

4th KIX Knowledge Café

Statistical Processes that make Education Data Usable and Credible

Date: 25 August 2026

Organizer: [African Union](#)'s Pan African Institute for Education for Development ([AU-IPED](#)), through the Global e-Schools and Communities Initiative ([GeSCI](#)), under the [Global Partnership for Education](#) – Knowledge and Innovation Exchange (GPE-KIX) Africa 19 Hub a joint endeavor with Canada's [International Development Research Centre](#) (IDRC)

Platform: Virtual - Zoom

Thematic Intervention: Polycarp Otieno, Education Specialist, UNICEF Eastern and Southern Africa Regional Office, KIX Africa 19 Hub.

Country presentations: Malawi, Sierra Leone, Somalia

Participating Countries: Sierra Leone, Uganda, Malawi, Somalia, Somaliland, Ethiopia, Kenya, Eswatini, Nigeria, South Sudan, Zimbabwe, The Gambia

GPE KIX AFRICA 19 HUB
KNOWLEDGE INNOVATION EXCHANGE

IDRC • CRDI

African Union

unesco
International Institute
for Capacity Building
in Africa

unicef
for every child

**4TH KIX
KNOWLEDGE CAFÉ
ON STATISTICAL PROCESSES**

Statistical Processes that Make
Education Data Credible and Usable

TUESDAY, 25 AUGUST 2026
2:00–4:00 P.M. EAT

COUNTRY SPOTLIGHTS
MALAWI • SOMALIA • SIERRA LEONE

THEMATIC INTERVENTION
UNESCO-UIS

PRESENTATION
EMIS Self-Assessment Platform by AU-IPED

REGISTER VIA THE LINK BELOW

African Union

unicef
for every child

unesco
International Institute
for Capacity Building
in Africa

gesci
Founded by UN ICT Task Force

Table of Contents

Executive summary.....	3
Background.....	3
Objectives.....	4
3. Opening and framing.....	4
3.1 Opening remarks.....	4
3.2 Thematic intervention: credibility is built during processing.....	4
4. Country experiences.....	5
4.1 Sierra Leone.....	5
4.2 Malawi.....	6
4.3 Education officials working in Somaliland.....	6
5. Discussion and cross-country learning.....	7
5.1 Validation should begin with use and work backwards.....	7
5.2 Feedback strengthens ownership and quality.....	7
5.3 Administrative and survey data should be reconciled carefully.....	7
5.4 Coordination requires shared structures and capacity.....	7
5.5 Technology must be useful and maintainable.....	8
5.6 Personal-data protection is an emerging priority.....	8
6. Cross-cutting challenges.....	8
7. Requests for support.....	9
8. Recommendations.....	10
8.1 Actions for ministries of education.....	10
8.2 Actions for governments and statistical systems.....	10
8.3 Actions for AU-IPED and the GPE KIX Africa 19 Hub.....	10
9. Follow-up action matrix.....	11
10. Conclusion.....	11

Executive summary

The 4th KIX Knowledge Café examined the statistical processes that determine whether education data can be trusted and used. Convened by AU-IPED under the GPE KIX Africa 19 Hub, the session brought together education planners, EMIS practitioners, statisticians, researchers and development partners for a country-led exchange on data validation, quality assurance, harmonization, documentation and institutional coordination.

The consistent message was that collecting more data, or collecting it more frequently, does not automatically result in better statistics or better decisions. Credibility is established in the processes between collection and use - through checks for completeness, plausibility, consistency, reproducibility and timeliness; through validation at source; and through clear documentation of how figures are produced.

“More frequent data or greater volumes of data do not necessarily produce better statistics or better decisions. Credibility is won or lost during processing...through validation, quality assurance, harmonization and documentation.”

— Polycarp Otieno, Education Specialist, UNICEF Eastern and Southern Africa Regional Office

Country presentations demonstrated different stages of EMIS development. Sierra Leone described layered quality controls and independent verification. Malawi presented a largely paper-based system undergoing digitization, supported by a strong institutional relationship with its National Statistical Office but challenged by fragmented subsector platforms. Education officials working in Somaliland described a web-based EMIS combined with paper collection in low-connectivity areas, while identifying gaps in formal validation, statistical capacity, documentation and coordination.

Participants explicitly requested technical and capacity-building support on EMIS interoperability, integrated education data, data-exchange mechanisms, EMIS–NSO collaboration, statistical analysis and validation. The discussion also identified personal-data protection as an emerging priority for future peer learning.

Background

AU-IPED leads the Meeting the Data Challenge thematic area within the GPE KIX Africa 19 Hub. Through EMIS self-assessments, peer reviews, national policy dialogues and Knowledge Cafés, AU-IPED supports GPE partner countries to strengthen the production, management and use of education data.

The Knowledge Café model was introduced as a structured but informal country-led peer-learning mechanism. It enables ministries, focal points, technical experts and partners to discuss shared priorities, compare national experiences and identify practical support needs. Earlier Cafés examined EMIS infrastructure, responsive national EMIS policies and sustainable EMIS financing.

The fourth Café responded to evidence from AU-IPED-supported processes showing that increased data collection and investment in digital platforms have not always produced statistics that are credible, comparable and confidently used. The discussion therefore focused on the statistical processes between collection and use: validation, quality assurance, harmonization, indicator construction, documentation, dissemination and institutional accountability.

Objectives

- Build a shared understanding of how statistical processes shape trust in education data.
- Document country practices, gaps and innovations in statistical production.
- Explore coordination models between EMIS units and National Statistics Offices.
- Identify practical actions and support required to strengthen statistical rigor within existing systems.
- Generate evidence for future KIX Africa 19 Hub knowledge products and peer-learning activities.

3. Opening and framing

3.1 Opening remarks

Speaking on behalf of the Head of AU-IPED, Lukman Jaji, Policy Officer for EMIS and KIX Africa 19 Focal Point, welcomed participants and situated the fourth Café within an ongoing series on the building blocks of effective education-data systems. He emphasized that African countries are at different stages of EMIS development and face different combinations of infrastructure, governance, capacity and data-use challenges.

“The more conversations we have, the easier it is for us to understand the pain points of different countries. These dialogues help us understand unique national situations, learn what is working and design issue-targeted interventions.”

- **Lukman O. Jaji, Policy Officer for EMIS, AU-IPED**

3.2 Thematic intervention: credibility is built during processing

Polycarp Otieno framed statistical processing as the stage at which education data either gains or loses credibility. He cautioned that neither the frequency nor the volume of collection guarantees better statistics. Without validation, quality assurance and harmonization, analysis may reproduce errors and provide a weaker basis for policy and resource-allocation decisions.

He proposed five tests for credible statistics:

- Completeness - whether all expected schools and populations are represented, and whether missing cases are known.

- Plausibility - whether reported values are reasonable in relation to context and expected ranges.
- Consistency - whether figures can be reconciled over time and across administrative, assessment and household sources.
- Reproducibility - whether definitions, methods and validation steps are stable, documented and repeatable.
- Timeliness - whether information reaches decision-makers in time to influence action.

Polycarp proposed four shifts which included the following:

- Validate at the source and institutionalize reconciliation through multi-agency technical structures
- Document methods so that every statistic can be reproduced, and judge data systems by whether they improve decisions.

He stressed that sophisticated platforms have limited value if users cannot generate information that helps them identify and resolve problems.

4. Country experiences

4.1 Sierra Leone

Emmanuel Vandy Deoud, KIX Focal Point and Development Partners' Coordination Desk Officer at Sierra Leone's Ministry of Basic and Senior Secondary Education, presented the statistical process used for the 2025 Annual School Census. The process combined office editing, back-checking, data cleaning and independent verification before the national database was finalized for analysis and reporting.

Practices highlighted

- Post-fieldwork checks for coverage, duplicated schools, non-response and incorrect values.
- Use of the school master list, GPS coordinates and enrolment records to investigate possible duplicates and verify school identity.
- Spot checks by district IT personnel and school quality-assurance officers, complemented by telephone verification when physical revisits were impractical.
- Automated cleaning using Stata and Python, combined with row-by-row and column-by-column manual review for cases that could not be resolved through code.
- Independent verification by the National Monitoring and Evaluation Agency in 345 schools across all districts, using both purposive and random selection.

His presented that cleaning removed duplicated school and teacher records before publication. The example demonstrated the value of applying different controls to different sources of error and of closing verification loops before report writing.

“No figure reaches publication untested. Editing, back-checking, cleaning and independent verification each remove a different source of error; together, they are what make reported statistics credible.”

- Emmanuel Vandy Deoud, Ministry of Basic and Senior Secondary Education, Sierra Leone

4.2 Malawi

James Namfuko, Deputy Director of Education Planning in Malawi’s Ministry of Education, described a largely manual, paper-based EMIS undergoing gradual digitization. Data collection is coordinated through zones and districts, after which officers meet physically to clean the data before further processing, analysis and report writing at central level.

Digitization efforts began as pilots and expanded from 2022–2023. Systems have been developed for secondary education, higher education and teacher management. However, project-based and subsector-specific development has created separate systems that are not yet fully integrated, and the primary subsector remains a major gap.

Institutional coordination

Malawi’s EMIS operates within the National Statistical System and is guided by the National Statistics Act of 2013. The National Statistical Office supports common standards, codes of practice and quality-assurance frameworks, while education statistics are compiled into the national statistical yearbook and used for national, regional and global reporting. This provides a strong institutional foundation, although validation and reconciliation between school-census and household-survey data require further attention.

Challenges and priorities

- Publication delays and avoidable errors associated with manual processing.
- Fragmented subsector platforms and the absence of a single reference point for education data.
- Differences between school-census and household-level education data.
- Need for automated completeness, range and logic checks at the point of entry.
- Need for common templates, standard operating procedures, a centralized data dictionary, metadata and continuity arrangements.
- Malawi has developed an EMIS governance framework that prioritizes data management, quality assurance and standardization. Its longer-term objective is an integrated EMIS that covers all education subsectors and provides a single reference point for the sector.

4.3 Education officials working in Somaliland

Hashi Hassan, supported by Safia Jibril Abdi, GPE Focal Point in Somaliland, and described a web-based EMIS managed centrally within the education ministry’s policy and planning function. Urban schools with internet access submit data online, while rural schools use paper forms that are later entered through district offices.

The system supports an annual statistical yearbook used for planning, financing and budget allocation. However, quality control between collection and publication has not yet been sufficiently formalized. The presentation identified missing or incorrect student information, duplicate learner records, connectivity limitations, limited verification of teacher data and heavy workloads for regional EMIS focal persons.

Priorities identified

- A centralized validation framework linking school, regional and national levels.
- Automated checks, manual review and periodic field spot checks.
- Stronger statistical analysis, visualization and institutional capacity.
- Improved data harmonization, documentation and coordination with statistical authorities.
- Further investment in tools, infrastructure, monitoring and evaluation functions, and regional capacity.

5. Discussion and cross-country learning

5.1 Validation should begin with use and work backwards

Participants emphasized that validation should not be treated as one final exercise. Built-in checks, school report cards, district performance reviews, census sample checks and feedback loops can enable errors to be identified while the data remains current. Starting with critical operational and strategic decisions - and then identifying the data required, can also prevent systems from collecting information that does not respond to policy needs.

5.2 Feedback strengthens ownership and quality

The discussion noted that data quality improves when the people entering information see it returned in a form that is useful to them. School report cards and local feedback can create incentives for accuracy, while data that only moves upwards risks becoming disconnected from school-level priorities.

5.3 Administrative and survey data should be reconciled carefully

Participants raised questions about differences between operational EMIS data, household surveys and learning assessments, including their implications for estimating out-of-school populations. The discussion clarified that sources may legitimately differ because of timing, definitions, sampling design and whether they measure enrolment or participation. Harmonization should therefore focus on transparency, comparability and triangulation rather than forcing a one-to-one match between unlike sources.

5.4 Coordination requires shared structures and capacity

Nwosu Rowland of Nigeria's National Bureau of Statistics requested capacity-building that jointly involves NSO and education-statistics personnel. Participants noted that education indicators frequently depend on population denominators and household data produced by

NSOs; stronger cooperation is therefore essential for official reporting, consistency and shared understanding.

5.5 Technology must be useful and maintainable

Discussion of DHIS2 and other digital platforms highlighted the value of configurable and locally maintainable systems. Sierra Leone referred to ongoing national core-team training on DHIS2 configuration, analytics and maintenance. Participants stressed that technology should support local problem-solving, feedback and data use rather than becoming a stand-alone project product.

5.6 Personal-data protection is an emerging priority

As ministries increasingly track individual learner and teacher records, participants proposed a future Knowledge Café on personal-data protection. Such a discussion should address education-sector responsibilities within wider national legal frameworks, statistical acts and civil-registration systems.

6. Cross-cutting challenges

Challenge	Policy significance
Late or informal validation	Errors are often discovered after data has moved from schools to districts and central ministries, making correction slower and more expensive.
Fragmented systems	Subsector- and project-specific platforms create parallel databases, inconsistent definitions and weak interoperability.
Manual workflows	Paper collection and physical cleaning exercises contribute to delays, transcription errors and limited real-time feedback.
Weak reconciliation	Administrative, household-survey and assessment data may differ because of timing, definitions or design, but reconciliation mechanisms are often limited.
Institutional gaps	EMIS–NSO coordination varies widely and may not extend to joint validation, indicator construction or dissemination.
Capacity constraints	Countries reported gaps in statistical analysis, visualization, quality assurance, documentation, system maintenance and regional support.
Connectivity and infrastructure	Rural schools may depend on hybrid paper and online approaches, creating uneven submission and validation processes.

Limited metadata	Insufficient documentation, data dictionaries and revision policies make statistics difficult to reproduce and defend.
Weak feedback and use	Data may travel upwards without returning to schools and districts as information that can guide action.
Privacy and governance	Individual-level systems are expanding faster than education-specific understanding of personal-data responsibilities.

7. Requests for support

The Café generated explicit country and participant requests that may inform follow-up planning. These requests do not constitute confirmed commitments. Country-specific support will require a formal written request, followed by review in relation to available resources.

Source	Support requested
Malawi	Technical, policy and capacity-building support for EMIS interoperability, data-exchange mechanisms and integrated education data for evidence-informed decision-making.
Malawi	Consideration of expert support or a peer exchange with a country demonstrating effective interoperability and inter-ministerial coordination.
Nigeria/NSO perspective	Joint capacity-building for National Statistics Office staff and education-statistics or EMIS personnel, and inclusion of NSO representatives in relevant Hub initiatives.
Education officials working in Somaliland	Technical engagement on validation, quality assurance, analytical capacity, harmonization, documentation and institutional strengthening, through channels consistent with AU and GPE KIX arrangements.
Participants	A future peer-learning session on personal-data protection practices for ministries collecting individual learner and teacher data.

“We are requesting technical and capacity-building support to strengthen not only the EMIS platform, but also interoperability, data-exchange mechanisms and the use of integrated education data for evidence-informed decision-making.”

- Caroline Chipinga-Mwale, KIX Africa 19 Focal Point, Malawi

8. Recommendations

8.1 Actions for ministries of education

- Embed completeness, range, logic and duplicate checks in data-collection tools and correct errors at source.
- Establish written validation protocols, escalation pathways and responsibilities from schools to national level.
- Maintain current school master lists, data dictionaries, metadata repositories and revision logs.
- Combine automated cleaning with risk-based manual review, spot checks and independent verification where appropriate.
- Return useful information to schools and districts through report cards, dashboards and performance reviews.
- Document how statistics influence planning, teacher deployment, resource allocation and support to underserved groups.

8.2 Actions for governments and statistical systems

- Formalize EMIS–NSO coordination through working groups, shared calendars, common definitions and joint validation arrangements.
- Clarify how administrative, survey and assessment data should be compared and used for different policy questions.
- Develop interoperability and data-exchange frameworks that connect education subsectors and relevant non-education systems.
- Strengthen staffing, professional development and retention for statistics, data engineering, analysis, visualization and quality assurance.
- Align individual-level education systems with national personal-data protection and civil-registration frameworks.

8.3 Actions for AU-IPED and the GPE KIX Africa 19 Hub

- Provide countries with a practical statistical-process checklist aligned with the AU EMIS Norms and Standards.
- Facilitate peer exchange on source-level validation, interoperability, EMIS–NSO coordination and sustainable use of DHIS2 or comparable platforms.
- Include NSO representatives in relevant capacity-building and inter-ministerial coordination activities.
- Review formal country requests and determine feasible support based on documented needs and available resources.
- Document country practices as reusable case studies rather than presenting one technology or model as universally applicable.

Convene a future dialogue on personal-data protection, legal mandates and responsible use of individual-level education data.

9. Follow-up action matrix

Action	Proposed lead	Timing	Status/condition
Submit a formal statement of Malawi's interoperability and capacity needs.	Malawi focal point/Ministry of Education	Soon	Requested during Café
Review the Malawi request and assess possible technical assistance or peer exchange.	AU-IPED/KIXA19 Hub partners	Following receipt of request	Subject to review and available resources
Document Sierra Leone's DHIS2 configuration, analytics, maintenance and core-team training as a possible peer-learning case.	AU-IPED / Sierra Leone	Future knowledge exchange	Peer-learning opportunity
Include NSO representatives in relevant education-statistics capacity building.	AU-IPED/KIXA19 Hub partners	To be determined	Recommendation accepted for planning
Reschedule the AU EMIS Self-Assessment Platform demonstration and circulate access guidance.	AU-IPED	Next engagement	Required following technical postponement
Develop the next Knowledge Café concept, including personal-data protection and legal mandates.	AU-IPED	Next Knowledge Café cycle	Proposed by participants

10. Conclusion

The fourth Knowledge Café confirmed that credible education statistics depend as much on governance and process as they do on technology. Countries are investing in digital platforms, but validation, documentation, interoperability, institutional coordination and professional capacity remain decisive gaps.

The country experiences also demonstrated that practical solutions already exist. Validation at source, layered back-checking, independent verification, strong links with National Statistics Offices, hybrid collection approaches and locally maintained digital platforms. The next step is to convert these lessons into targeted support and peer learning that responds to each country's context.

For AU-IPED and the GPE KIX Africa 19 Hub, the discussion provides a clear agenda. To support countries to institutionalize statistical quality, connect fragmented systems,

strengthen EMIS-NSO collaboration and ensure that education data can be defended and used in decisions affecting learners, teachers and resources.