifi (HDI 0.551, Security 0.38) and Vihiga (HDI 0.425, Security 0.38) show weak enforcement of basic security measures. Overall, these findings highlight the urgent need for national coordination and capacity-building efforts at the county level to improve data security compliance, training, and resilience, especially in underserved or high-risk regions.

Data Accessibility and Human Development Index

Data accessibility is a key driver of human development, as it enables timely and inclusive decision-making by ensuring that stakeholders can access, interpret, and utilize data effectively to improve education, health, and economic outcomes. The study revealed that there is a generally positive relationship between data accessibility and HDI across Kenyan counties. High-performing counties such as Kiambu (HDI 0.634, Accessibility 0.93), Narok (HDI 0.601, Accessibility 0.87), and Kakamega (HDI 0.542, Accessibility 0.88) have invested in robust ICT infrastructure and user-friendly data systems, allowing both internal and external stakeholders to access updated information easily. These counties likely maintain digital repositories, provide regular training, and reduce legal or technical barriers to access, which enhances evidence-based planning. Kisumu (HDI 0.554, Accessibility 0.84) and Kericho (HDI 0.626, Accessibility 0.78) also show strong alignment, suggesting effective integration of digital tools and stakeholder support mechanisms that reinforce inclusive governance.

Figure 5: Data Accessibility and Human Development Index

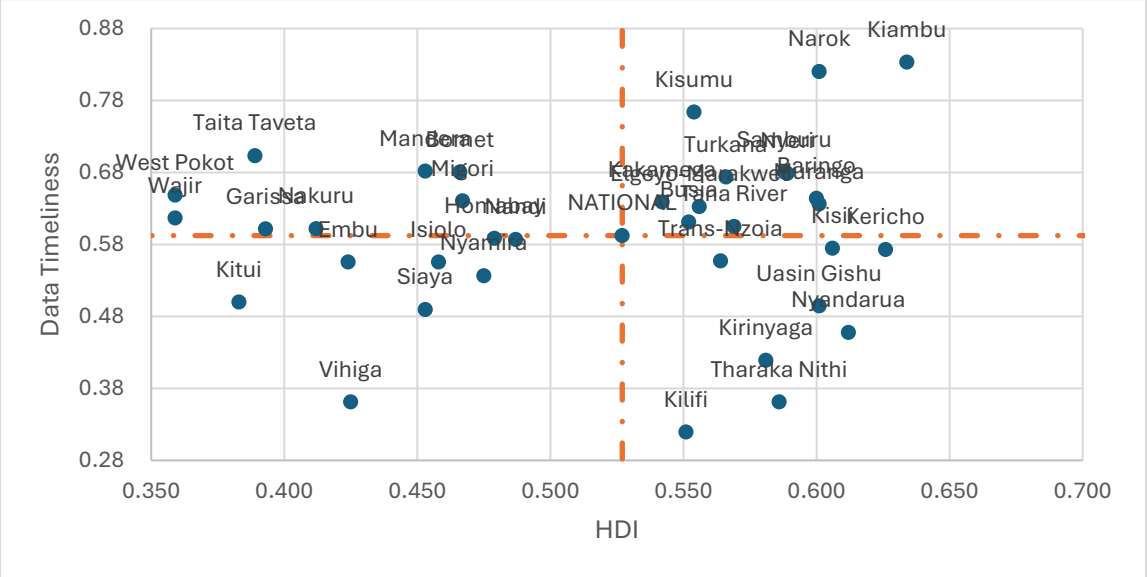


In contrast, counties like Isiolo (HDI 0.458, Accessibility 0.36), Kilifi (HDI 0.551, Accessibility 0.38), and Tharaka Nithi (HDI 0.586, Accessibility 0.36) report low accessibility scores despite moderate HDI levels, indicating possible challenges such as underdeveloped ICT infrastructure, lack of training, or restrictive data access policies. Similarly, Embu (HDI 0.424, Accessibility 0.50) and Uasin Gishu (HDI 0.601, Accessibility 0.49) reflect inconsistencies where development gains are not matched by data openness, potentially limiting their ability to sustain or scale progress. The national average of 0.65 suggests that while most counties are performing moderately well, there is still a critical need for policy harmonization and investment in systems that enhance equitable data access across all levels of government.

Data Timeliness and Human Development Index

Data timeliness, as a key component of data governance, plays a critical role in supporting responsive decision-making and public service delivery, particularly in the education sector. Timely data ensures that interventions align with current needs and reduces inefficiencies caused by outdated information. This dimension is positively associated with the Human Development Index (HDI), as counties with higher scores often exhibit better infrastructure, staffing, and coordination mechanisms for routine reporting. Counties such as Kiambu (HDI 0.634, Timeliness 0.83) and Narok (HDI 0.601, Timeliness 0.82) demonstrate how strong local systems—backed by investment in digital tools, dedicated personnel, and active partnerships—can maintain high data currency. Their performance suggests the presence of structured timelines for data submission, digital education dashboards, and responsive audit cycles.

Figure 6: Data Timeliness and Human Development Index



On the other end of the spectrum, counties like Kilifi (HDI 0.551, Timeliness 0.32), Tharaka Nithi (HDI 0.586, Timeliness 0.36), and Vihiga (HDI 0.425, Timeliness 0.36) reflect weaker systems likely characterized by staffing shortages, limited audit frequency, and low technology adoption, contributing to data delays. Counties such as Mandera (0.68), Bomet (0.68), and Taita Taveta (0.70), despite moderate HDI scores, have relatively strong data timeliness—this may be due to external collaboration with development partners or proactive county leadership introducing incentives for reporting and using mobile data collection tools. The national average of 0.59 suggests room for improvement across most counties. Enhancing mechanisms such as user feedback loops, real-time dashboards, and structured periodic audits will be essential to boost timeliness and, in turn, improve the planning and delivery of critical development services.

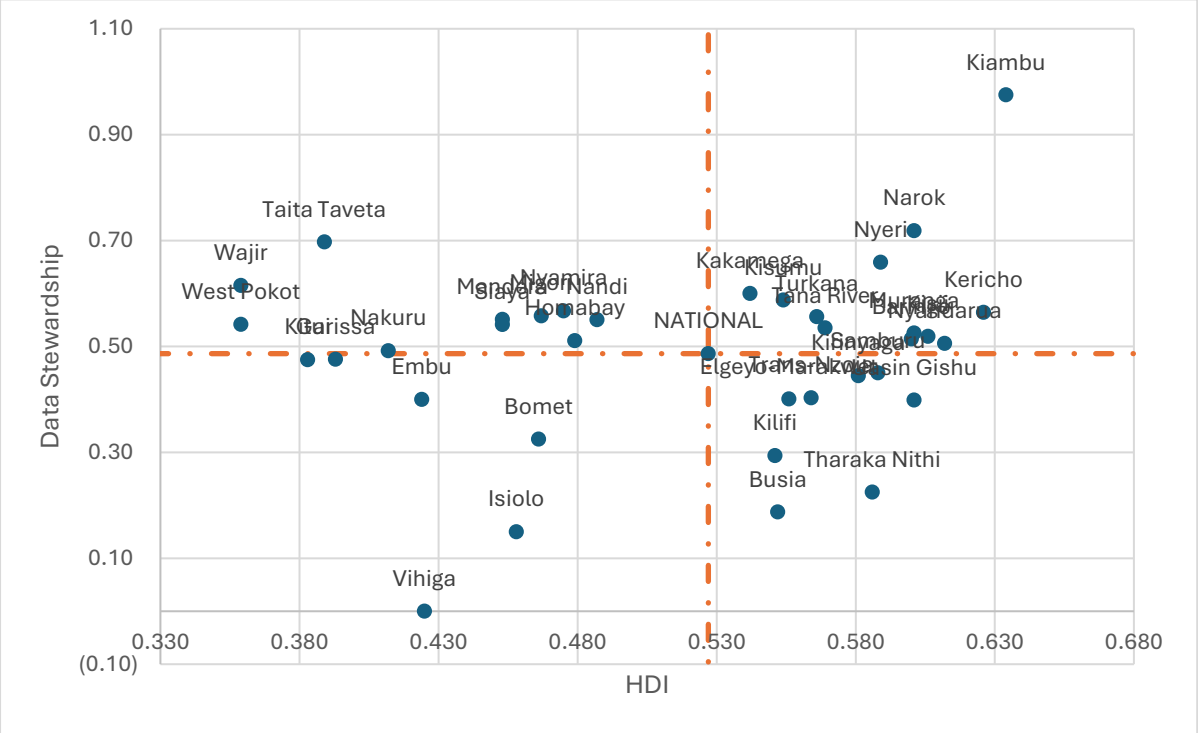
Data Stewardship/Leadership and Human Development Index

Strong data stewardship and leadership are essential for driving human development, as they ensure data is collected, managed, and used with transparency, accuracy, and accountability. Analysis of county-level data in Kenya shows a clear positive relationship between well-established data stewardship practices and higher Human Development Index (HDI) scores. Counties such as Kiambu (HDI 0.634, Stewardship 0.98), Narok (0.601, 0.72), and Nyeri (0.589, 0.66) have demonstrated strong governance by establishing clear roles, regular staff training, and functional education data management committees. These counties use reliable, timely data to inform policies and resource allocation, contributing to better outcomes in health, education, and income.

In contrast, counties like Isiolo (HDI 0.458, Stewardship 0.15), Busia (0.552, 0.19), and Tharaka Nithi (0.586, 0.23) reveal major weaknesses in data leadership. These gaps are often due to undefined responsibilities, limited funding, and lack of operational frameworks, which hamper the ability to use data effectively for development planning. While examples like Wajir (HDI 0.359, Stewardship 0.62) show that targeted support can improve data practices even in low-HDI contexts, the national average of 0.49 suggests a systemic need for investment in

stewardship capacity. Strengthening data leadership through policy frameworks, funding, and institutional roles is critical for improving development outcomes across all counties.

Figure 7: Data Stewardship/Leadership and Human Development Index



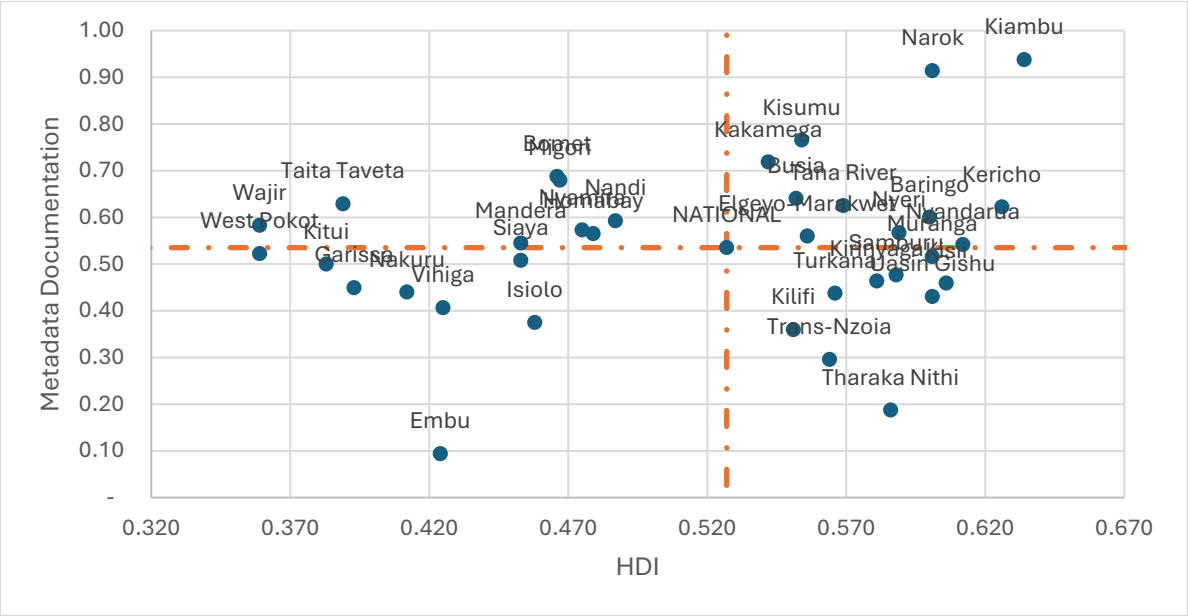
Metadata Documentation and Organization and Human Development Index

Metadata documentation and organization are foundational to high-quality data governance, as they enable clarity, consistency, and usability across all levels of the education system. Well-maintained metadata ensures users understand the meaning, structure, and source of data fields, supporting accurate analysis and policy decisions that directly impact development outcomes. In Kenya, counties with strong performance in this area—such as Kiambu (HDI 0.634, Metadata 0.94), Narok (HDI 0.601, 0.91), and Kisumu (HDI 0.554, 0.77)—are likely benefiting from the use of standardized metadata frameworks, training of education staff, and established protocols for updating and using metadata. These counties may also employ mechanisms for user feedback, ensuring that metadata practices evolve to meet the needs of various stakeholders. Their metadata coverage likely includes comprehensive documentation of data collection methods, sources, and definitions, contributing to more reliable education data and ultimately better human development outcomes.

Conversely, counties like Embu (HDI 0.424, Metadata 0.09), Tharaka Nithi (HDI 0.586, 0.19), and Trans-Nzoia (HDI 0.564, 0.30) reflect critical weaknesses in metadata systems. These counties may lack documented metadata standards, formal training for data users, or clear guidance for retrieving and interpreting data—leading to inconsistent, error-prone, or underutilized datasets. Nationally, Kenya's average metadata documentation score of 0.54 suggests that many counties operate without fully established systems, despite efforts by the Ministry of Education to promote structured data practices. While the National Education Management Information System (NEMIS) offers metadata support for primary and secondary schools, its reach and application in county-level TVET and ECDE data systems remain limited. Strengthening metadata organization through the adoption of national metadata standards

(such as ISO 11179), improving staff training, and enhancing system interoperability will be key to building a unified and actionable education data landscape that supports sustainable development across all counties.

Figure 8: Metadata Documentation and Organization and Human Development Index



Data Governance Index and Human Development Index

The correlation analysis in Table 7 reveals a positive but modest relationship between HDI and all six dimensions of data governance, suggesting that improvements in data governance are associated with higher levels of human development across counties. The observed correlations between data governance dimensions and HDI align with existing literature emphasizing the critical role of data systems in advancing human development outcomes, particularly in low- and middle-income countries. Scholars such as Kitchin (2014) and Jerven (2013) argue that high-quality, timely, and accessible data are foundational for effective public service delivery, policy formulation, and monitoring of socio-economic development. The relatively stronger correlation between data quality (r = 0.24) and HDI echoes findings by UNICEF (2017) and World Bank Group (2017), which underscore that robust Education Management Information Systems (EMIS) directly enhance the capacity of governments to target educational interventions and allocate resources equitably, thereby improving health, education, and income outcomes reflected in HDI scores.

Moreover, metadata documentation and accessibility, which also show notable associations with HDI, have been identified as key enablers of transparency, accountability, and evidence-based planning (World Bank, 2014; OECD, 2021). These dimensions facilitate interoperability and comparative analysis across regions, allowing sub-national governments to benchmark performance and replicate best practices. On the other hand, the weaker correlation observed for data timeliness (r = 0.08) is consistent with prior studies (e.g., Mo Ibrahim Foundation, 2019), which suggest that while timely data is crucial for real-time decision-making, its impact on long-term development outcomes may be mediated by institutional readiness and the presence of effective feedback loops.

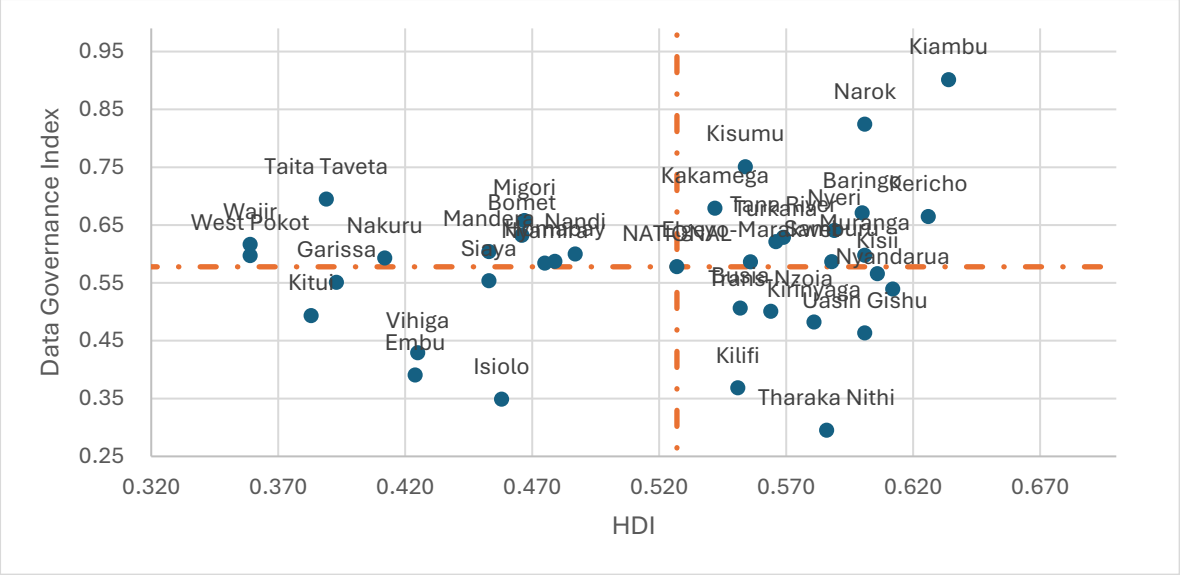
Table 7: Pearson Correlations Between Data Governance Dimensions and HDI

	1.	2.	3.	4.	5.	6.	7.	8.
1. HDI	1.00							
2. Data Quality	0.24	1.00						
3. Data Security	0.14	0.61	1.00					
4. Data Accessibility	0.19	0.75	0.80	1.00				
5. Data Timeliness	0.08	0.62	0.70	0.75	1.00			
6. Data Stewardship/Leadership	0.15	0.37	0.74	0.71	0.68	1.00		
7. Metadata Documentation and Organization	0.20	0.73	0.71	0.79	0.69	0.62	1.00	
8. Overall-Data Governance Index	0.19	0.75	0.90	0.92	0.85	0.83	0.88	1.00

On the other hand, the strong internal correlations among the governance dimensions themselves—such as between data security and data accessibility (r = 0.80)—reflect the integrated nature of data systems, where improvements in one area often reinforce others. This supports the integrated governance approach promoted by frameworks such as the OECD Data Governance Model (2021) and the African Union Digital Transformation Strategy (2020–2030), which advocate for systemic, coordinated data reforms to drive inclusive development. Thus, the empirical patterns observed in the Kenyan context are firmly rooted in and supported by global research on data governance and human development linkages.

Counties with stronger data governance frameworks tended to achieve higher levels of development. For instance, counties like Kiambu (HDI 0.634, DGI 0.90), Narok (HDI 0.601, DGI 0.82), and Kisumu (HDI 0.554, DGI 0.75) have some of the highest scores in both metrics. These counties have likely invested in integrated data systems, capacity-building for staff, robust stewardship mechanisms, and routine data audits. As a result, their ability to generate accurate, timely, and actionable data contributes directly to improved planning, transparency, and service delivery in education, health, and other human development sectors.

Figure 9: Data Governance Index and Human Development Index



Conversely, counties like Tharaka Nithi (HDI 0.586, DGI 0.29), Isiolo (HDI 0.458, DGI 0.35), and Kilifi (HDI 0.551, DGI 0.37) illustrate how weak data governance structures may constrain development efforts. These counties may lack standardized metadata, face irregular data reporting, or operate with minimal oversight and accountability in education data systems. The national average DGI of 0.58 indicates moderate performance overall, but also highlights the need for targeted policy interventions to bridge gaps in lower-performing counties. Notably, counties show that even in regions with lower HDI, strong governance practices can emerge—often due to focused donor support or local leadership reforms. This divergence between relatively strong data governance and lower HDI scores in certain counties invites closer examination. For instance, counties like Taita Taveta (HDI 0.389, DGI 0.69) and West Pokot (HDI 0.359, DGI 0.60) demonstrate that robust data governance practices can exist even in regions with lower overall development outcomes. This may be attributable to targeted donor interventions, pilot initiatives, or progressive county-level leadership that prioritize digital infrastructure and data reforms regardless of broader socio-economic challenges. However, the absence of a strong correlation in these cases suggests that improved data systems alone may not be sufficient to immediately shift HDI metrics, which are influenced by long-term structural factors like poverty, infrastructure, and health access. It would therefore be valuable to track the HDI trajectories of such counties over time to determine whether sustained data governance efforts eventually lead to corresponding gains in human development. This could provide critical insights into lag effects and help refine policy timelines and expectations for reform impact.

Strengthening county-level data governance across all dimensions—quality, security, timeliness, accessibility, stewardship, and metadata—will be essential for achieving equitable and sustainable development in Kenya.

5. Conclusion and Recommendations

5.1. Conclusion

The analysis of Kenya's Data Governance Index reveals a strong correlation between robust data governance practices and improved Human Development Index (HDI) outcomes. Counties with well-structured data systems—characterized by high data quality, accessibility, timeliness, security, metadata documentation, and stewardship—demonstrate stronger planning capabilities, more efficient service delivery, and better outcomes in health, education, and economic sectors. High-performing counties like Kiambu, Narok, and Kisumu exemplify how institutionalized data leadership, regular audits, digital tools, and user engagement can collectively drive meaningful development. Conversely, counties with weaker data governance often struggle with data inconsistencies, limited capacity, and poor coordination, which hamper evidence-based decision-making and equitable progress.

To achieve sustained national development, Kenya must prioritize investment in comprehensive data governance reforms across all counties. This includes adopting national standards, strengthening institutional roles, ensuring consistent training and funding for data activities, and enforcing compliance with existing laws such as the Data Protection Act (2019). A harmonized, inclusive, and well-resourced data governance ecosystem will not only improve transparency and accountability but also empower counties to better respond to citizens' needs. Strengthening these foundations is essential for closing development gaps and ensuring that all counties, regardless of starting point, have the tools and systems necessary to contribute to Kenya's Vision 2030 and the Sustainable Development Goals. A summary of data governance issues and recommendations is provided in Table 2.

5.2. Data Governance Issues and Recommendations

To strengthen Kenya's education data ecosystem, it is essential to address the persistent challenges that undermine effective data governance across counties and institutions. These issues span multiple domains—including inconsistent data quality, poor data security practices, limited access to data, delays in reporting, weak leadership structures, and inadequate metadata documentation. Each of these governance gaps hinders the country's ability to make evidence-based decisions, ensure accountability, and deliver equitable services in the education sector. Based on the analysis above, the basic issues and recommendations for each of the pillars can be summarized as in Table 8. These proposals aim to improve data reliability, safeguard privacy, enhance accessibility, and promote institutional leadership, while assigning clear roles to responsible actors at both national and county levels. This supports the development of a cohesive and responsive data governance framework that aligns with Kenya's national development goals and global best practices.

Table 8: Data Governance Issues and Recommendations

Data Governance Pillar	Issues	Recommendations	Roles
Data Quality	- Inconsistent data validation and cleaning in low-DGI counties - Manual data collection limits reliability - Variability in	- Implement mandatory data quality checks and validation protocols - Automate collection processes to reduce manual errors - Provide training on	County Education Directors, MoE Quality Assurance Unit, ICT Officers

	completeness and consistency across counties	data accuracy and integrity	
Data Security	- Poor encryption and data protection practices in low-performing counties (e.g., Isiolo, Busia) - Inadequate training on security protocols - Minimal compliance with the Data Protection Act (2012)	- Enforce encryption, access controls, and multi-factor authentication - Provide regular staff training on cybersecurity - Monitor and report compliance with data protection laws	Office of the Data Protection Commissioner (ODPC), MoE Legal Units, County ICT Teams
Data Accessibility	- Weak digital infrastructure in several counties - Limited user access and transparency - Lack of open data platforms and feedback mechanisms	- Develop open-access platforms for education data - Standardize accessibility policies across counties - Create user feedback channels for improvement	MoE ICT Directorate, County Governments, Civil Society & Researchers
Data Timeliness	- Delays in reporting cycles - Insufficient staff and technological tools for real-time updates - Irregular audit mechanisms for data currency	- Enforce routine data submission timelines - Invest in mobile data collection tools and dashboards - Schedule periodic data audits	MoE Planning Unit, County ICT Units, Sub-County Education Officers
Data Stewardship/Leadership	- Undefined roles and responsibilities in counties - Lack of dedicated data teams and committees - Inconsistent funding for data leadership and capacity building	- Institutionalize stewardship roles and guidelines - Form cross-sectoral education data committees - Allocate dedicated budget lines for data leadership activities	County Directors of Education, MoE EMIS Directorate, County Executive Committees

Metadata Documentation and Organization	- Incomplete or outdated metadata in counties - Lack of standard documentation guidelines and tools - Limited staff awareness and technical capacity	- Adopt national metadata standards - Train staff on metadata use and documentation - Establish centralized metadata repositories per county	KNBS, MoE EMIS Officers, County Data Stewards

References

African Union. (2020). *Digital Transformation Strategy for Africa (2020–2030)*. African Union Commission. Retrieved from https://au.int/sites/default/files/documents/38507-doc-dts-english.pdf

Barrera-Osorio, F., & Linden, L. (2009). *The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia*. World Bank Group.

Becker, G. S. (2009). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. University of Chicago Press.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). *Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood*. American Economic Review, 104(9), 2633–2679.

Hanushek, E. A., & Woessmann, L. (202215). *The Knowledge Capital of Nations: Education and the Economics of Growth*. MIT Press.

Heckman, J. J. (2008). *Schools, Skills, and Synapses*. Economic Inquiry, 46(3), 289–324.

Jerven, M. (2013). *Poor Numbers: How We Are Misled by African Development Statistics and What to Do About It*. Cornell University Press.

Kenya National Bureau of Statistics. (2024). Statistical Abstracts. https://www.knbs.or.ke/wp-content/uploads/2024/12/2024-Statistical-Abstract1.pdf

Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. SAGE Publications.

Materu, P. (2007). Higher education quality assurance in sub-Saharan Africa: Status, challenges, opportunities, and promising practices. World Bank. https://openknowledge.worldbank.org/server/api/core/bitstreams/e9904501-5711-52ae-86b7-b860bb730adb/content

Mo Ibrahim Foundation. (2019). *Assessing Africa's Data Landscape*. Retrieved from https://mo.ibrahim.foundation/sites/default/files/2019-11/African_Governance_Report_2019_EN.pdf

OECD. (2022). Going Digital to Advance Data governance for growth and well-being: Data governance for data-driven innovation. https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/12/going-digital-to-advance-data-governance-for-growth-and-well-being_246d8cab/e3d783b0-en.pdf

OECD. (2021). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-Use Across Societies*. OECD Publishing. https://doi.org/10.1787/276aaca8-en

Raudonyte, I., & Foimapafisi, T. (2022). Using learning assessment data for educational planning in Sub-Saharan Africa: a comparative analysis. UNESCO International Institute for Educational Planning, Paris.

Schultz, T. W. (1961). *Investment in Human Capital*. American Economic Review, 51(1), 1–17.

Tchamyou, V. S., Asongu, S. A., & Odhiambo, N. M. (2019). The role of ICT in modulating the effect of education and lifelong learning on income inequality and economic growth in Africa. *African Development Review*, *31*(3), 261-274. https://doi.org/10.1111/1467-8268.12388

UNICEF. (2017). *Data for Children Strategic Framework*. United Nations Children's Fund. https://data.unicef.org/wp-content/uploads/2017/04/Data-for-Children-Strategic-Framework-UNICEF.pdf

Wamutoro, M., Kessio, D. K., & Wambua, B. K. (2022). Effectiveness of EMIS for student information management on management of public secondary schools in Uasin Gishu County, Kenya. *Reviewed Journal International of Business Management*, *3*(1), 122-133. https://doi.org/10.36348/jaep.2022.v06i11.002

World Bank Group. (2017). Data for Learning: Building a Smart Education Data Ecosystem. *Directions in Development- Human Development.* https://openknowledge.worldbank.org/bitstreams/7661f023-72c7-5ece-be00-75921a193562/download

World Bank. (2014). *Open Data for Resilience Initiative: Field Guide*. The World Bank. https://openknowledge.worldbank.org/server/api/core/bitstreams/81009d98-8871-54be-98c1-be605001a7b6/content

World Bank. (2018). *World Development Report 2018: Learning to Realize Education's Promise*. Washington, DC: World Bank

World Economic Forum. (2023). How Africa's youth will drive global growth. https://www.weforum.org/stories/2023/08/africa-youth-global-growth-digital-economy/

Mission

To strengthen local capacity for conducting independent,
rigorous inquiry into the problems facing the management of economies in sub-Saharan Africa.

The mission rests on two basic premises: that development is more likely to occur where there is sustained sound management of the economy, and that such management is more likely to happen where there is an active, well-informed group of locally based professional economists to conduct policy-relevant research.

Bringing Rigour and Evidence to Economic Policy Making in Africa

- Improve quality.
- Ensure Sustainability.
- Expand influence.

www.aercafrica.org

Learn More

www.facebook.com/aercafrica

www.instagram.com/aercafrica_official/

twitter.com/aercafrica

www.linkedin.com/school/aercafrica/

Contact Us
African Economic Research Consortium
Consortium pour la Recherche Economique en Afrique
Middle East Bank Towers,
3rd Floor, Jakaya Kikwete Road
Nairobi 00200, Kenya
Tel: +254 (0) 20 273 4150
communications@aercafrica.org