# African Open Data Readiness Index for AI (AODRIA)

**Official Methodology Version 1.0**
Date: 27 February 2026
Developed by: AI Made in Africa / GIZ African Union
For the Africa Data Stewardship Award 2026
Status: Public Release
License CC-BY 4.0

**Executive summary**

The African Open Data Readiness Index for AI(AODRIA) is a continental benchmarking instrument designed to evaluate, rank, and recognize open datasets relevant to Africa based on their fitness for Artificial Intelligence (AI) applications, data stewardship standards, and real-world impact. This methodology (v1.0) underpins the Africa Data Stewardship Award, transforming subjective assessment into a transparent, reproducible, and actionable scoring framework.

Our goal is not merely to rank datasets, but to incentivize improvements in data quality, openness, governance, sustainability, and socio-economic value across the African data ecosystem.

## 1. Introduction

The **African Open Data Readiness Index for AI (AODRIA)** is the continent's first dedicated, dataset-level benchmarking tool that measures how ready open datasets are for responsible and high-impact AI applications.

AODRIA evaluates individual Africa-relevant open datasets across five weighted pillars and produces a final score out of 100. The index directly informs the selection of winners for the **Africa Data Stewardship Award**, which highlights organizations demonstrating exemplary stewardship of open datasets in Africa.

The methodology is designed to be transparent, replicable, and adaptable, enabling future updates as the African data ecosystem evolves.

### 1.1. Objectives

The African Open Data Readiness Index for AI aims to:

- Provide transparent and objective scoring of open datasets relevant to AI innovation in Africa.
- Establish a continental benchmark for dataset quality and accessibility.
- Recognize and reward excellent data stewardship.
- Highlight gaps and drive improvements in African open data quality.

### 1.2. Alignment with International Data Principles

The AODRIA framework aligns with internationally recognized data governance standards, including the FAIR Data Principles (Findable, Accessible, Interoperable, Reusable).

Rather than treating FAIR compliance as a standalone pillar, AODRIA integrates FAIR principles across multiple dimensions including Openness, AI Readiness, socio-economic impact and sustainability.

This ensures that datasets evaluated under AODRIA support discoverability, interoperability, and responsible reuse.

## 2. Scope & Eligibility

To be evaluated under AODRIA v1.0, a dataset must meet the following baseline criteria:

### 2.1. **Mandatory Eligibility Criteria**

| Criterion | Requirement | Verification Method |
|---|---|---|
| **Africa Relevance** | Data must originate from, describe, or be directly applicable to African contexts (continental, regional, or national) | Publisher declaration + content review |
| **Accessibility** | Data must be accessible to the public (Open Data) or available via a streamlined access request process (Managed Access) | Access test by jury |
| **Machine-Readability** | Data must be available in a structured format (CSV, JSON, XML, GeoJSON, Parquet, etc.). Unstructured PDFs alone are ineligible | Format check |
| **Active Status** | Dataset must have been updated or verified within the last 3 years (unless archival/historical with documented value) | Metadata timestamp review |
| **Language** | Metadata must be available in at least one widely understood language (English, French, Portuguese, Arabic, or Swahili) | Landing page review |

### 2.2. Exclusion criteria

Datasets will be excluded if they:

- Contain undisclosed personal data (PII) without ethical clearance.
- Violate national data protection laws or AU Data Policy Framework
- Are behind prohibitive paywalls (> $50 USD for access)
- Have no identifiable publisher or custodian.
- Are duplicates of already submitted datasets

## 3. Pillar Framework

Each dataset will be evaluated across five core pillars. Each pillar will be scored on a 1–5 scale, where 1 = Poor, 3 = Good / Acceptable and 5 = Excellent.

### 3.1. AI Readiness (Weight: 30%)

Measures how easily the data can be ingested and used for machine learning models.

| Score | Description |
|---|---|
| 5 | **Excellent** – Fully structured (CSV, JSON, API, GeoJSON, Parquet), clean, large volume, well-labelled, almost no preprocessing needed. |
| 3 | **Good** – Structured format, moderate cleaning or volume, usable for AI with some effort. |
| 1 | **Poor** – Mostly unstructured (PDF, images, tables without structure), very heavy preprocessing required or not machine-readable. |

### 3.2. Openness (Weight: 20%)

Measures accessibility and legal reusability.

| Score | Description |
|---|---|
| 5 | **Excellent** – Immediate public download, fully permissive license (CC0/CC-BY/Open Data Commons), no registration, rich metadata & documentation. |
| 3 | **Good** – Publicly downloadable with minor barriers, open license, adequate documentation. |
| 1 | **Poor** – Not truly open (paywall, proprietary license, registration required, or no clear license). |

### 3.3. Governance and FAIR Compliance (Weight: 20%)

Measures ethical standards, transparency and quality of stewardship AND compliance with FAIR principles (Findable, Accessible, Interoperable, Reusable).

| Score | Description |
|---|---|
| 5 | **Excellent** – No PII (no names, phones, exact GPS), includes data dictionary & consent statement AND clear provenance/collection methodology AND meets all FAIR principles. |
| 3 | **Good** – Anonymized but incomplete documentation on collection methodology or partial FAIR compliance. |
| 1 | **Poor** – Contains raw PII or no governance/FAIR information. |

### 3.4. Socio-economic impact (Weight: 15%)

Measures real-world value created by the dataset in Africa.

| Score | Description |
|---|---|
| 5 | **Excellent** – Clear evidence of real-world use (cited in policy papers, used in national AI projects, generated jobs/reports, or measurable SDG impact). |
| 3 | **Good** – Some documented use (academic citations, NGO reports, or downloads >1,000). |
| 1 | **Poor** – No evidence of use or impact. |

### 3.5. Sustainability (Weight: 15%)

Measures long-term viability and maintenance.

| Score | Description |
|---|---|
| 5 | **Excellent** – Regularly updated (at least quarterly), stable institutional hosting, active community/funding, versioning in place. |
| 3 | **Good** – Updated within last 18 months, reasonably stable hosting and some maintenance. |
| 1 | **Poor** – Last update >24 months ago, unstable hosting, or abandoned/one-time dataset. |

## 4. AODRIA Score Calculation

**Official Formula**

**AODRIA Score (out of 100) =** [(AI Readiness × 0.30) + (Openness × 0.20) + (Governance & FAIR × 0.20) + (Socio-economic impact × 0.15) + (Sustainability × 0.15)] × 20

## 5. Tier Classification

| Tier | AODRIA Score | Recognition level |
|---|---|---|
| Platinum | 90.0 – 100.0 | Africa Data Stewardship Award winner |
| Gold | 80.0 – 89.9 | Winner / Highly Commended |
| Silver | 70.0 – 79.9 | Strong Recognition |
| Bronze | 60.0 – 69.9 | Recognized open dataset |
| Needs Improvement | Below 60.0 | Feedback only |

## 6. Evaluation Process

The evaluation process follows four steps:

1. Dataset identification through open call, nomination, or expert curation
2. Eligibility screening based on baseline criteria
3. Expert Jury scores **independently** using the AODRIA framework.
4. Final score = weighted formula above
5. Validation & publication of all raw scores

## 7. Transparency & Reproducibility

To ensure transparency:

- All datasets and final AODRIA scores will be published openly on a public leaderboard.
- The AODRIA methodology will remain publicly accessible and is archived with a persistent identifier (DOI) for citation and academic use.

## 8. Limitations & Future Improvements

- The simplified 1–3–5 scoring scale is used to reduce evaluator bias and speed up review.
- Some indicators rely on public metadata and publisher declarations.
- Future versions of AODRIA may incorporate automated indicators and expanded scoring ranges.

## 9. Versioning and Updates

This document represents Version 1.0 of the African Open Data Readiness Index for AI methodology.

Future versions may introduce:

- Additional technical indicators
- Sector-specific scoring
- Automated data quality metrics
- Expanded jury evaluation procedures

**Approved by:** AI Made in Africa **Date:** 27 February 2026