Big Data and Data Science for Official Statistics in Africa

Oliver Chinganya
Director
African Centre for Statistics
United Nations Economic Commission for Africa
Addis Ababa, Ethiopia

CONTENTS

- Africa Centre for Statistics focus areas
- Official Statistics
- Terms and definitions
- ECA Data Science and Big Data initiatives
- Regional hub for Big Data and Data Science for Africa
- Big Data and Data Science situation in Africa
- Key takeaways in Big Data and Data Science for official statistics

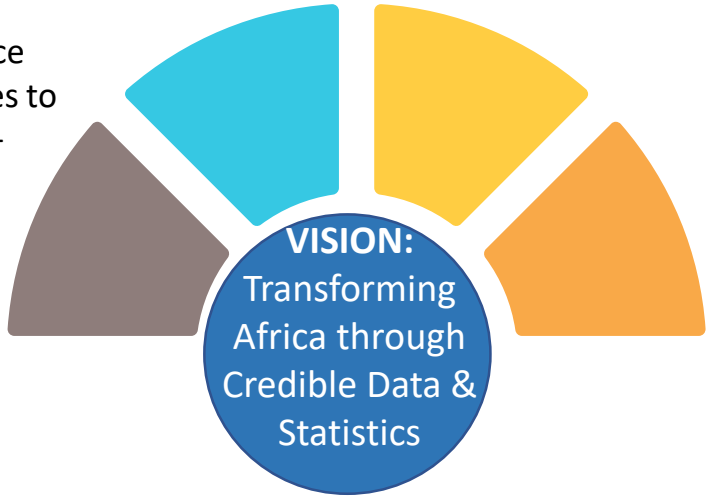
African Centre for Statistics - Focus

Programmatic focus

Capacity development: Enhance the capacities of member States to respond to demands for fit-for-purpose data

Governance & policy: Promote statistical governance mechanisms

Data services: Provide authoritative development information on Africa & support in data production for ECA research & policy analysis work.

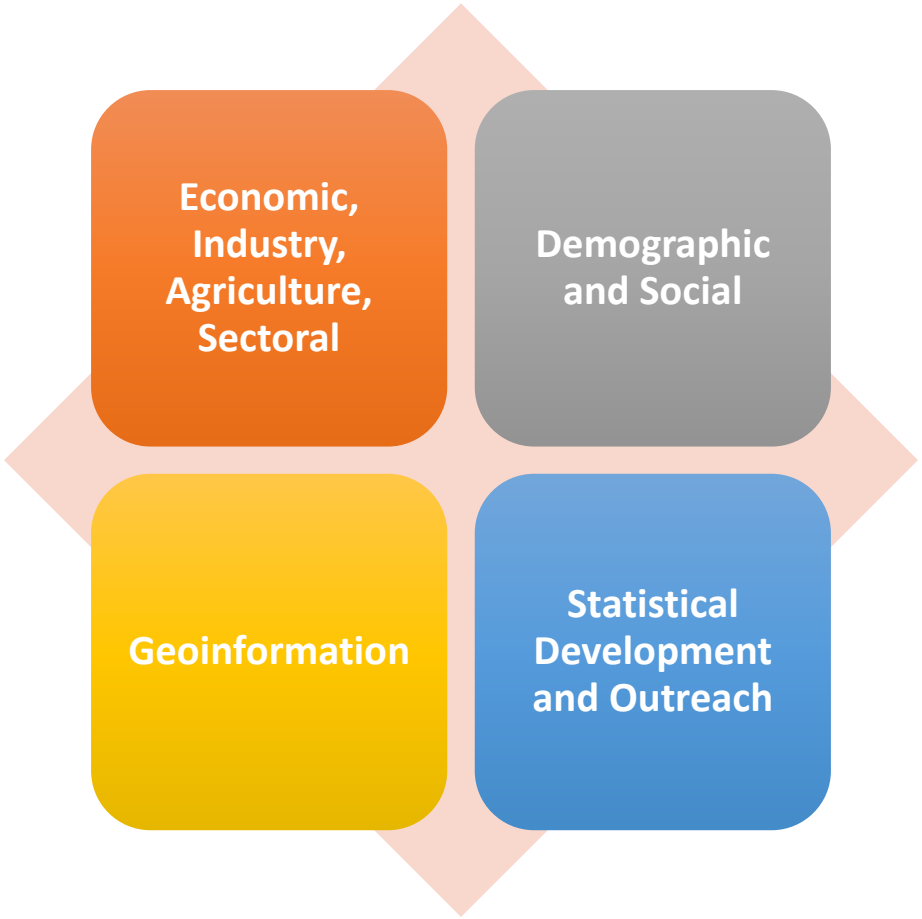


Modernization of statistical production & services: Conduct & promote research on methods, concepts, definitions & classifications

Outreach & partnership: Build partnerships, nurture technical assistance & advisory services, collaboration, & networking with national, regional & international entities.

African Centre for Statistics - areas

To improve the production, dissemination & use of quality data & statistics in the framework of the Agenda 2030 & AU 2063 to support economic and social development in Africa.



ECA DATA SCIENCE AND BIG DATA INITIATIVES

"To support Member States in effectively capturing alternative data sources, applying appropriate data science tools and techniques in providing credible and insightful data and statistics"

ECA Data Science initiatives - Strategic Focus

Leadership and coordinaiton

- Take a leading role across Africa in the promotion of Big Data and data science methods and techniques for use within member States

Research and projects

- To undertake research,
- To coordinate member state partnerships,
- To assist with the build of prototypes,
- The provision of demonstrations, and
- The packaging and dissemination of solutions for all member States

Capacity building

- To provide various capability building activities that build big data and data science knowledge and skills for member states.

Leadership

Impact

- The African statistical system builds data science knowledge and skills in a coherent way, clear guidance and frameworks on how to use new data and how to apply new methods in data science to policy questions

Research and Project Implementation

Impact

- Research activities will enable the data science work to quickly demonstrate the benefits of using data science locally
- Delivering smaller projects designed to build coding skills, confidence and impact

CAPABILITY BUILDING

Impact

- Will upskill and equip member states to independently undertake data science and to grow their own Big Data and data science capability and capacity.

- Coordinate and/or provide various training activities, ranging from classroom events, workshops, hackathons, and mentoring and coaching. The Centre will also provide details of courses available from other approved partners, such as academia.

What has been done & planned?

Big Data & Data Science activities undertaken

CAPACITY BUILDING:

- **Enabling Data Science in Africa - A series of on-line conferences:** Promote the ECA's Data Science and Big Data programme and the activities available to support member states; Share and understand local challenges in using and applying data science methods and techniques; Raise awareness of the value of data science, the different applications and the benefits available; Identify potential research projects that could be taken forward in partnership with ACS.
- **Data Science Leadership Training Workshop:** The workshop aims to lay down important ground information about how we should think and lead data science initiatives; to critically about the non-conventional data ecosystem such as Big Data and other data sources. The online workshop will discuss the key issues surrounding data science leadership and challenges, including AI and Big Data.
- **Data Science Capability Building: Python Training Workshop On-line Training Workshop -** series of virtual workshops – Two staff from each member States invited.
- **Data Science Art of the Possible – For ECA management and staff**
- **Python courses for ECA staff**

ONGOING PROJECT

- Using web scraping for CPI
- Scanner data for CPI

PROJECT PROPOSALS

- Training - Machine Learning for Official Statistics
- Six-month Big Data and Data science capacity building with capstone project

Partnership and Coordination

Building on this experience and working with partners to establish Big Data and data science methods and culture will be crucial to the success of the mission/vision for the Centre.

- ***Skills*** – partners will have skills and experience in different aspects of Big Data and data science
- ***Data*** – access to new data sources is key and different partners will be able to unlock access to data sources.
- ***IT infrastructure*** – the local ambition for Big data and data science will need an appropriate IT infrastructure and partners will be able to advise and develop on which IT infrastructure is most appropriate.
- ***Expertise and ideas*** – a fresh and experienced perspective is a critical tool to help identify new solutions for where Big Data and data science can inform or solve these data problems.
- ***Funding and resources*** – Big Data and data science will require new funding and resources and partnerships will be able to bring either or both to assist the African Centre and member states meet its data science aims.
- ***Advocacy*** – lastly, partners can help support member states' ambitions for Big Data and data science work through promoting and demonstrating the value of data science, within country, through political lobbying.

Current and potential partners

Member States (MS)	Stat Afrique	African Development Bank	ONS Data Science Campus	GPSDD
Other UN organisations:	FAO, UNHCR, UNICEF, UNDP, UNFPA, UNSD, UN Global Pulse Lab, UNSD, ILO, …	RCMRP	UN Global Platform	Academia
AIMS	University of KwaZulu-Natal	Regional Training Centres	Data First	EUROSTAT
OECD	Private Sectors	MTN	GEOPOLL	ACBF
	Regional Economic Communities	WB, IMF		

The Regional Hub for Africa is a collaboration between the

National Institute of Rwanda (NISR)

and

United Nations Economics for Africa (UNECA).

The partnership is established via a Memorandum of Understanding among NISR, ECA, and DESA.

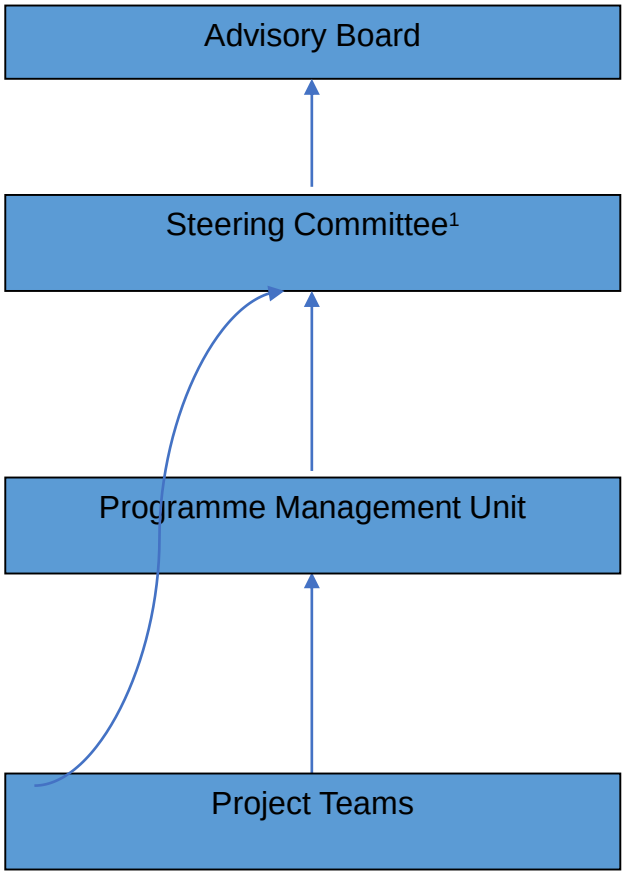
Regional Hub for Big Data and Data Science for Africa

- Aid the facilitation of cross-border collaboration projects that seek to apply big data, alternative data sources, and data science, to complement official statistics and fill data gaps for SDG indicators,

Provide knowledge on newly developed methods, algorithms and tools, via webinars, seminars and the UN Global Platform, and

- Provide training in the use of big data and data science for the community of official statisticians in Africa.

Governance and reporting structure for the un regional hub for Africa



Membership: NISR, UNECA & UNDESA
Role: Oversee the implementation of the MOU
Frequency: Annual

Membership: NISR, UNECA, UNDESA & Five Member states representing all African regions
Role: To approve and oversee the implementation of the roadmap, prioritise resources, select research projects / workstreams
Frequency: Quarterly
Reporting to the Annual Advisory Board

Membership: NISR, UNECA & ONS
Role: Responsible for delivering the roadmap -receiving and validating research project proposals, drawing on experts, enabling partnerships, managing resources, enabling the project, implementing any training to support project teams
Frequency: regular stand ups
Reporting to the Steering Committee, as requested

Membership: Member states, experts, data scientists, others, as needed
Role: To deliver specific solutions
Frequency: Agile
Reporting to the Programme Management Unit and Steering Committee, as requested

- **Uganda** - Use of satellite imagery to identify metal roofs and as a proxy-indicator for poverty by the Uganda Bureau of Statistics
- **Namibia** – Using anonymized call detail records (CDRs) to measure migration in Namibia - internal migration estimates can be derived and modeled from CDRs at subnational and annual scales, and how precision and accuracy of these estimates compare to census-derived migration statistics.
- **South Sudan** – use of Earth Observation data for estimating cattle census
- **Senegal** – use of Earth Observation data for mapping crops at national level.

Big Data and Data Science projects situation

- Donors led
- Low uptake, capacity and infrastructure gap
- At prototype and proof of concept

Key takeaways

- Access to big data sources – public & private --- partnership
- Methodology and tools - machine learning models, statistical methods, reusable programming codes/scripts
- Capacity building – training and knowledge/experience sharing
- Infrastructure - local and cloud based ICT/software services

ideas for a prosperous Africa

IdeastoAction

THANK YOU!

Ideas to Action