

TECHNISCHES POLITIKPAPIER

Die fehlende Ebene im KI-Ökosystem Afrikas

Von standardmäßig offen zu KI-nutzbar durch Design

Mitwirkende

Yonas Gebremichael

Saba Tiku

Olamide Goriola

Lorna Munanie

GIZ Afrikanische Union

GIZ Afrikanische Union

ConvergenceAI

ConvergenceAI

Zusammenfassung

Afrika generiert mehr Daten als je zuvor. Dennoch steht der Kontinent vor einer kritischen Usability-Lücke: Die Infrastruktur, die benötigt wird, um Rohdatensätze in vertrauenswürdige, kontextualisierte und pipeline-reife Ressourcen für die KI-Entwicklung zu verwandeln, ist weitgehend unterentwickelt. Dieses Briefing, das gemeinsam von ConvergenceAI und GIZs AI Made in Africa-Initiative erstellt wurde, stützt sich auf Ökosystemintelligenz aus mehr als 40 afrikanischen Ländern sowie Datensatzanalysen aus 30+ Ländern, um diese Lücke zu diagnostizieren und konkrete Maßnahmen zu empfehlen.

Ländern mit fortschrittlichen KI-Strategien fehlt häufig die grundlegende Dateninfrastruktur, um diese umzusetzen. Datensätze in Prioritätssektoren – Landwirtschaft, Gesundheit, Klima – sind fragmentiert, schlecht dokumentiert und rechtlich unklar. Die Herausforderung ist weder ein Mangel an Daten noch ein Mangel an Strategie. Es fehlt an einer verbindenden Schicht: die Kuratierung, Annotation, Dokumentation und Interoperabilitätsstandards, die Daten von Natur aus nutzbar machen.

Dieses Briefing fordert Regierungen, Entwicklungspartner, Forscher und Plattformen auf, die Datenbereitschaft als zentrale nationale Infrastruktur und nicht als sekundäre Nachgedanken zu betrachten.

Zentrale Erkenntnisse

01	Datenverfügbarkeit ist nicht gleich Datennutzbarkeit. Das Volumen offener Daten in Afrika wächst, doch die meisten Datensätze können ohne erhebliche Vorverarbeitung keine verlässlichen KI-Pipelines speisen.
02	KI-Strategie und Dateninfrastruktur sind nicht aufeinander abgestimmt. Den nationalen KI-Prioritäten in Landwirtschaft, Gesundheit und Klima stehen keine einsatzbereiten Datensätze in ebendiesen Sektoren gegenüber.
03	Vier strukturelle Fehlermuster treten wiederholt auf: Fragmentierung, mangelnde Aufbereitung, Unzugänglichkeit und die Nutzbarkeitslücke selbst.
04	Die geografische Konzentration verschärft die Ungleichheit. Aktive Datenökosysteme konzentrieren sich stark auf Kenia, Nigeria, Südafrika, Ruanda und Ägypten.
05	Die Sprachlücke ist gravierend. Die Abhängigkeit von ausschließlich englischsprachigen Metadaten schließt den größten Teil der sprachlichen Vielfalt Afrikas strukturell aus.
06	Katalogisierung ist keine Transformation. Zu wissen, dass ein Datensatz existiert, macht ihn noch nicht nutzbar – es fehlt die interpretierende und kontextualisierende Ebene.

Untersuchungsrahmen und Datenquellen

- **Geografische Abdeckung:** über 40 afrikanische Länder (Ökosystem-Monitoring von ConvergenceAI); über 30 Länder (Bewertungen durch Open Data Africa / AODRIA)
- **Erfasste Sektoren:** Landwirtschaft, Gesundheit, Klima, Sprache sowie Wirtschafts- und Bevölkerungsdaten
- **Primärquellen:** Intelligence-Plattform von ConvergenceAI, African Open Data Readiness Index for AI (AODRIA), Erhebungen der Weltbankgruppe, Broken-Links-Index der Open Government Partnership
- **Zielgruppe:** politische Entscheidungsträgerinnen und Entscheidungsträger, Verantwortliche für KI-Strategien, Entwicklungsfinanzierungsinstitutionen, Programmverantwortliche im Bereich offener Daten

Die Frage hinter der Frage

Fast das gesamte vergangene Jahrzehnt hindurch kreiste die lauteste Debatte in afrikanischen KI-Kreisen um die Verfügbarkeit: Erzeugt der Kontinent genügend Daten, um Modelle zu trainieren, Analysen durchzuführen oder öffentliche Dienstleistungen zu betreiben? Die Antwort lautet zunehmend: ja. Von landwirtschaftlichen Erträgen im ländlichen Äthiopien über Gesundheitsdaten in kenianischen Countys bis hin zu Satellitenbildern aus der Sahelzone – die Datenerzeugung auf dem Kontinent beschleunigt sich rasant. Nationale Datenportale haben sich vervielfacht. Open-Government-Initiativen haben an Zugkraft gewonnen. Forschungseinrichtungen, zivilgesellschaftliche Organisationen und statistische Ämter haben das online verfügbare Angebot stetig erweitert.

Die folgenreichere Frage lautet: Lassen sich diese Daten tatsächlich nutzen?

Zu häufig bleiben vielversprechende Datensätze in Silos gefangen. Sie sind fragmentiert, unzureichend dokumentiert oder technisch nicht auf moderne KI-Pipelines vorbereitet. Daraus entsteht eine kritische Nutzbarkeitslücke. Was Afrika braucht, sind nicht schlicht mehr offene Daten; es braucht eine Infrastruktur für Datenbereitschaft, die Datensätze auffindbar, rechtlich nachnutzbar, technisch interoperabel und für KI-Systeme inhaltlich aussagekräftig macht.

Die Intelligence-Plattform von ConvergenceAI und die Initiative „AI Made in Africa“ der GIZ bei der Afrikanischen Union eröffnen einander ergänzende Perspektiven auf diese Herausforderung. ConvergenceAI erfasst die Ebene des Gesamtökosystems – Politikrahmen, Innovationskennzahlen und Maßnahmen zur Kapazitätsentwicklung in mehr als 40 afrikanischen Ländern – und zeichnet so ein laufend aktualisiertes Bild davon, wo die KI-Bereitschaft steht und wo sie ins Stocken gerät. „AI Made in Africa“ bewertet über seinen Arbeitsstrang zur Datenzugänglichkeit die Verfügbarkeit, Qualität und KI-Tauglichkeit von Datensätzen in über 30 Ländern; als Bewertungsrahmen dient dabei der [African Open Data Readiness Index for AI \(AODRIA\)](#). Dieser beurteilt Datensätze anhand von Dimensionen wie KI-Bereitschaft, Offenheit, Governance und FAIR-Konformität, sozioökonomischer Wirkung und Nachhaltigkeit.

Nebeneinander betrachtet zeichnen beide Plattformen ein ernüchterndes, aber handlungsleitendes Bild. Afrika erzeugt Daten und entwickelt Strategien, doch die verbindende Ebene – die Überführung von Rohdaten in vertrauenswürdige und kontextualisierte Ressourcen für KI – bleibt unterentwickelt. Die Frage ist nicht länger, ob Daten offen sind; sie lautet, ob Daten von ihrer Anlage her nutzbar sind.

Das strukturelle Problem

ConvergenceAI verfolgt die grundlegende Entwicklung von KI-Ökosystemen in mehr als vierzig afrikanischen Staaten anhand von drei Kernsäulen: politische Aktivität, also die Erfassung nationaler KI-Strategien, sektoraler Politiken und sich wandelnder Regulierungsrahmen; Entwicklung von KI-Fachkräften, also die Kartierung akademischer Bildungswege, universitärer Programme und staatlich geführter Initiativen zur digitalen Grundbildung; sowie KI-Infrastruktur und Innovation, also die Erhebung physischer und digitaler Ressourcen wie KI-Hubs, Innovationslabore, Rechenzentren und Rechenkapazitäten, ergänzt um die Katalogisierung lokal entstandener Innovationen in verschiedenen Sektoren.

„AI Made in Africa“ konzentriert sich auf die Angebotsseite der Daten. Die Initiative katalogisiert Hunderte offener Datensätze aus den Bereichen Landwirtschaft, Gesundheit, Sprache, Klima und weiteren Feldern und bewertet deren KI-Tauglichkeit mittels AODRIA. Dazu zählen Faktoren wie KI-Bereitschaft, Offenheit, Governance und FAIR-Konformität, sozioökonomische Wirkung und Nachhaltigkeit.

Die Überschneidung beider Perspektiven offenbart ein strukturelles Missverhältnis. Das länderübergreifende Monitoring von ConvergenceAI zeigt, dass sich die Reife der Ökosysteme stark auf wenige Zentren konzentriert:

Kenia, Nigeria, Südafrika, Ruanda und Ägypten (siehe Anhang, Abbildung 1). Gewichtiger noch: Die kombinierte Betrachtung beider Plattformen zeigt, dass Länder mit vergleichsweise fortgeschrittener KI-Politik dennoch häufig mit fragmentierten öffentlichen Datenökosystemen zu kämpfen haben. Nationale Strategien betonen vielfach den Einsatz von KI in Landwirtschaft, Gesundheit und Klimaanpassung. Den entsprechenden Datensätzen in diesen Sektoren fehlen jedoch regelmäßig die Vorverarbeitung, Dokumentation oder Interoperabilität, die für praxistaugliches Modelltraining und dessen Einsatz erforderlich wären.

Dies ist weder allein ein Datenproblem noch allein ein Politikproblem. Es ist ein Problem der Integration des Ökosystems.

Sind die verfügbaren Daten nutzbar?

Nutzbarkeit ist keine Ja-oder-Nein-Frage. Ein Datensatz kann technisch herunterladbar und zugleich für jede KI-Pipeline praktisch unbrauchbar sein. Zusammengenommen legen beide Plattformen vier wiederkehrende Fehlermuster offen, die die Nutzbarkeitslücke in afrikanischen Open-Data-Ökosystemen kennzeichnen.

Vier wiederkehrende Fehlermuster

- Fragmentierung – Daten liegen in Silos bei Ministerien, Nichtregierungsorganisationen, Forschungseinrichtungen und privaten Plattformen.
- Mangelnde Aufbereitung – Vielen Datensätzen fehlen grundlegende Beschreibungen, sie verwenden uneinheitliche Metadatenformate oder liegen in rohem, unverarbeitetem Zustand vor.
- Unzugänglichkeit – Selbst offiziell veröffentlichte Ressourcen verschwinden häufig; inaktive Repositorien und tote Links sind weit verbreitet.
- Die Nutzbarkeitslücke – Selbst offene, aktuelle und intakte Datensätze erfordern erhebliche Bereinigung, Annotation und Kontextualisierung, bevor sie verlässliche Modelle speisen können.

Fragmentierung

Daten liegen in Silos bei Ministerien, Nichtregierungsorganisationen, Forschungseinrichtungen und privaten Plattformen, verstreut über unverbundene Portale, wissenschaftliche Datenbanken und Unternehmenswebsites. Da diese Quellen unabhängig voneinander betrieben werden, ist eine Zusammenführung strukturell schwierig: Ein Gesundheitsdatensatz einer Behörde lässt sich womöglich nicht mit Landwirtschafts- oder Klimadaten einer anderen in Einklang bringen, was sektorübergreifende Lösungen wie Modelle für klimaangepasste Landwirtschaft erschwert.

Eine [Erhebung der Weltbankgruppe](#) stellt fest, dass wesentliche Informationen zu Politiken und Rahmenwerken regelmäßig in unkoordinierten digitalen Silos auf Länder- und Sektorebene gefangen bleiben, die auf veralteter Infrastruktur betrieben werden – eine Schwachstelle, die dadurch verschärft wird, dass der Kontinent eine unmittelbare Steigerung der grundlegenden Kapazitäten für Datenhosting und Konnektivität um 75 % benötigt.

Mangelnde Aufbereitung

Vielen Datensätzen fehlen grundlegende Beschreibungen, sie verwenden uneinheitliche Metadatenformate oder liegen in rohem, unverarbeitetem Zustand vor, der für KI-Pipelines ungeeignet ist. Lizenzrechtliche Unklarheit ist verbreitet. Manche Datensätze enthalten keine eindeutigen Nutzungsbedingungen, was jede darauf aufbauende Entwicklung rechtlich riskant macht.

Unzugänglichkeit

Selbst offiziell veröffentlichte Ressourcen verschwinden häufig. Inaktive Repositorien und tote Links sind ein anhaltendes Problem, insbesondere auf Regierungsportalen. Bemerkenswert ist: Selbst unter den afrikanischen Mitgliedstaaten der Open Government Partnership – also Ländern, die sich formell zu Open-Data-Standards verpflichtet haben – zeigt der [Broken-Links-Index der Open Government Partnership](#), dass Regierungen es durchgängig versäumen, die laufende Infrastruktur zu pflegen, die veröffentlichte Ressourcen zugänglich hält. Wenn dies bereits für Länder gilt, die sich Open-Government-Verpflichtungen angeschlossen haben, dürfte die Lücke auf dem übrigen Kontinent nahezu sicher noch deutlich größer sein.

Die Nutzbarkeitslücke

Selbst Datensätze, die offen, aktuell und intakt sind, erfordern erhebliche Bereinigung, Annotation und Kontextualisierung, bevor sie verlässliche Modelle speisen können. Es geht nicht darum, dass jeder Datensatz in ML-fertiger Form veröffentlicht werden müsste. Die derzeitige Kluft zwischen „veröffentlicht“ und „pipelinefähig“ ist jedoch so groß, dass sie alle bis auf die technisch am besten ausgestatteten Akteure systematisch von der Nutzung afrikanischer offener Daten für KI ausschließt.

Beispiele auf Sektorebene

- **Landwirtschaft:** Datensätze zu Ernteerträgen und Witterung werden häufig in unregelmäßigen Abständen erhoben, ohne einheitliche Geodaten-Identifikatoren veröffentlicht oder über getrennte meteorologische, Landnutzungs- und Ertragssysteme hinweg isoliert. Entwicklerinnen und Entwickler können Teilaspekte des Problems erkennen, aber keine Modelle verlässlich auf das gesamte Entscheidungsumfeld trainieren, mit dem Landwirtinnen und Landwirte tatsächlich konfrontiert sind.
- **Gesundheit:** Gesundheitsdaten in Ländern wie Kenia und Südafrika weisen großes Potenzial auf, es fehlen jedoch häufig Standards zur Anonymisierung oder eine mehrsprachige Dokumentation. Datensätze, die diagnostische Anwendungen tragen könnten, sind dadurch entweder rechtlich riskant im Einsatz oder sprachlich zu eng gefasst, um den Bevölkerungsgruppen zu dienen, die sie abbilden.
- **Sprache:** Sprachdatensätze für afrikanische Sprachen sind das deutlichste Beispiel. Ohne ausreichenden Umfang und ausreichende Vielfalt weichen NLP-Modelle auf Englisch oder andere große Weltsprachen aus – mit dem Ergebnis, dass die entstehenden Werkzeuge strukturell an jenen Gemeinschaften vorbeigehen, die sie am dringendsten benötigen.

Wo die Lücken tatsächlich liegen

Die kombinierte Betrachtung beider Plattformen macht mehrere kritische blinde Flecken sichtbar:

Sektorale Entkopplung

Landwirtschaft, Gesundheitsversorgung und Klimaresilienz nehmen in den von ConvergenceAI erfassten nationalen KI-Strategien einen prominenten Platz ein. Frühe Bewertungen von Open Data Africa finden in diesen Bereichen jedoch weniger hochwertige, trainingsfertige Datensätze als bei allgemeineren Wirtschafts- oder Bevölkerungsdaten. Ein Teil der Erklärung liegt in der Art, wie Strategien entstehen: KI-Politiken werden häufig isoliert von Digital- oder Technologieministerien entwickelt, losgelöst von den umfassenderen nationalen Entwicklungsplänen, die von Planungskommissionen und Finanzministerien koordiniert werden. Wenn das finanzielle, regulatorische und infrastrukturelle Gewicht der primären Entwicklungsagenda eines Landes anderswo liegt, bleiben strategische KI-Zusagen in prioritären Sektoren in der Entwurfsphase stecken, statt sich in Mandate für Datenerhebung und -kuratierung zu übersetzen.

Geografische Konzentration

Das Monitoring von ConvergenceAI zeigt, dass sich aktive Rahmenwerksentwicklung und Datensatzverfügbarkeit ganz überwiegend auf wenige Zentren konzentrieren – Kenia, Nigeria, Südafrika, Ruanda und Ägypten (siehe Anhang, Abbildung 1). In weiten Teilen Zentral- und Westafrikas dünnt die Abdeckung deutlich aus; dort verfügen einige Länder über erste Politikrahmen, aber über nur sehr begrenzte Sichtbarkeit offener Datensätze. Das hat kumulative Folgen: Internationale Investoren, Entwicklungspartner und Wagniskapital orientieren sich an positiven Signalen aus dem Ökosystem. Länder ohne sichtbare Dateninfrastruktur erzeugen ein Informationsvakuum, das routinemäßig als hohes Risiko gedeutet wird; Kapital umgeht deshalb systematisch vielversprechende Innovationen in Regionen mit geringer Sichtbarkeit. Das ist nicht nur eine Datenlücke, sondern eine Gerechtigkeitslücke. Eine KI-Entwicklung, die auf jenen Datensätzen aufbaut, die zufällig gut dokumentiert sind, begünstigt systematisch die Bevölkerungsgruppen, die von der bestehenden Infrastruktur ohnehin am besten versorgt werden. Verschärft wird diese Konzentration durch eine sprachliche Sichtbarkeitsverzerrung: Frankophone und lusophone afrikanische Staaten sind in globalen Open-Data-Indizes aufgrund von Sprachbarrieren in der Metadatenberichterstattung systematisch unterrepräsentiert.

Die Sprachlücke

Die starke Abhängigkeit von ausschließlich englischsprachigen Metadaten und Dokumentationen schließt weite Teile der sprachlichen Vielfalt Afrikas aus und begrenzt die kulturelle Passgenauigkeit der Modelle. Weil es für ressourcenarme afrikanische Sprachen an sauberen Daten mangelt, greifen globale Technologiekonzerne häufig auf synthetische Daten für das Modelltraining zurück – was zu kulturellen Halluzinationen führt.

Das Aktualitätsproblem

Viele Datensätze sind gemessen am Veränderungstempo auf dem Kontinent veraltet – etwa Gesundheitsdaten aus der Zeit nach COVID-19 oder sich rasch wandelnde Klimamuster. Veraltete Daten mindern den Wert selbst gut strukturierter Datenbestände für zeitkritische KI-Anwendungen.

Diese Lücken bergen das Risiko, KI-Systeme hervorzubringen, die die Wirklichkeit Afrikas nicht abbilden.

Katalogisierung ist keine Transformation

Eine Plattform, die Ihnen sagt, dass ein Datensatz existiert, sagt Ihnen nicht, ob Sie ihn nutzen können. Ein Ranking, das die KI-Bereitschaft eines Landes bewertet, verbessert diese Bereitschaft nicht automatisch. Katalogisierung ist

notwendig, aber sie ist keine Transformation. Was in den meisten Fällen fehlt, ist die interpretierende Ebene: jene Kontextualisierung, Kuratierung und Annotation, die Rohverzeichnisse handlungsfähig machen. Nutzerinnen und Nutzer auf Daten zu verweisen, ohne Kontext, lokale Besonderheiten, bekannte Verzerrungen, Erhebungsmethoden und dokumentierte Einschränkungen mitzuliefern, birgt das Risiko fehlgeleiteter Anwendungen und fehlerhafter Modelle.

AODRIA ist nicht nur als Ranking konzipiert, sondern als Spiegel für Datenherausgeber – als Mechanismus, der sie dazu anleitet, Metadaten zu verbessern, Dokumentation zu ergänzen und Interoperabilitätsstandards zu übernehmen. Sein Wert liegt weniger in der vergebenen Punktzahl als in dem Verbesserungspfad, den er aufzeigt. Bereitschaft ist in diesem Verständnis eine Praxis, kein Status.

Die Ökosystem-Intelligence von ConvergenceAI geht über reine Dashboard-Sichtbarkeit hinaus, indem sie Zusammenhänge sichtbar macht: welche Politiken mit stärkeren Datenpraktiken einhergehen, wo Innovationskapazität auf Datenverfügbarkeit trifft und wo Investitionen den höchsten entwicklungspolitischen Ertrag versprechen. Ziel ist es, sich von einem Monitoring-Instrument zu einer Analyseplattform zu entwickeln, die staatlichen Akteuren und regionalen Institutionen hilft, strukturelle Lücken zu diagnostizieren, operative Entwicklungspfade zu vergleichen und evidenzbasierte Technologiepolitik zu gestalten.

Gemeinsam arbeiten beide Plattformen auf ein geteiltes Ziel hin – nicht schlicht auf mehr Daten oder bessere Rankings, sondern auf eine Infrastrukturebene, die die Dynamik des Ökosystems mit der Bereitschaft der Datensätze verbindet, und zwar auf eine Weise, die für Politik, Entwicklung und Förderung gleichermaßen handhabbar ist.

Echter Fortschritt verlangt den Übergang vom Prinzip „Open by Default“ zum Prinzip „AI-Usable by Design“.

Empfehlungen

Die Nutzbarkeitslücke ist nicht das Ergebnis eines einzelnen Versäumnisses. Sie spiegelt ein diffuses Bündel von Fehlanreizen im gesamten Datenökosystem wider und lässt sich von keinem Akteur im Alleingang schließen. Ihre Überwindung erfordert einen grundlegenden Wandel darin, wie öffentliche, private und zivilgesellschaftliche Einrichtungen ihre jeweilige Verantwortung verstehen.

Akteur	Empfehlung
Regierungen, Datenherausgeber und Wissenschaft	Technologiepolitik nicht länger als Öffentlichkeitsarbeit betreiben, sondern als zentrale nationale Infrastruktur steuern. Datenverantwortung von Anfang an in Erhebungsmandate einbetten – standardisierte Formate, Metadatenanforderungen und lizenzrechtliche Klarheit. AODRIA-ähnliche Bereitschaftsbewertungen in nationale Open-Data-Portale integrieren.
Entwicklungspartner	Kuratierung, Annotation, Bereinigung und langfristige Datenpflege mit derselben Priorität finanzieren wie die Ersterhebung. Die Förderung mehrsprachiger Dokumentation und kontextualisierender Ebenen ist keine ergänzende Nebenaufgabe; sie ist der Unterschied zwischen Daten, die existieren, und Daten, die genutzt werden können.
Forschende und Entwickelnde	Aktiv beitragen, indem bekannte Probleme dokumentiert, Skripte zur Vorverarbeitung geteilt, Fallstudien zur Nachnutzung veröffentlicht und Verzerrungen oder Einschränkungen gekennzeichnet werden. Dies sind Formen von Infrastrukturinvestitionen, deren Wirkung sich über die Zeit summiert.

Akteur	Empfehlung
Plattformen (ConvergenceAI und Open Data Africa)	Sich zu größerer Interoperabilität zwischen Ökosystem-Intelligence und Signalen zur Datensatzbereitschaft verpflichten und gemeinsam Standards für die fehlende Ebene entwickeln.

Politische Implikationen

Nebeneinander betrachtet machen die Intelligence-Plattform von ConvergenceAI und der AODRIA-Index von Open Data Africa sowohl das enorme Potenzial Afrikas als auch die genaue Beschaffenheit der fehlenden Ebene sichtbar: die Brücke zwischen der Existenz von Daten und ihrem Nutzen für KI.

Dies ist keine Geschichte des Defizits, sondern die eines Ökosystems im Aufbau, das seinen Tritt findet. Die künftige Herausforderung besteht nicht allein darin, mehr Daten zu erzeugen, sondern Ökosysteme zu schaffen, in denen Daten auffindbar, vertrauenswürdig, nachnutzbar, kontextualisiert und fortlaufend verbesserbar sind. Beide Plattformen haben sich verpflichtet, sich im Dienst dieses Ziels zu öffentlicher Infrastruktur weiterzuentwickeln.

Werden diese Lücken nicht geschlossen, wird die KI-Entwicklung auf dem Kontinent weiterhin vorrangig von jenen Datensätzen geprägt, die zufällig gut dokumentiert sind – zum systematischen Vorteil der ohnehin am besten versorgten Bevölkerungsgruppen und unter Umgehung jener Gemeinschaften und Regionen mit dem größten Bedarf und dem größten Potenzial. Regionale Institutionen, darunter die Afrikanische Union, die Afrikanische Entwicklungsbank (AfDB), die ECOWAS und die IGAD, kommt eine entscheidende ermöglichende Rolle zu: beim Setzen von Interoperabilitätsstandards, bei der Finanzierung grenzüberschreitender Datenkuratierung und bei der Koordinierung von Rahmenwerken zur KI-Bereitschaft, die über Strategiepapiere hinausgehen. Diese Anstrengungen sollten auf bestehenden kontinentalen Initiativen aufbauen. So wird die Umsetzung der Kontinentalen KI-Strategie der AU dauerhafte Investitionen in die zugrunde liegende Dateninfrastruktur, in Praktiken der Datenverantwortung und in Interoperabilitätsstandards erfordern, die vertrauenswürdige KI-Systeme erst ermöglichen. Ebenso bietet das Datenpolitische Rahmenwerk der AU (AU Data Policy Framework) eine Grundlage für bessere Daten-Governance, Datenaustausch und grenzüberschreitende Nachnutzung, während die Strategie der AU zur digitalen Transformation Afrikas (2020–2030) den Aufbau digitaler öffentlicher Infrastruktur und von Datenökosystemen als zentrale Voraussetzungen digitaler Transformation hervorhebt. Die Stärkung KI-tauglicher Datenökosysteme stellt damit einen praktischen Weg dar, diese umfassenderen kontinentalen Verpflichtungen voranzubringen.

Wir laden politische Entscheidungsträgerinnen und Entscheidungsträger, Forschende, Entwickelnde und Datenverantwortliche ein, beide Plattformen zu erkunden, unsere Befunde zu hinterfragen und aktiv beizutragen. Die Daten sind vorhanden. Die Strategien nehmen Gestalt an. Jetzt ist der Zeitpunkt, die fehlende Ebene gemeinsam aufzubauen – und vom Prinzip „Open by Default“ zum Prinzip „AI-Usable by Design“ überzugehen.

Glossar

Begriff	Definition
AODRIA	African Open Data Readiness Index for AI (Afrikanischer Index für die KI-Bereitschaft offener Daten). Ein von Open Data Africa / GIZ entwickelter Bewertungsrahmen, der Datensätze anhand von KI-Bereitschaft, Offenheit, Governance, FAIR-Konformität, sozioökonomischer Wirkung und Nachhaltigkeit beurteilt.
FAIR-Konformität	Findable, Accessible, Interoperable, Reusable (auffindbar, zugänglich, interoperabel, nachnutzbar). Ein Satz von Datenprinzipien, der Datensätze maschinell verarbeitbar und über verschiedene Systeme und Kontexte hinweg nachnutzbar machen soll.
Interoperabilität	Die Fähigkeit von Datensätzen oder Systemen unterschiedlicher Herkunft, ohne besonderen Aufwand zusammenzuwirken. Im KI-Kontext bedeutet dies, dass Datensätze über Plattformen und Pipelines hinweg kombiniert oder verwendet werden können, ohne manuell umformatiert werden zu müssen.
Nutzbarkeitslücke	Die Kluft zwischen einem technisch veröffentlichten (offenen) Datensatz und seiner praktischen Verwendbarkeit in einer KI-Pipeline – bedingt durch fehlende Dokumentation, mangelhafte Metadaten, lizenzrechtliche Unklarheit oder unzureichende Vorverarbeitung.
ConvergenceAI Data Intelligence	Eine auf Afrika ausgerichtete Plattform für KI-Ökosystem-Intelligence, die Politik, Innovation und Kapazitätsentwicklung in 54 afrikanischen Ländern erfasst (intelligence.convergenceai.io).
AI Made in Africa / Open Data Africa	Eine Initiative der GIZ bei der Afrikanischen Union, die darauf abzielt, die Voraussetzungen für einen wirksamen Einsatz von KI in der afrikanischen Wirtschaft zu stärken. Open Data Africa ist Teil der Initiative „AI Made in Africa“ und bewertet die Verfügbarkeit, Qualität und KI-Bereitschaft offener Datensätze in Afrika.

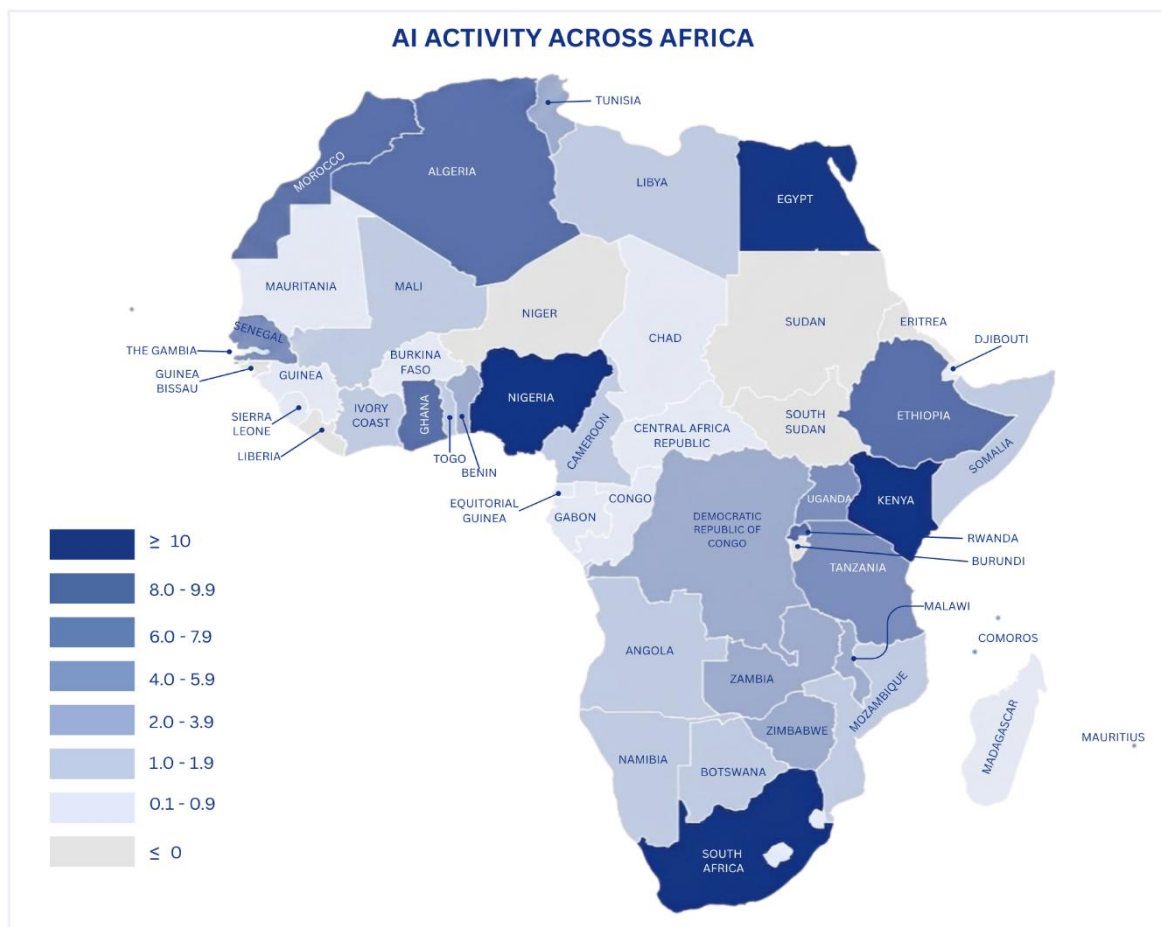
Die Plattformen entdecken

ConvergenceAI: convergenceai.io · [Africa AI Data Intelligence](#)

Open Data Africa und AODRIA: open-data-africa.github.io

ANHÄNGE

Abbildung 1



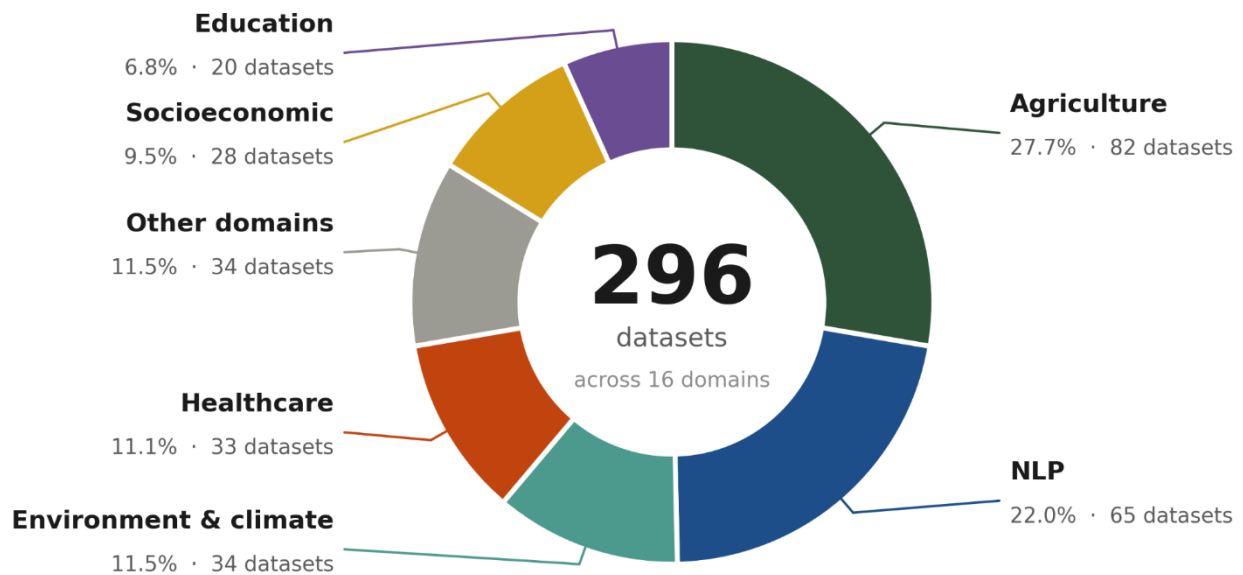
Quelle: [ConvergenceAI Data Intelligence](#)

Dokumentierte Aktivität im KI-Ökosystem nach Land, gemessen am Umfang der auf der Intelligence-Plattform von ConvergenceAI erfassten Datenpunkte zu Politik, Innovation und Kapazitätsentwicklung. Die Darstellung verdeutlicht die starke Konzentration sichtbarer Aktivität auf wenige Zentren neben großen Gebieten, in denen Aktivität nur spärlich oder gar nicht dokumentiert ist; die Werte spiegeln den Dokumentationsgrad wider und stellen kein Ranking der Ökosystemreife dar. Hinweis: Die Beschriftungen der Originalgrafik liegen in englischer Sprache vor.

Abbildung 2

What kind of data is open in Africa?

Domain distribution of cataloged open datasets across the continent



Source: Open Data Africa dataset catalogue, snapshot July 2026 (n = 296 cataloged datasets)

Quelle: Datensatzkatalog von Open Data Africa, Stand Juli 2026 (n = 296 katalogisierte Datensätze).

Verteilung der katalogisierten offenen Datensätze nach Themenbereich. Landwirtschaft und Sprachverarbeitung (NLP) stellen zusammen rund die Hälfte aller erfassten Datensätze, während Gesundheit, Umwelt und Klima – trotz ihrer Priorität in nationalen KI-Strategien – jeweils nur etwa ein Neuntel des Bestands ausmachen.