SCALING AI FOR SMEs

Insights Into South Africa's AI Data Governance Environment

GSMA Mobile for Development

M4D

GSMA

The GSMA is a global organisation unifying the mobile ecosystem to discover, develop and deliver innovation foundational to positive business environments and societal change. Our vision is to unlock the full power of connectivity so that people, industry, and society thrive. Representing mobile operators and organisations across the mobile ecosystem and adjacent industries challenges, the GSMA delivers for its members across three broad pillars: Connectivity for Good, Industry Services and Solutions, and Outreach. This activity includes advancing policy, tackling today's biggest societal issues, underpinning the technology and interoperability that make mobile work, and providing the world's largest platform to convene the mobile ecosystem at the MWC and M360 series of events.

We invite you to find out more at www.gsma.com

This material is funded by the UK government's Foreign, Commonwealth & Development Office (FCDO) through its Global Research and Technology Development portfolio. The views expressed in this report are those of the authors and do not necessarily reflect the official policy or positions of the UK government's Foreign, Commonwealth & Development Office (FCDO) or its partners.

GSMA Central Insights Unit

The Central Insights Unit (CIU) sits at the core of GSMA Mobile for Development (M4D) and produces in-depth research on the role and impact of mobile and digital technologies in advancing sustainable and inclusive development. The CIU engages with public and private sector practitioners to generate unique insights and analysis on emerging innovations in technology for development. Through our insights, we support international donors to build expertise and capacity as they seek to implement digitisation initiatives in low- and middle-income countries through partnerships within the digital ecosystem.

Contact us by email: centralinsights@gsma.com

GSMA Intelligence is the definitive source of global mobile operator data, analysis and forecasts, and publisher of authoritative industry reports and research. Our data covers every operator group, network and MVNO in every country worldwide – from Afghanistan to Zimbabwe. It is the most accurate and complete set of industry metrics available, comprising tens of millions of individual data points, updated daily.

GSMA Intelligence is relied on by leading operators, vendors, regulators, financial institutions and third-party industry players, to support strategic decision making and long-term investment planning. The data is used as an industry reference point and is frequently cited by the media and by the industry itself.

Our team of analysts and experts produce regular thought-leading research reports across a range of industry topics.

www.gsmaintelligence.com
info@gsmaintelligence.com

Authors
Tanvi Deshpande and Emma Leering, GSMA Mobile for Development

Contributors
Ruth Orbach, Ambar Khawaja and Nigham Shahid, GSMA Mobile for Development
Tim Hatt and Silvia Presello, GSMA Intelligence

Acknowledgements
This report draws on research conducted for the GSMA by Axum and the Global Center on AI Governance. We would like to thank Robin Miller (Axum), Alim Ladha (Axum), Mark Gaffley (GCG), Fola Adeleke (GCG), and Rachel Adams (GCG).

We would also like to thank the many individuals and organisations that contributed to the research. A full list of the organisations consulted for the research can be found at the end of the report.

CONTENTS

Acronyms and abbreviations

AAQ	Ask-A-Question
AI	Artificial Intelligence
BaaS	Banking as a Service
DCDT	Department of Communications and Digital Technologies
DFI	Development Finance Institution
DTIC	Department of Trade, Industry and Competition
EU	European Union
FSCA	Financial Sector Conduct Authority
GDP	Gross Domestic Product
GDPR	General Data Protection Regulation
GenAI	Generative AI
GPU	Graphics Processing Unit
ICT	Information and Communications Technology
MNO	Mobile Network Operator
MSME	Micro, Small and Medium Enterprise
NCA	National Credit Act
NDoH	National Department of Health
NLP	Natural Language Processing
PC4IR	Presidential Commission on the 4th Industrial Revolution
POPIA	Protection of Personal Information Act
SME	Small and Medium Enterprise
SMS	Short Message Service
UNESCO	United Nations Educational, Scientific and Cultural Organization

Definitions

AI: Artificial intelligence (AI) is comprised of widely different technologies that can be broadly defined as "self-learning, adaptive systems."[1] AI has the capability to process language, solve problems, recognise pictures and learn by analysing patterns in large sets of data.

Accountability: The expectation that organisations or individuals will ensure the proper functioning, throughout their life cycle, of the AI systems that they design, develop, operate or deploy, in accordance with their roles and applicable regulatory frameworks. Accountability is demonstrated through their actions and decision-making processes.[2]

Data governance: Data governance encompasses technical, policy and regulatory frameworks to manage data along its value cycle – from creation to deletion – and across policy domains including health, research, public administration and finance.[3]

Explainability: Enabling those who have been affected by the outcome of an AI system to understand how it occurred.[4] This involves providing easy-to-understand information to enable them to challenge the outcome and, to the extent practicable, the factors and logic that led to an outcome.

Local language: A language that is spoken within a specific community, region or country, often distinct from the dominant or national language. It may or may not be officially recognised and is typically central to cultural and social identity.

Local language AI: In this report, local language AI refers to AI systems that are designed, trained or adapted to work in local languages. This includes tools and models that understand, generate or translate local languages, making AI more accessible and relevant to speakers of those languages.

Model drift: The degradation of AI model performance due to changes in data or in the relationships between input and output variables.[5] Model drift can result in faulty decision-making and poor predictions.

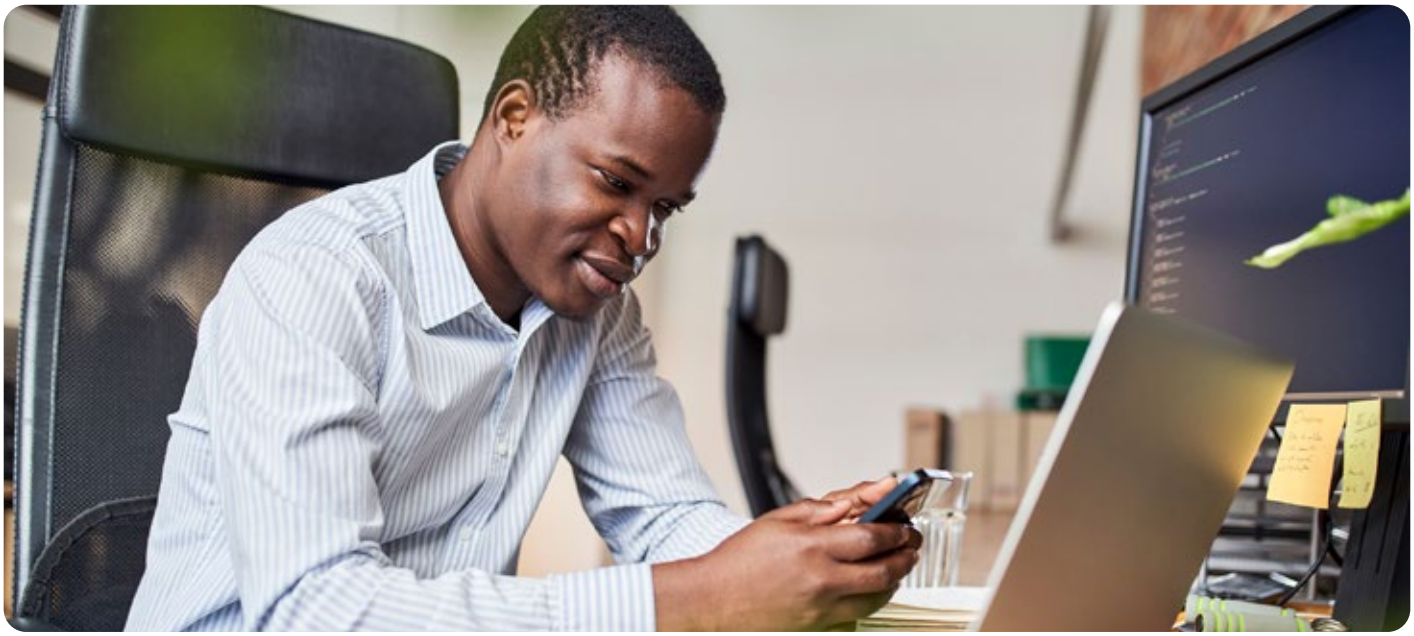
Proportionality: A principle that restricts authorities in the exercise of their powers by requiring them to strike a balance between the means used and the intended aim.[6] In the context of fundamental rights, such as the right to the protection of personal data, proportionality is key for any limitation placed on these rights.

1 International Telecommunication Union (ITU). (2026). "Artificial intelligence for good".
2 Organisation for Economic Co-Operation and Development (OECD) AI principles: Accountability.
3 OECD AI principles: Data Governance.
4 OECD AI principles: Transparency and Explainability.
5 IBM. (2026). "What is model drift?"
6 European Data Protection Supervisor. (n.d.). "Necessity and Proportionality".

List of figures

List of tables

EXECUTIVE SUMMARY



South Africa has one of the most comprehensive data governance environments on the African continent. The use of artificial intelligence (AI) is protected and governed by a constitutional right to privacy, a robust data protection policy and a growing body of AI-related policy frameworks, with a dedicated national AI policy underway. While this policy environment has enabled the steady growth of AI-enabled businesses in the South African economy, small and medium enterprises (SMEs) and startups developing or deploying AI are constrained by several barriers that prevent them from reaping its benefits.

This report explores the legal foundations of data governance for AI in South Africa, describes how SMEs and startups are currently using AI and analyses the key constraints around data governance for AI. SMEs and startups are increasingly undergoing digital transformation, but AI adoption is still nascent and not yet integrated meaningfully into products or business processes. Enforcement of existing data governance policies remains limited, with organisations investing in governance only after they reach a certain size and scale. Smaller enterprises frequently lack the organisational capacity, dedicated roles and resources to meet compliance obligations, leaving them poorly positioned to meet the more demanding requirements of AI-specific data governance.

Capital constraints are a key barrier to investing in data governance processes. SMEs and startups are chronically undercapitalised, and financial resources are directed towards operational activities over governance functions. Upkeep of digital processes and data management requires specialised skills that are costly for early-stage organisations to develop. Compounding this, SMEs also report that they lack sufficient support from the AI ecosystem since there is no ecosystem actor currently dedicated to providing AI governance support to SMEs. This structural absence requires direct investment from intermediary organisations like technology hubs, incubators and accelerators, and industry associations to reach SMEs at the scale the governance challenge requires.

The four case studies in this report highlight other sector-specific data governance constraints. There is a lack of sector-specific guidance on data management practices and standards, leaving organisations to rely on international standards or build their own compliance benchmarks without strong domestic reference points. This directly affects the training data used by AI systems, which can lead to bias and inaccurate representation of last-mile populations that would benefit from impact-focused, AI-led solutions. There are no requirements to evaluate AI model performance after deployment, which can lead to the risk of model drift, divergence from real-world performance and growing bias over time, beyond the scope of existing regulatory or investor oversight frameworks.

The findings point to specific design considerations for AI data governance policy in South Africa. Governance frameworks must ensure AI adoption is equitable and proportionate to the size and maturity of an enterprise and accompanied by practical enabling mechanisms. AI-specific governance must consider factors like data readiness, risk and bias assessment and the appropriateness of AI for a given use case.

As South Africa finalises its National AI Policy, this is a timely opportunity for it to strengthen the national data governance and policy environment. This report outlines how collaborative actions between regulatory bodies, international development organisations and local ecosystem players can help address structural and behavioural constraints and data governance challenges and enable SMEs and startups to adopt AI meaningfully and at scale.

Table 1: Key recommendations

Improve the operationalisation of data governance policies to improve compliance

Enable regulatory sandboxes with special access for SMEs and develop practical, sector-specific data governance toolkits.

Relevant stakeholders: Government departments, institutional regulators, enterprise development organisations, SME associations

Issue AI-specific regulatory guidance

Clarify provisions around automated decision-making, focus on monitoring the performance of AI models and strengthen coordination between regulatory bodies.

Relevant stakeholders: Government departments, institutional regulators, technology hubs, academic and research institutions

Establish formal data access and sharing mechanisms

Enable data sharing between large institutions and SMEs and startups by creating incentives for large players and developing sector-specific data-sharing mechanisms.

Relevant stakeholders: Industry associations, big tech companies, institutional regulators, sector-specific government bodies

Build the AI-related capacity of ecosystem intermediaries

Support ecosystem intermediaries, including technology hubs and startup accelerators, to provide guidance on AI governance to SMEs.

Relevant stakeholders: International development organisations, technology hubs, startup incubators and accelerators, SME associations, industry associations

Leverage market mechanisms to improve data governance compliance

Incentivise SMEs to invest in data governance processes by introducing AI-specific due diligence requirements to access capital and as part of public procurement processes.

Relevant stakeholders: Institutional regulators, development finance institutions (DFIs), capital providers, industry associations, startup accelerators

1. INTRODUCTION

Digital and AI adoption among SMEs and startups in South Africa

Small and medium enterprises (SMEs) are a thriving and significant part of the South African economy, accounting for 91% of formal businesses and a third of the country's GDP.[7] According to a 2024 survey of micro, small and medium enterprises (MSMEs), formal MSMEs in South Africa contribute ZAR 3 trillion (USD 185 billion) to the economy, while informal and township businesses together account for ZAR 4.7 trillion (USD 265 billion).[8]

South Africa also has a growing startup ecosystem that is increasingly technology-oriented and concentrated in sectors including financial services, e-commerce, health and agriculture. The startup ecosystem is geographically concentrated, with formal SMEs and startups clustered in Johannesburg, Pretoria, Cape Town and Durban. There is significant disparity in access to infrastructure, skills and capital between urban and rural areas.

Digital maturity and AI adoption across SMEs and startups is also extremely uneven, with around 30% having adopted digital tools in their operations.[9] In terms of AI adoption, while a small subset of SMEs and startups are deploying AI as a core part of their value proposition, most use AI indirectly through third-party tools and platforms. Even with those that identify as AI-enabled, the quality of AI use lacks depth and meaningful engagement.

Stakeholder interviews revealed that SMEs and startups are producing high volumes of AI-generated outputs without the necessary oversight, governance structures or data complexity to ensure the outputs of AI models are accurate, unbiased and representative of their end users. One reason may be that SMEs and startups believe they must adopt AI or show AI capabilities in their teams to get noticed by investors or donors. While some barriers to AI development, such as poor infrastructure conditions, have improved over the past few years with a reduction in load shedding and increased availability of GPU-as-a-service,[10] other barriers persist, such as high cloud and compute costs (especially for smaller SMEs), gaps in urban-rural connectivity and unclear regulations for cross-border data processing for organisations using international infrastructure.

Approaches to data governance policies in the age of AI

South Africa has a strong policy environment for data governance and was ranked the highest-performing African country in the 2024 Global Index for Responsible AI.[11] However, the existing policy environment does not include explicit provisions for the use of data. The Protection of Personal Information Act (POPIA) covers the core dimensions of data governance with respect to AI: consent, automated decision-making, data quality, cross-border processing and accountability, and its provisions are broadly aligned with international standards. However, there is no AI-specific guidance clarifying how POPIA's provisions apply to AI systems, and no sector-specific codes of conduct for high-risk use cases, such as credit scoring or healthcare support.

Coordination across the multiple regulatory bodies with AI-relevant mandates remains limited, leaving SMEs navigating overlapping obligations without a clear reference point. There are also no SME-specific compliance pathways that translate uniform legal obligations into proportionate operational requirements. For SMEs and startups that play a key role in the country's technology innovation agenda, there is a wide gap between policy and implementation.

This report examines how the current AI regulatory landscape of South Africa shapes the ability of SMEs and startups to develop, adopt and scale AI applications. It provides recommendations to address these constraints and to strengthen the country's AI data governance ecosystem.

7 The Banking Association South Africa. (2025). "SME".
8 FinMark. (2024). FinScope MSME Survey South Africa 2024.
9 Profitshare Partners. (2025). "How AI is opening the funding floodgates for South Africa's SMEs".
10 Graphics processing units, which are necessary in the training and deployment of AI models.
11 UNESCO. (2025). Global AI Ethics and Governance Observatory: South Africa.

2. RESEARCH OBJECTIVES AND METHODOLOGY

2.1 Research objectives and methods

This report aims to explore how AI data governance mechanisms can be strengthened for SMEs and startups operating in South Africa. It unpacks how data governance impacts different stages of the AI life cycle, from generating data to developing, deploying and scaling AI applications, with a focus on creating sustainable socio-economic impact in the country.

The main methods used for this research are outlined below, and a list of stakeholders consulted for this research can be found in Annex 1.

- **Desk research:** A literature review of the current data regulatory and innovation landscape for AI in South Africa and relevant international data governance strategies.
- **Key informant interviews (KIIs):** Interviews with AI-enabled startups, investment agencies and donors, research organisations and think tanks, capacity building organisations and government agencies.
- **SME survey:** Quantitative survey of 200 SMEs across multiple sectors to understand their use of AI and the key data governance constraints faced by AI-enabled SMEs.
- **Roundtable discussion:** Multistakeholder discussion with government representatives, founders of AI startups, health companies and fintechs, investors and policy research organisations to validate research findings and identify gaps for further study.

In addition, the report highlights four case studies in the priority sectors of health and financial inclusion. These case studies examine various governance constraints faced by startups, as well as data access partnerships and data management practices, to present overarching considerations for data governance in South Africa.

Definitions of SMEs and startups

This report treats startups and small and medium enterprises (SMEs) as distinct but related segments within South Africa's innovation and enterprise ecosystem. Startups are defined as early-stage, innovation-driven, formally registered enterprises developing scalable, tech-enabled products, services or business models. SMEs are formally registered enterprises that may or may not be innovation-led, but play a critical role in employment, service delivery and sectoral value chains. Despite differences in maturity, risk profile and growth trajectory, both groups face overlapping challenges when developing or adopting AI: accessing data, meeting regulatory compliance obligations, managing reliance on infrastructure and building organisational capacity. This shared experience among South African SMEs and startups is the basis for treating them as the primary unit of analysis in this report.

Priority sectors for the study

The two priority sectors selected for this study are health and financial services. Selection was based on criteria detailed in Annex 2, with both sectors chosen because AI adoption is already underway and because potential for impact, data complexity and data sensitivity are all high.

- **Healthcare** is a priority impact sector given the potential for AI-enabled solutions to improve diagnostics, access to service delivery and resource allocation, including addressing specialist shortages and extending coverage to rural populations. It also has the most complex governance challenges, spanning lawful basis for processing, consent, secondary data use, security safeguards and automated decision-making risks.
- **Financial services (fintech)** offers high potential for impact through expanded access to payments, credit and financial services for underserved and excluded groups. The sector involves significant data concentration, with high-value transactional and behavioural data often controlled by regulated incumbents and large platforms, making data access and governance accountability particularly acute challenges for SMEs.

2.2 Survey methodology

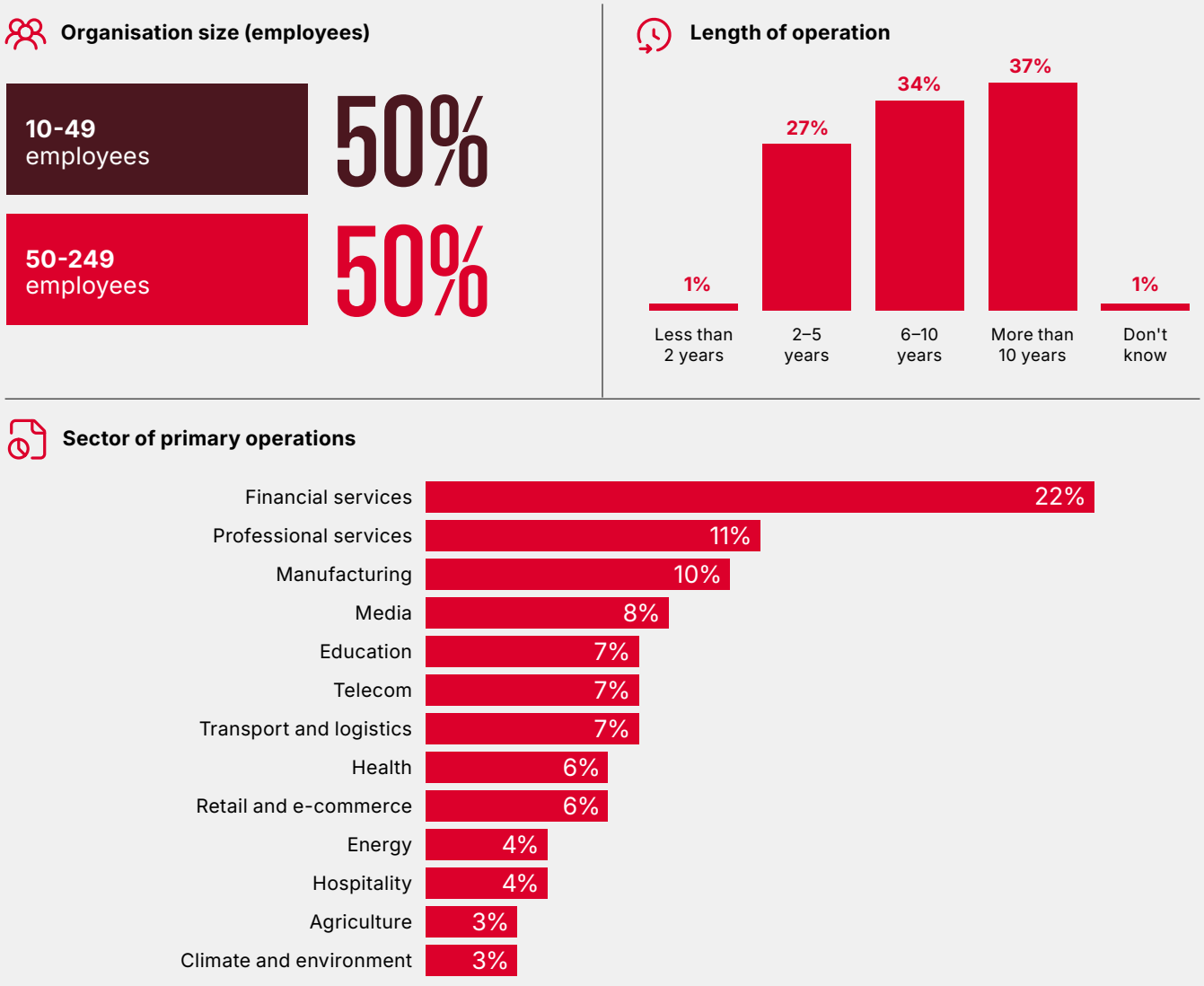
In February 2026, GSMA Intelligence (GSMAi) surveyed 200 director-level and above professionals working for South African organisations with between 10 and 249 employees. The survey draws on a representative sample of formal SMEs to assess AI adoption, data governance and policy in South Africa. All respondents had direct responsibility for the management of data for AI, and all organisations were required to be actively using AI in their operations, services or products at the time of the survey.

The sample is evenly split between small (10–49 employees) and medium (50–249 employees) enterprises, and spans a range of sectors including financial services, manufacturing, agriculture and telecoms, among others, ensuring coverage of South Africa's main areas of economic activity. More than 70% of surveyed organisations have been in operation for at least six years, with 37% operating for more than 10 years.

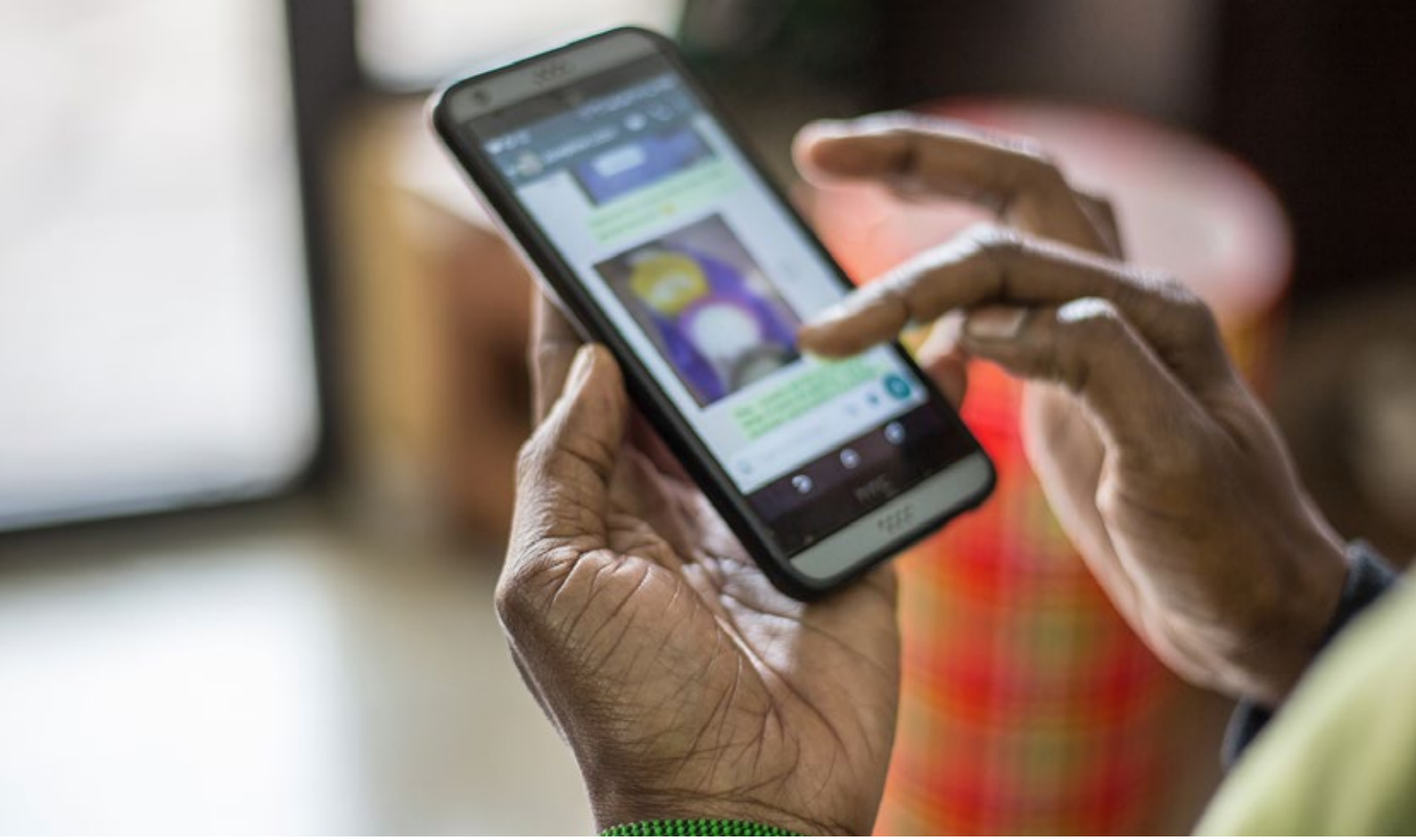
The survey is intentionally confined to organisations with existing AI engagement. As such, the findings reflect the experience of AI-active SMEs and do not represent the broader SME population, which includes businesses not currently using AI. Findings should be interpreted accordingly – statements such as "Most SMEs in South Africa have a strong capacity to invest in AI" apply specifically to this subset. Where sector- or deployment-specific analyses are made, smaller sub-sample sizes may affect the depth of analysis and should be read with this caveat in mind.

Figure 1

Profile of AI-enabled SMEs covered in the GSMAi survey



N=200
Source: GSMA AI Governance 2026



2.3. Limitations of this research study

This study focuses on formally incorporated startups and SMEs that are already engaged with South Africa's formal innovation ecosystem and operating within the regulatory and institutional structures this report examines. The findings are predominantly drawn from urban-based organisations and do not extend to township-based innovators or informal sector enterprises. The startup case studies are among the most governance-conscious organisations in their respective sectors. The gaps that persist here are almost certainly more acute across the broader SME and startup population, where capacity is lower and institutional support is absent.

Township-based and informal sector actors represent a significant and growing dimension of South African AI activity, building solutions adapted to informal markets, low-connectivity environments and non-English-speaking populations. However, the governance challenges they face around data access, infrastructure, regulatory exposure and ecosystem support are materially different from those this report examines, and their experience is not captured in the primary research.

Existing research shows that AI adoption rates among informal businesses in South Africa are minimal but evolving, and further research on this segment would provide a more complete picture of AI governance constraints in the economy. Data findings and insights from this report can also shed light on how strengthening data governance processes may help small businesses increase meaningful AI adoption.

3. AI DATA GOVERNANCE ENVIRONMENT IN SOUTH AFRICA

3.1 Key actors in the AI data governance ecosystem

South Africa's AI data governance ecosystem comprises government departments, independent statutory regulators and a range of research, civil society and academic institutions. Government and regulatory bodies hold oversight responsibility for AI-related policy development and implementation. Research and academic institutions contribute evidence on the impact of digital technologies, including AI, on the economy, and are increasingly active in developing governance frameworks for responsible AI use.

Government departments

South Africa's AI data governance agenda is led at the national level by the Department of Communications and Digital Technologies (DCDT), which published the South Africa National AI Policy Framework in August 2024. The Department of Science and Innovation (DSI) plays a complementary role through research funding, innovation strategy and oversight of bodies such as the Council for Scientific and Industrial Research (CSIR) and the National Research Foundation (NRF).[12] Sector departments include the Department of Health, Department of Trade, Industry and Competition (DTIC) and the National Treasury, which together shape AI-relevant policy. Coordination across these departments remains a structural challenge; fragmentation across ministerial mandates and insufficient cross-departmental collaboration are mentioned as persistent weaknesses in South Africa's approach to AI data governance.[13]

Independent statutory regulators

Several independent statutory bodies hold regulatory functions directly relevant to AI. The Information Regulator, established under POPIA and accountable to the National Assembly, is the primary oversight authority for data protection and privacy, including enforcement of POPIA's automated decision-making provisions.[14] The Competition Commission South Africa (CCSA) examines how AI and algorithmic systems affect market dynamics and the effects of ranking algorithms on competition and freedom of expression.[15] The Financial Sector Conduct Authority (FSCA) and the National Credit Regulator (NCR) exercise oversight in financial services, where AI use in credit scoring, insurance and payments intersects with both consumer protection and data governance obligations.[16]

Research, civil society and academic institutions

A range of research and civil society institutions contribute to South Africa's AI data governance ecosystem. The AI Institute of South Africa (AIISA), established in response to the Presidential Commission on the 4th Industrial Revolution (PC4IR), operates through hubs at the University of Johannesburg, Tshwane University of Technology and the Central University of Technology, and is mandated to foster ethical AI research and develop national AI talent.[17] The Centre for Artificial Intelligence Research (CAIR) and the Centre for the Fourth Industrial Revolution (C4IR) contribute further to knowledge generation and applied AI development.[18] Research ICT Africa provides independent policy research and evidence on the digital economy, including on AI governance gaps. Several universities have developed institutional frameworks for responsible AI use, including the University of Cape Town, whose Senate Ethics Committee issued guidelines on generative AI (GenAI) for research.[19]

Regionally, the African Observatory on Responsible Artificial Intelligence provides benchmarking and coordination across the continent.[20] A report by UNESCO notes that while this ecosystem is comparatively active by African standards, coordination between these actors and formal regulatory processes remains limited, and the intermediary infrastructure needed to translate research into practical governance guidance for smaller enterprises is largely absent.[21]

12 Department of Communications and Digital Technologies (DCDT). (2023). South Africa's artificial intelligence (AI) planning.
13 UNESCO. (2025). South Africa artificial intelligence readiness assessment report.
14 See POPIA, section 71.
15 UNESCO. (2025). South Africa artificial intelligence readiness assessment report.
16 Ibid.
17 Ibid.
18 Ibid.
19 Ibid.
20 See: African Observatory on Responsible AI website.
21 UNESCO. (2025). South Africa artificial intelligence readiness assessment report.

3.2 Policies governing AI use in South Africa

South Africa's AI governance environment is shaped primarily by general data protection laws rather than AI-specific regulation. The constitutional right to privacy provides the foundational basis, operationalised through the Protection of Personal Information Act (POPIA), which is the primary legal instrument governing data collection, processing, storage and sharing across sectors. POPIA addresses several aspects relevant to AI, including restrictions on automated decision-making with legal or significant impacts, requirements for documented accountability and security measures, and controls on cross-border transfers of personal information.

Sector-specific legislation, including the National Credit Act, Consumer Protection Act, Electronic Communications Act and Cybercrimes Act, governs aspects of AI-relevant data use across financial services, health and telecommunications. Together, these instruments establish a broad compliance baseline, but none were designed with AI systems in mind. Key governance questions around algorithmic bias, explainability and model accountability remain either unaddressed or inapplicable to nascent AI contexts.

South Africa is nonetheless among the more advanced African countries in AI regulation. Following the PC4IR and a National AI Summit in April 2024, the National AI Policy Framework was published and the National Policy on Data and Cloud was finalised, addressing related questions of data infrastructure and localisation. The key challenge identified through stakeholder interviews is that South Africa's policy framework does not yet function as an enabling environment for AI in practice. Policy instruments remain largely strategic: no AI-specific regulatory guidance has been issued, no sector codes of conduct exist for automated decision-making, and POPIA's application to AI systems remains interpretively unsettled. A comprehensive mapping of relevant legal and regulatory instruments is provided in Annex 3.

Stakeholders consulted during the research process have highlighted that the enforcement of policies around data protection and AI has been uneven. Coordination across the Information Regulator, sector regulators and the DCDT remains limited, and SMEs frequently face uncertainty about which regulator's guidance applies to a given AI use case. Maximum statutory penalties are up to ZAR 10 million (USD 0.6 million) under POPIA, but the absence of enforcement capacity creates an environment of uncertainty.[22]

Despite the breadth of South Africa's data governance framework, several structural gaps constrain SMEs and startups engaging with AI. These are examined in depth in section 5, which draws together findings from the survey, stakeholder interviews and case studies to present the key constraints and their policy implications.

22 POPI Act Compliance: Offences, Penalties, and Administrative Fines.

4. AI-ENABLED DATA USE IN PRACTICE

SMEs and startups engage with data across several stages of the AI life cycle depending on their modes of data sourcing, model design and initial deployment, as well as their level of AI adoption. When mapped onto the AI data life cycle, engagement typically occurs in three primary forms: 1) data holders, 2) data processors and 3) data deployers (Table 2). Each form has distinct data governance obligations and degrees of leverage in the ecosystem. It is also important to acknowledge a fourth category – data subjects – which are the persons or entities that the data is about, such as medical patients, business customers, payment platforms or social messaging app users.

Table 2: Modes of engagement with AI and data

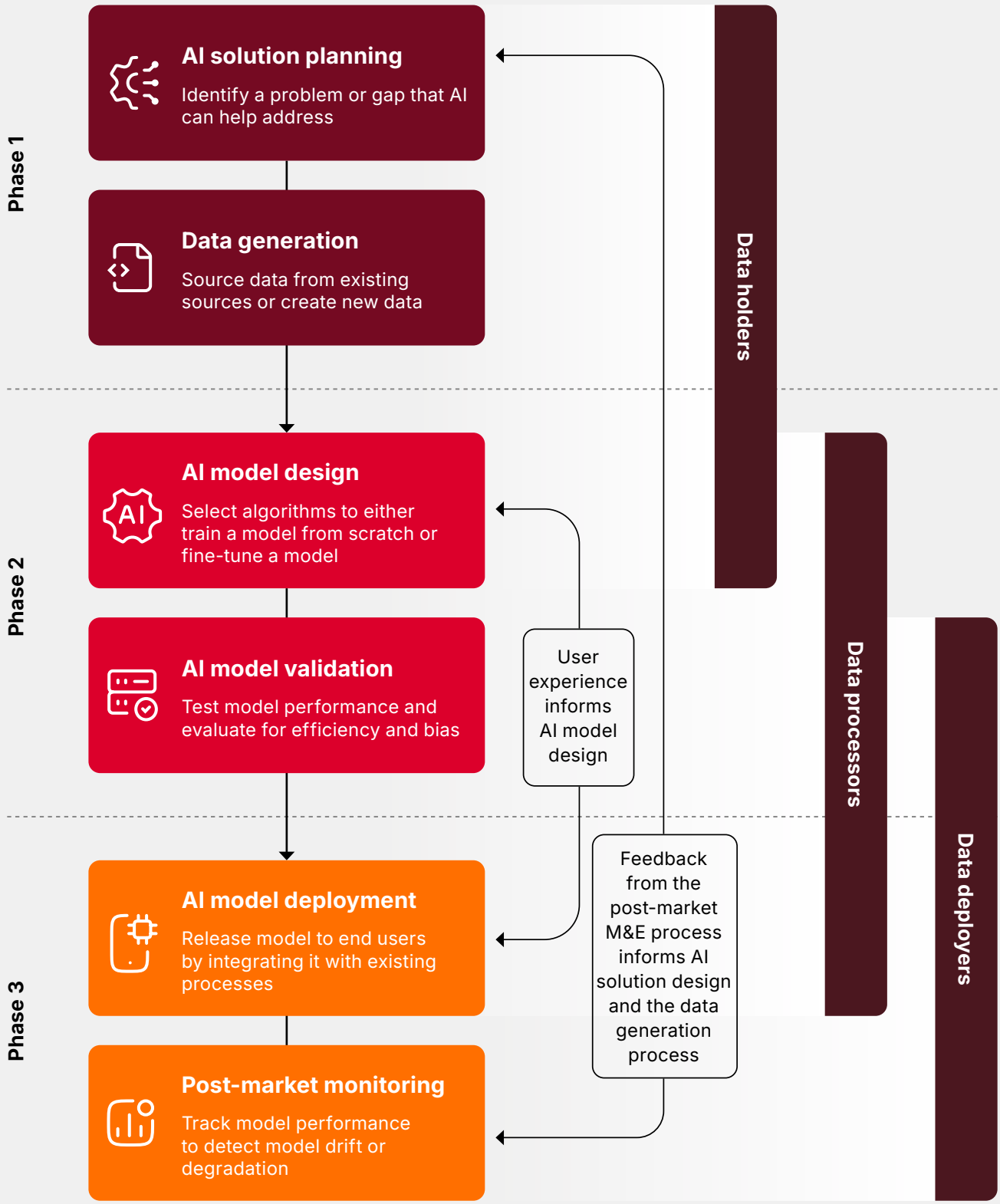
	Data holders	Data processors	Data deployers
Description	Large institutions that generate or collect primary data from users, devices or the environment.	Organisations that process, clean or add value to data to build AI models and systems.	Companies that use data and AI models to deliver products or services.
Examples	Financial institutions, MNOs, public health systems and public hospitals, government departments, etc.	Startups, academic institutions, tech hubs.	Established SMEs, fintech providers, digital health or education service providers.
Ecosystem position	Make the most consequential AI data governance decisions by setting the terms such as pricing, partnerships and contractual sharing agreements.	Depend on data holders for access to foundational data and deliver AI models to deployers. Face constraints like capital and governance capacity.	They are closest to the end users of AI models and do not have control over data collection or AI model design, which may create constraints for governance compliance.

Source: Authors' analysis.

The AI data life cycle (as seen in Figure 2) describes the process of how data is gathered, processed, designed into AI models and deployed in the market for end users. Data holders are linked to the first stage of problem identification and data generation or collection, and act as custodians of the data, responsible for its safekeeping and sharing. Data processors are involved in cleaning and formatting the data to design and develop AI models and then evaluating and validating the AI models for their defined purpose and end users. Data deployers engage with data at the end of the value chain, mainly through deploying pre-built AI models for their own goals or engaging as customers of finished AI products. The final stage of the AI data life cycle is post-market monitoring and maintenance, where feedback loops back into the initial stage of data identification and collection, contributing to AI solution design. The case studies in section 4.4 illustrate how data moves across these stages in the healthcare and financial services sectors.

Figure 2

The AI data life cycle



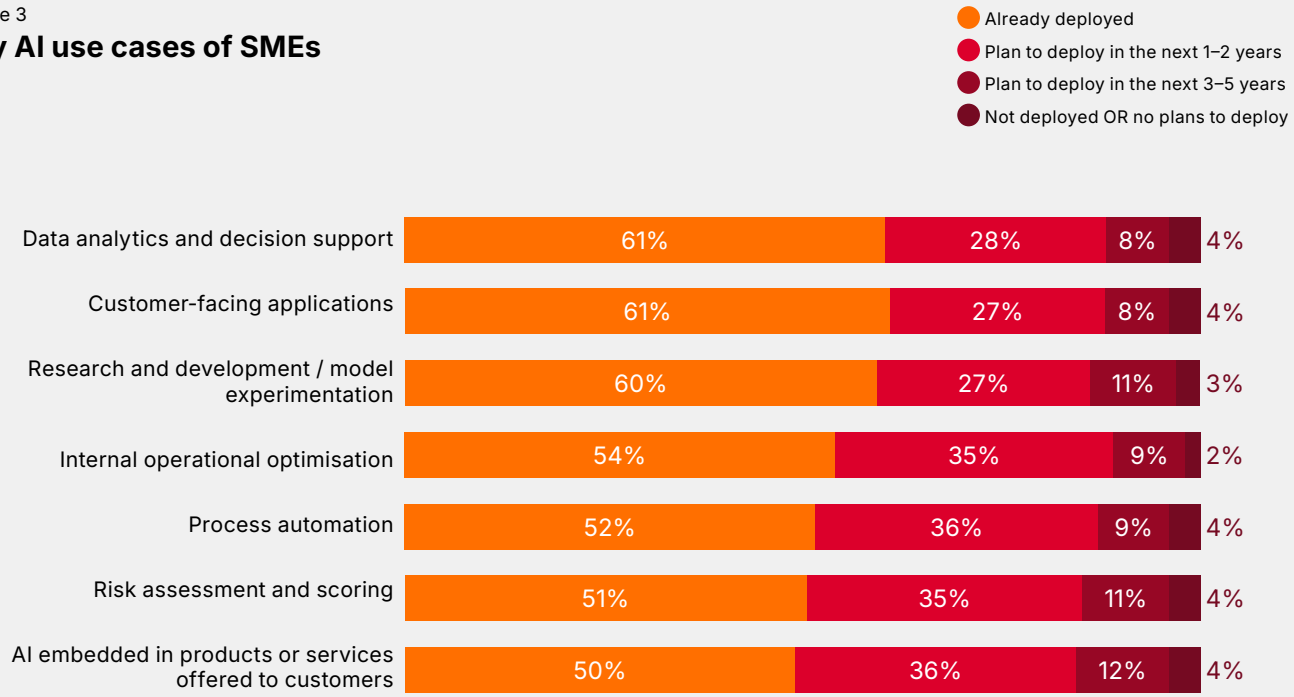
Source: Authors' analysis.

4.1 AI adoption by SMEs and key use cases

AI adoption among surveyed SMEs is broad, with 61% of SMEs having already deployed data analytics and customer-facing applications (see Figure 3). Each category of use cases shows at least 50% already deployed, with another 27–36% planning deployment within one to two years.

Figure 3

Key AI use cases of SMEs



N=200
Source: GSMA AI Governance 2026

For governance, this breadth of AI use shows that SMEs are already managing consent, data sourcing and automated decision-making obligations across multiple use cases, without the governance pathways that would allow for more staged adoption. Government intervention would be beneficial in several ways. First, it would create a more level playing field for SMEs and startups that lack the capital to build data governance processes on their own, reducing the current compliance burden. It could also lower the costs of implementing policies for SMEs by making ready-made government frameworks available. Finally, it would be easier to bring in responsible AI principles to be followed as an obligation as part of the enforcement of data governance policies.

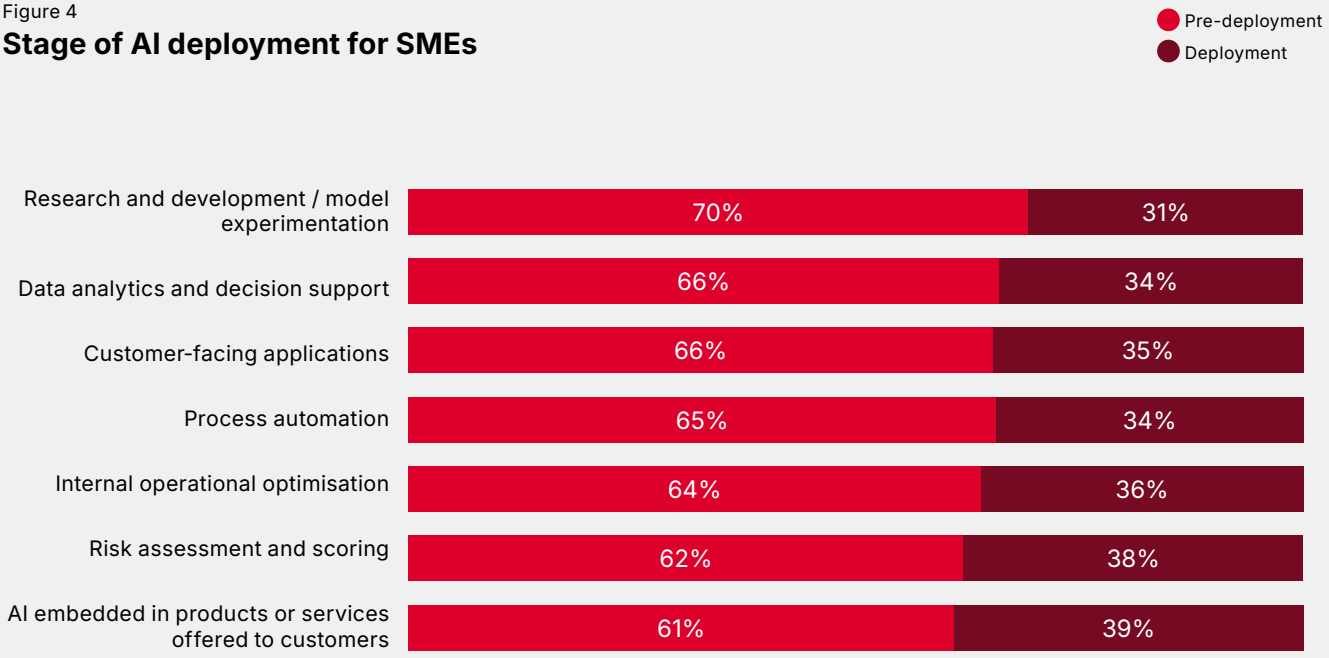
Across all use cases, most surveyed organisations (between 61% and 70%) are still in pre-deployment stages, with deployment rates ranging from just 31% to 39% (see Figure 4). Organisations in the hospitality, transport/logistics, professional services, health and education sectors are most advanced in the deployment of AI-enabled solutions. Further, SMEs in the research and development/model experimentation stage show the highest pre-deployment activity at 70%, while AI embedded in products or services has the highest deployment rate at 39%.

This pattern suggests that AI is currently in the development and experimentation stage, which means that governance decisions being made now around consent, data sourcing, model design and testing will determine the compliance behaviours of organisations creating AI systems that have not yet reached the market.

Based on this data, addressing governance gaps at the pre-deployment stage may be more efficient than retrofitting compliance to systems already in use. Interventions targeting pre-deployment governance are therefore more likely to be effective than those focused on correcting practices after deployment.

Figure 4

Stage of AI deployment for SMEs

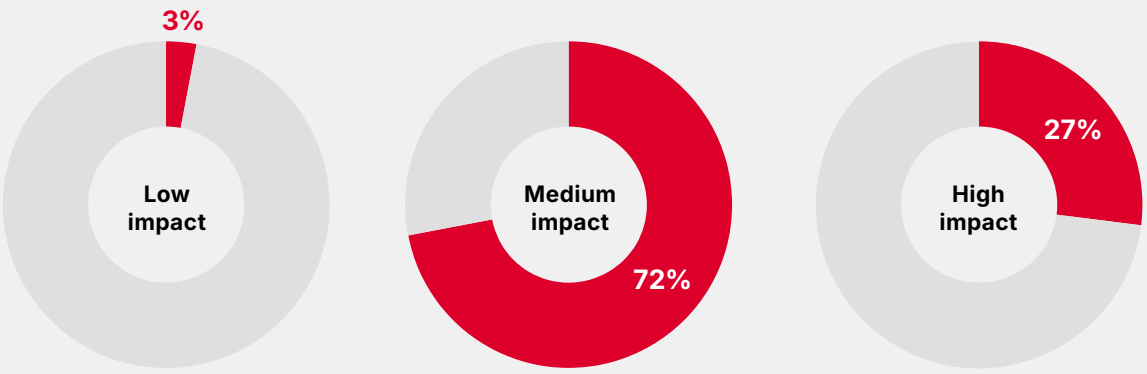


N=200
Source: GSMA AI Governance 2026

Assessing their existing AI deployment, 72% of surveyed organisations fall into the medium impact[23] category, with 27% achieving high impact and only 3% reporting low impact (The AI Impact Index is explained in Annex 4). This finding indicates that most organisations have adopted and deployed AI but have not yet embedded it deeply enough to produce transformative business outcomes. This gap between deployment and impact is consistent with the finding that most AI use cases are still in pre-deployment or early deployment stages, and reinforces the case for governance interventions that support organisations to deepen AI integration responsibly rather than simply accelerating initial adoption.

Figure 5

AI Impact Index for surveyed SMEs



N=200
Source: GSMA AI Governance 2026

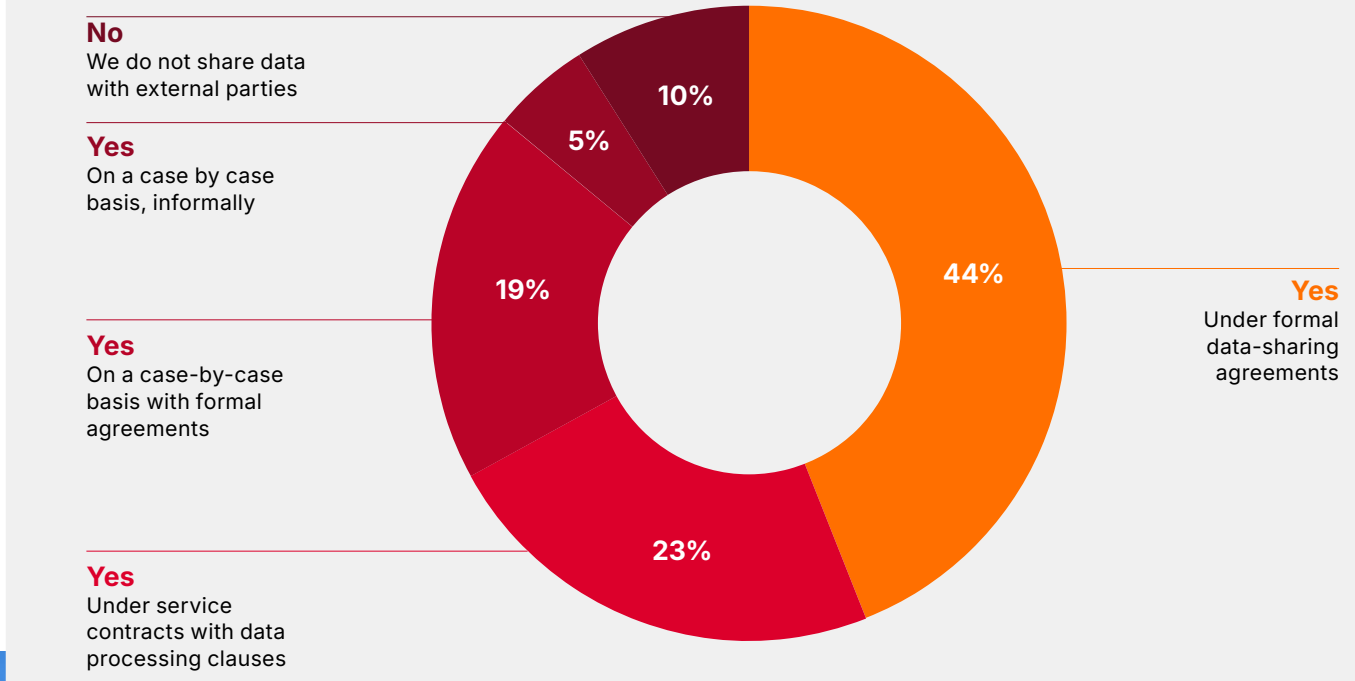
23 AI impact in this context measures the effect of AI on an organisation, including its capacity to invest in AI, the extent to which AI is adopted and applied and the deployment status of AI use cases within the organisation.

4.2 Data-specific insights and challenges

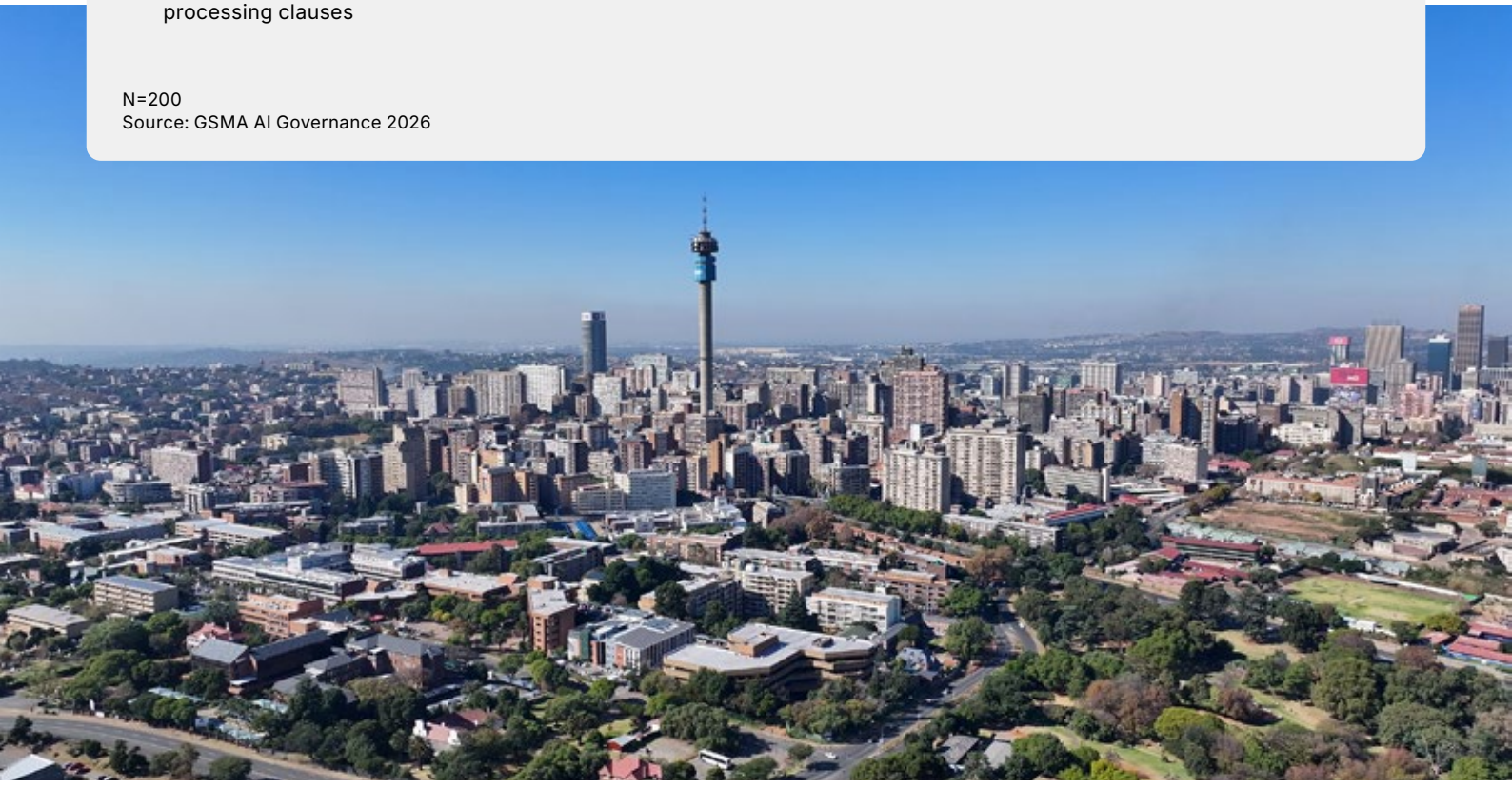
In terms of formal data sharing, 44% of surveyed organisations share data under formal agreements, and a further 23% do so under service contracts (see Figure 6). However, the existence of formal agreements does not indicate that those agreements address AI-specific uses. Across the four case studies examined in section 4.4, consent frameworks and data-sharing contracts were consistently found to predate the AI applications subsequently built on that data, meaning that formal arrangements exist, but their coverage of AI-relevant secondary uses is largely undocumented.

Figure 6

Data-sharing practices of SMEs with external organisations



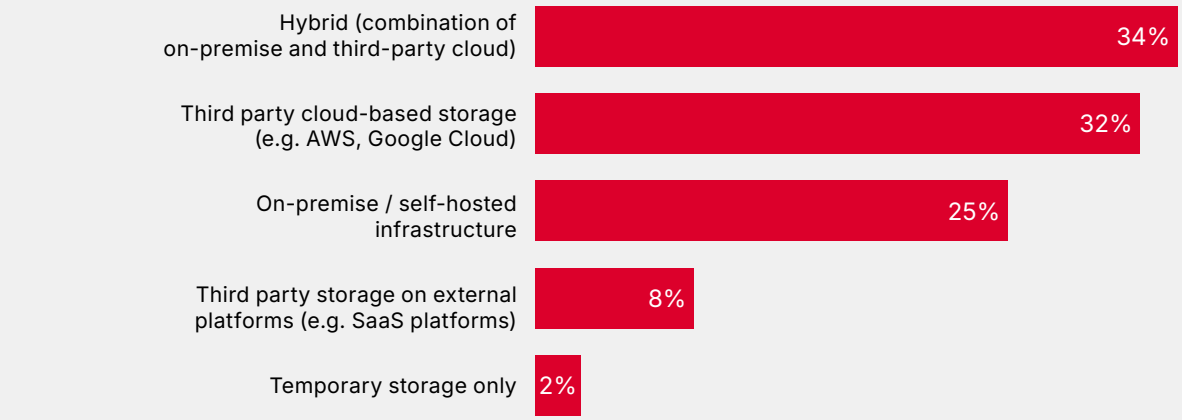
N=200
Source: GSMA AI Governance 2026



As seen in Figure 7, data storage among surveyed SMEs is split across hybrid infrastructure (34%), third-party cloud providers such as AWS and Google Cloud (32%) and on-premise/self-hosted infrastructure (25%). Two-thirds of the organisations are therefore using at least some third-party or cloud-based storage, much of it hosted on international infrastructure. Under POPIA, cross-border transfers of personal information require either an adequacy determination or a data transfer agreement, neither of which most SMEs using international cloud providers are likely to have fully documented. This means that a significant proportion of surveyed organisations are likely operating with unresolved cross-border data processing obligations because of using the most commercially accessible infrastructure available.

Figure 7

Dominant approach to data storage among SMEs

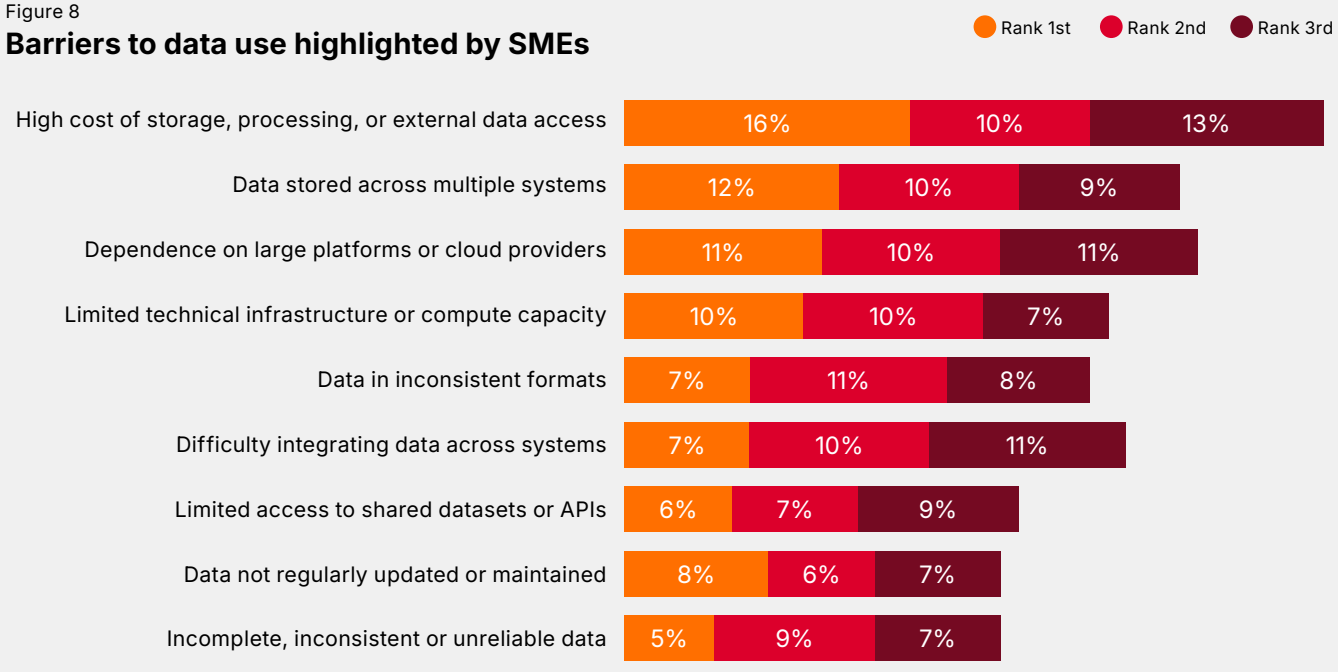


N=200
Source: GSMA AI Governance 2026

The leading barrier to data use is the high cost of storage, processing or external data access. This was cited as the top barrier by 16% of respondents and within the top three barriers by 39% of respondents (see Figure 8). Data stored across multiple systems follows at 31%, reflecting the fragmented data environments in which most SMEs operate. Following these barriers are dependence on large platforms or cloud providers and limited technical infrastructure or compute capacity, which together point to a structural problem that is primarily financial and infrastructure-based, rather than regulatory.

Figure 8

Barriers to data use highlighted by SMEs



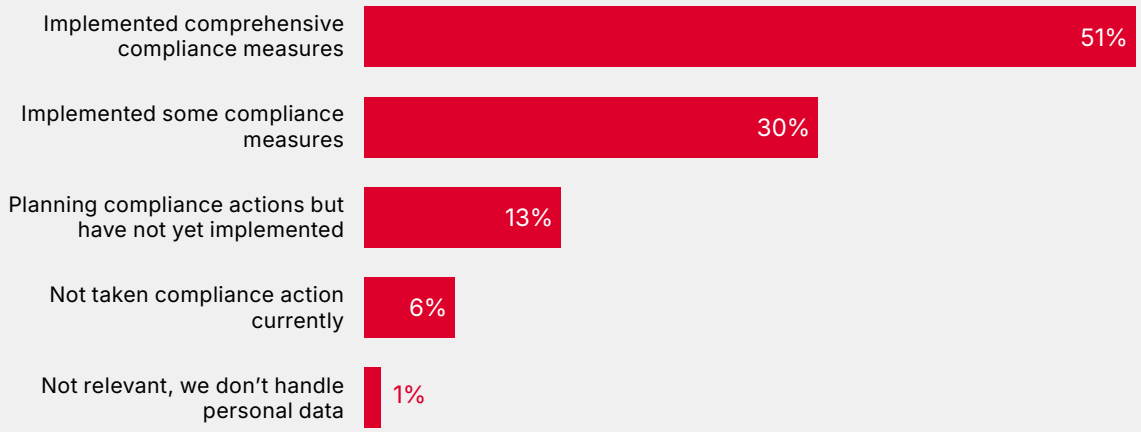
N=200
Source: GSMA AI Governance 2026

4.3 Data governance practices

Just over 80% of SMEs surveyed claim varying levels of POPIA compliance measures (see Figure 9). For organisations already using AI across multiple functions, partial compliance means that the data practices underpinning active AI systems may not meet the requirements for consent, security safeguards or automated decision-making accountability that POPIA imposes.

Figure 9

SME organisational compliance with POPIA

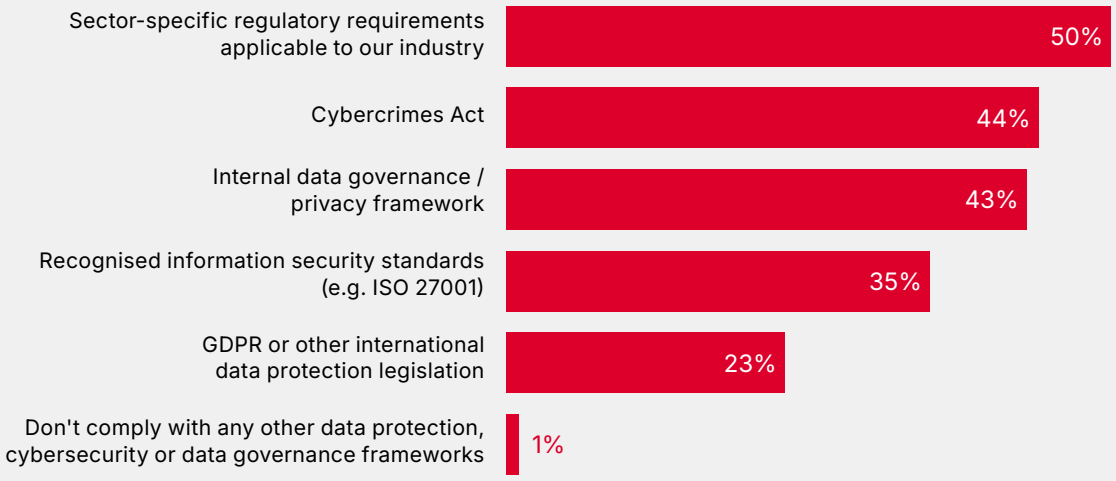


N=200
Source: GSMA AI Governance 2026

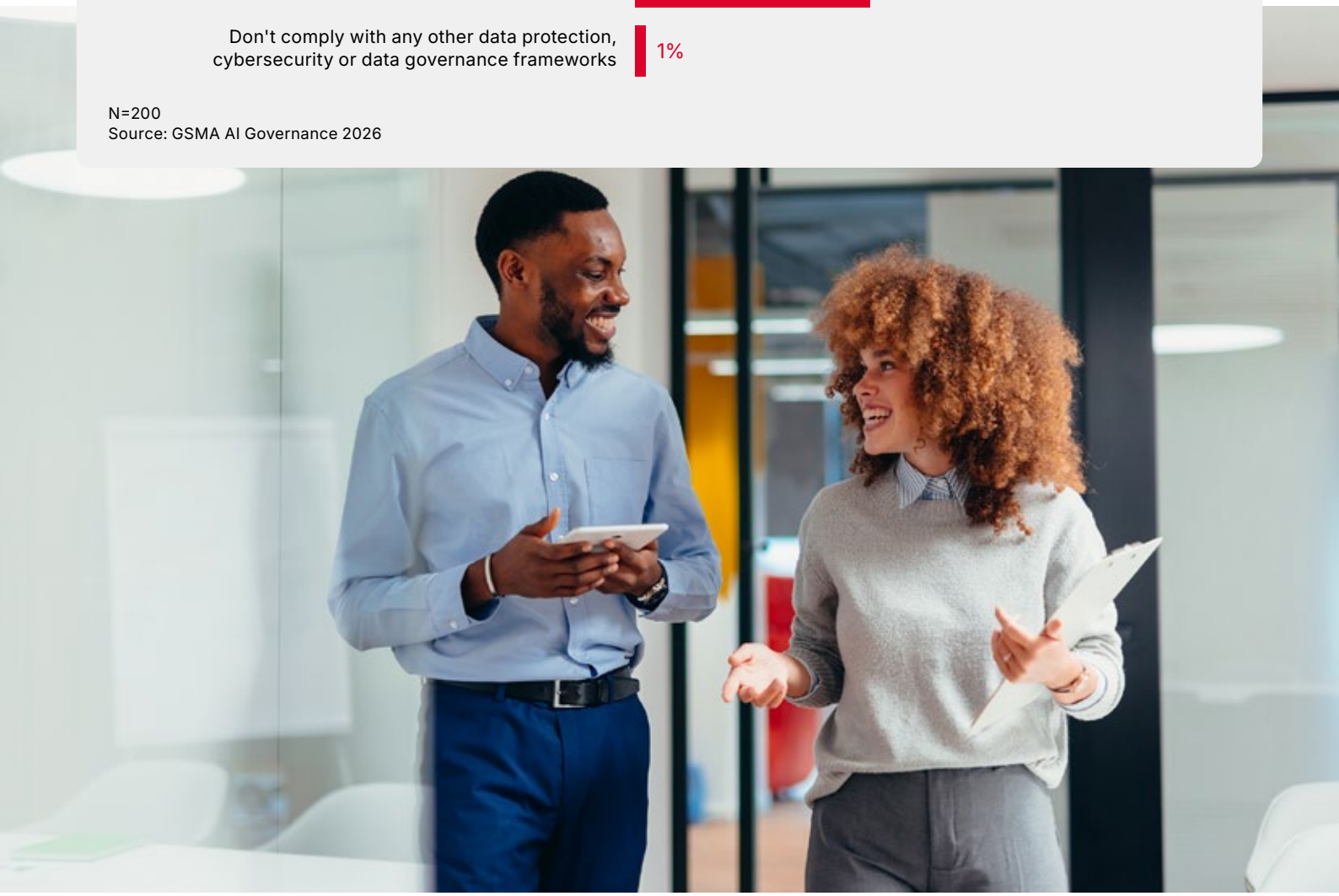
Beyond POPIA, the survey indicates that sector-specific regulatory requirements are the most adhered-to framework at 50%, followed by the Cybercrimes Act at 44% and international data governance or privacy frameworks at 43% (see Figure 10). Recognised information security standards are in use by 35% of surveyed SMEs, and 23% comply with the General Data Protection Regulation (GDPR) or other international data protection legislation, reflecting the proportion of SMEs with EU-facing operations or international investor requirements. Only 1% report complying with no additional frameworks. This data reveals layered compliance, which is consistent with the fragmented regulatory landscape where SMEs are navigating multiple overlapping frameworks simultaneously, without a coordinating interface that clarifies interactions or conflicts across different legislation.

Figure 10

SME adherence to sector-specific regulations (in addition to POPIA)



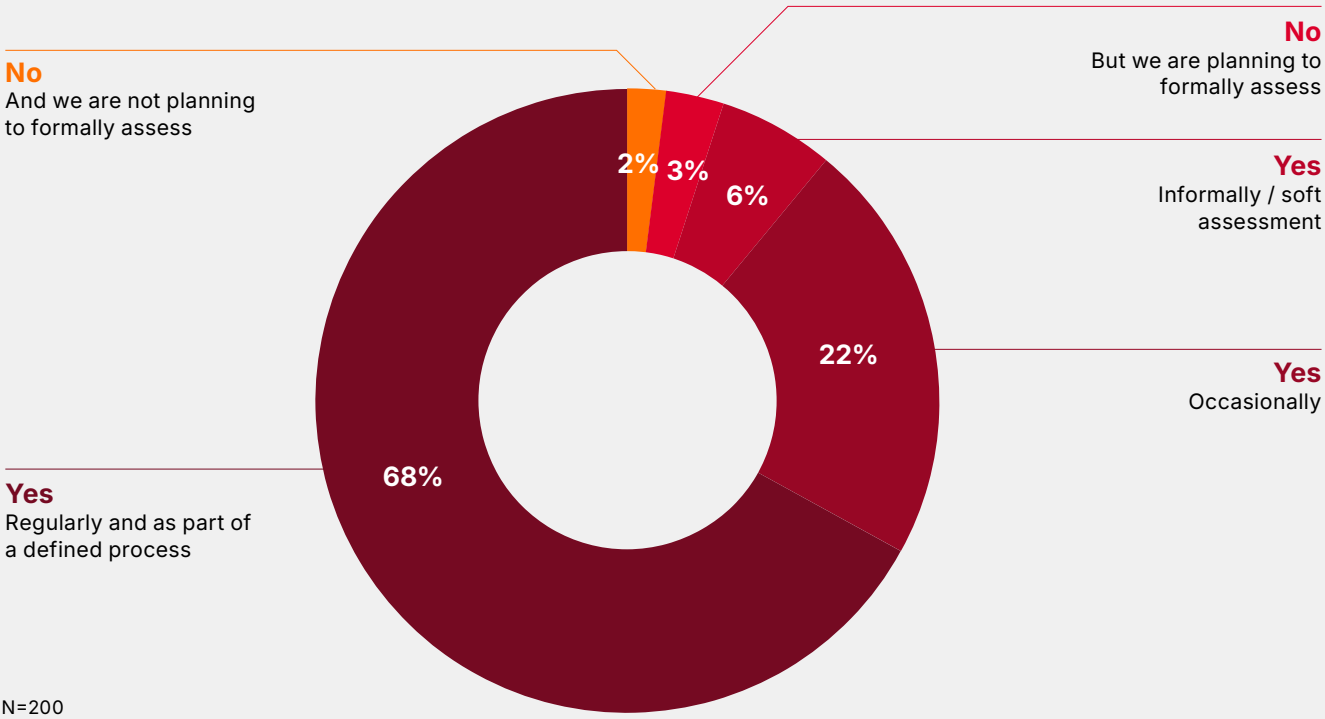
N=200
Source: GSMA AI Governance 2026



Many surveyed organisations with dedicated compliance budgets assess AI risks regularly and as part of a defined process, while informal and occasional assessment accounts for a smaller share and a minority of SMEs report no current plans to formally assess (see Figure 11). The data suggests that among AI-active SMEs that have made a financial commitment to compliance, structured risk assessment follows the initial investment. This suggests that the AI data governance problem in South Africa is fundamentally a financing problem, and that interventions targeting compliance culture or awareness are likely to be less effective than those that address the capital constraints preventing governance investment from occurring at all.

Figure 11

Proportion of SMEs with high AI capacity that assess AI risks regularly



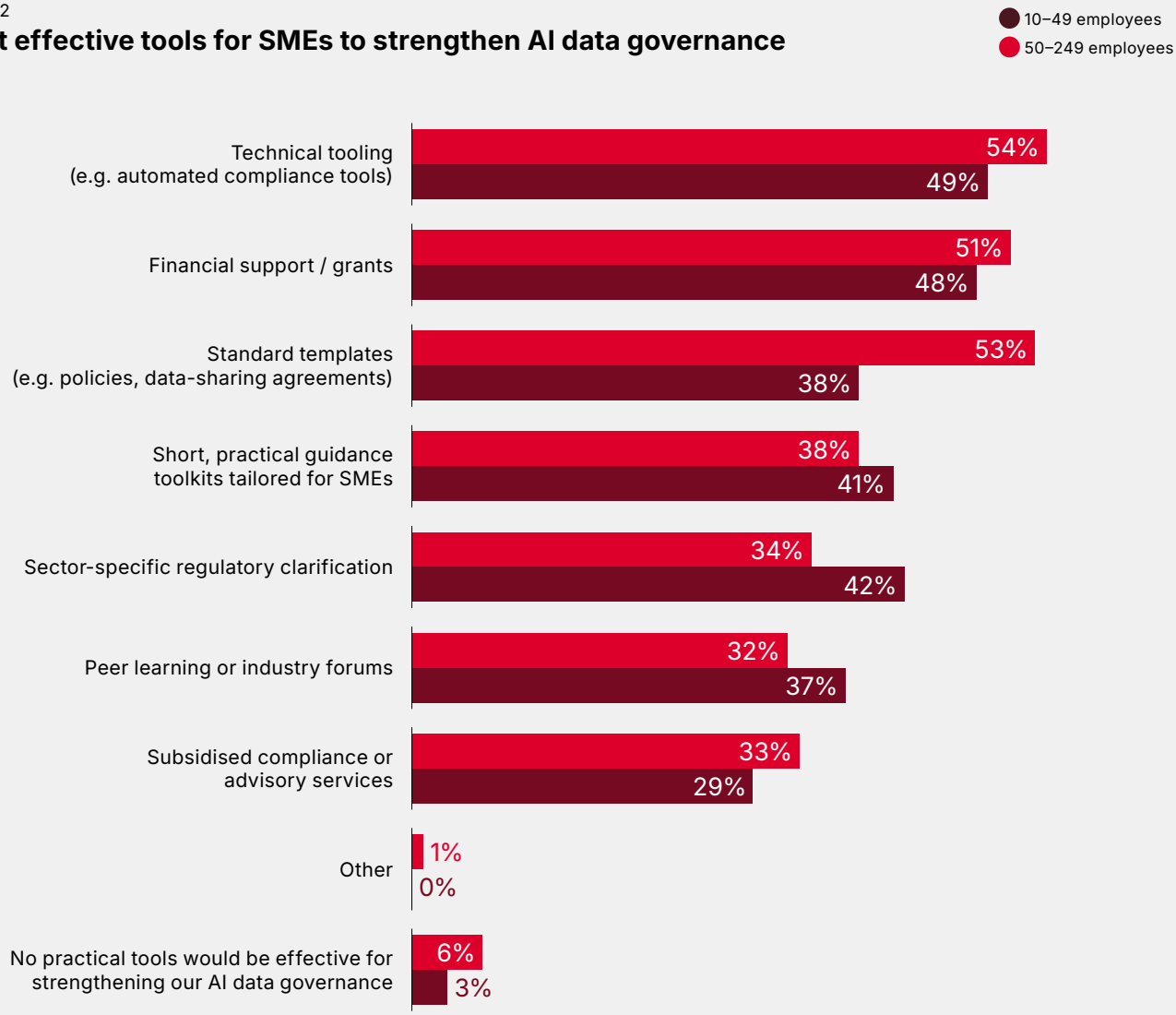
N=200
Source: GSMA AI Governance 2026

The data suggests that among AI-active SMEs that have made a financial commitment to compliance, structured risk assessment follows the initial investment.

Among the medium-sized SMEs surveyed, technical tooling and financial support were ranked as the most effective tools to strengthen organisational AI data governance, at 49% and 48%, respectively (see Figure 12). For smaller SMEs, technical tools, standard templates and financial support were cited by more than 50% of respondents. The consistency of financial support for both SMEs directly reflects the capital constraints identified throughout this report, where governance is not being deferred because organisations lack awareness or intent but rather because they lack the resources to act.

Figure 12

Most effective tools for SMEs to strengthen AI data governance



N=200
Source: GSMA AI Governance 2026

4.4 Case study insights: health and financial services

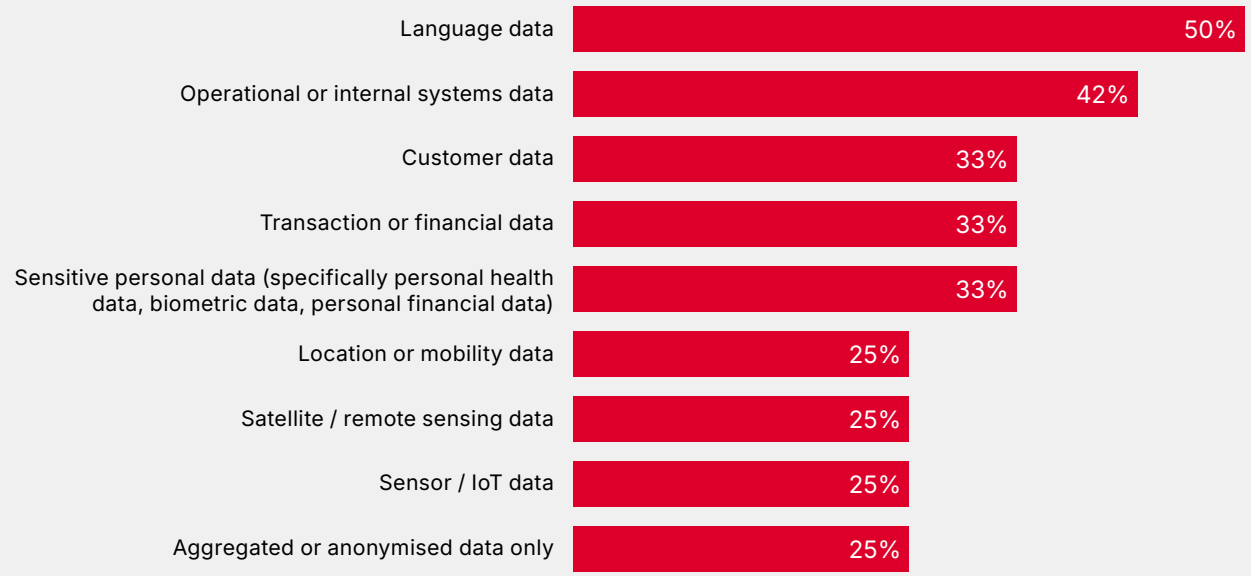
4.4.1 AI healthcare diagnostic support

South Africa faces persistent healthcare system challenges, including shortages of trained clinicians, uneven access to specialist diagnostics and significant disparities between urban and rural service provision. AI-enabled diagnostic support tools for medical imaging analysis, clinical decision support, symptom triage and self-care companions have emerged as promising mechanisms to extend scarce expertise, support frontline health workers and reach populations that formal health systems do not consistently serve.

For SMEs and startups developing or deploying these tools, data governance is a core determinant of whether an AI system can be built, deployed at scale, trusted by clinical and public health partners and sustained over time. Healthcare data is classified as special personal information under POPIA, leading to higher compliance obligations, including explicit consent and prohibition on data processing without prior authorisation. The key types of data that AI-enabled SMEs engage with include language data, operational data and customer data (Figure 13). AI-driven clinical support decisions may engage automated decision-making provisions, while accountability chains across developers, healthcare institutions, funders and patients are among the most complex of any use case examined in this study.

Figure 13

Types of data used by AI-enabled SMEs in the health sector



N=200
Source: GSMA AI Governance 2026

AI healthcare diagnostic support

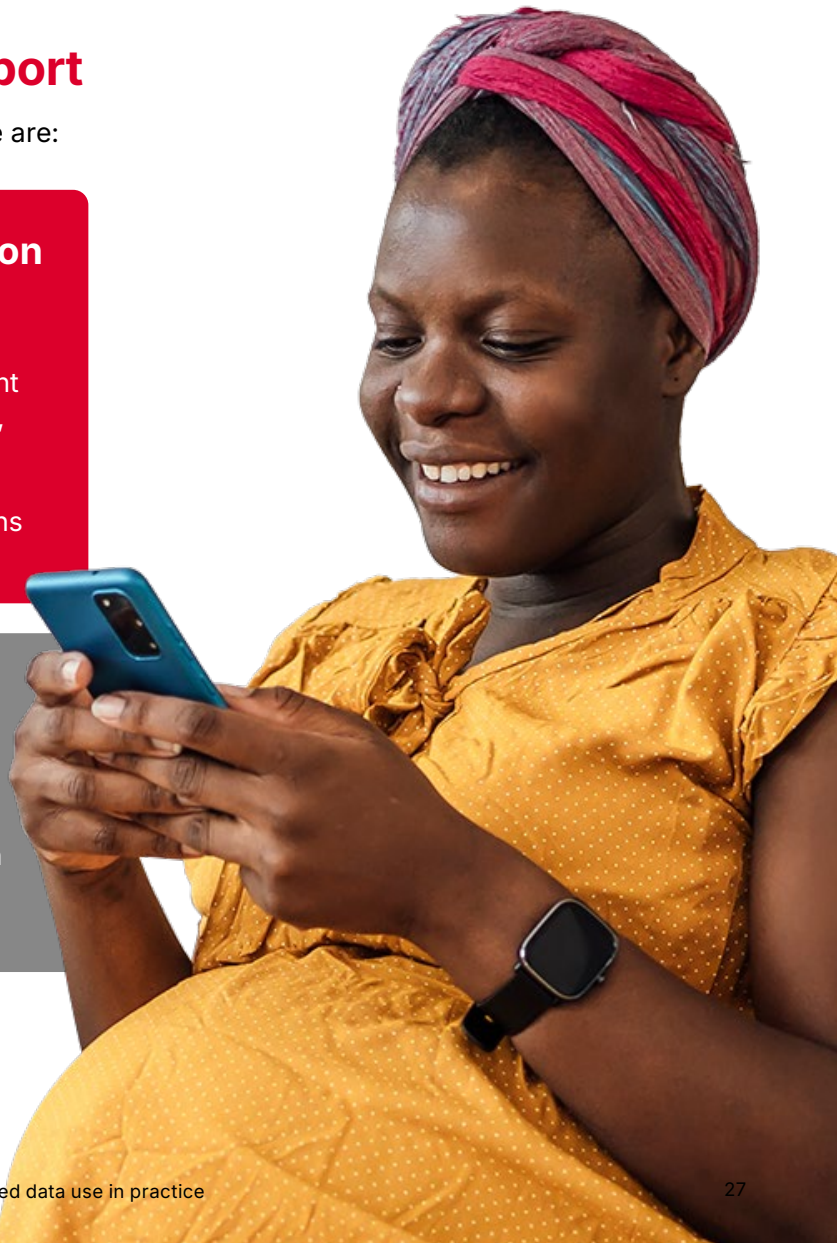
The two case studies presented under this use case are:

MomConnect Ask-A-Question (AAQ)

Operated by Reach Digital Health in partnership with the National Department of Health. MomConnect is a large-scale, established deployment that illustrates how a mature, governance-conscious organisation navigates AI data obligations in a public health context.

Audere

A nonprofit AI health technology organisation operating in South Africa, demonstrates how a growth-stage developer builds governance-by-design in a context where sector-specific standards do not yet exist.



CASE STUDY 1
MomConnect Ask-A-Question (AAQ)

A WhatsApp and SMS AI-powered natural language processing (NLP) tool to triage and respond to health queries from pregnant women and new mothers.

Key lesson

Even organisations with mature data governance can encounter structural consent gaps when AI is layered onto programmes rather than being built in to the design.

Sector:
Maternal health

AI maturity stage:
Scale

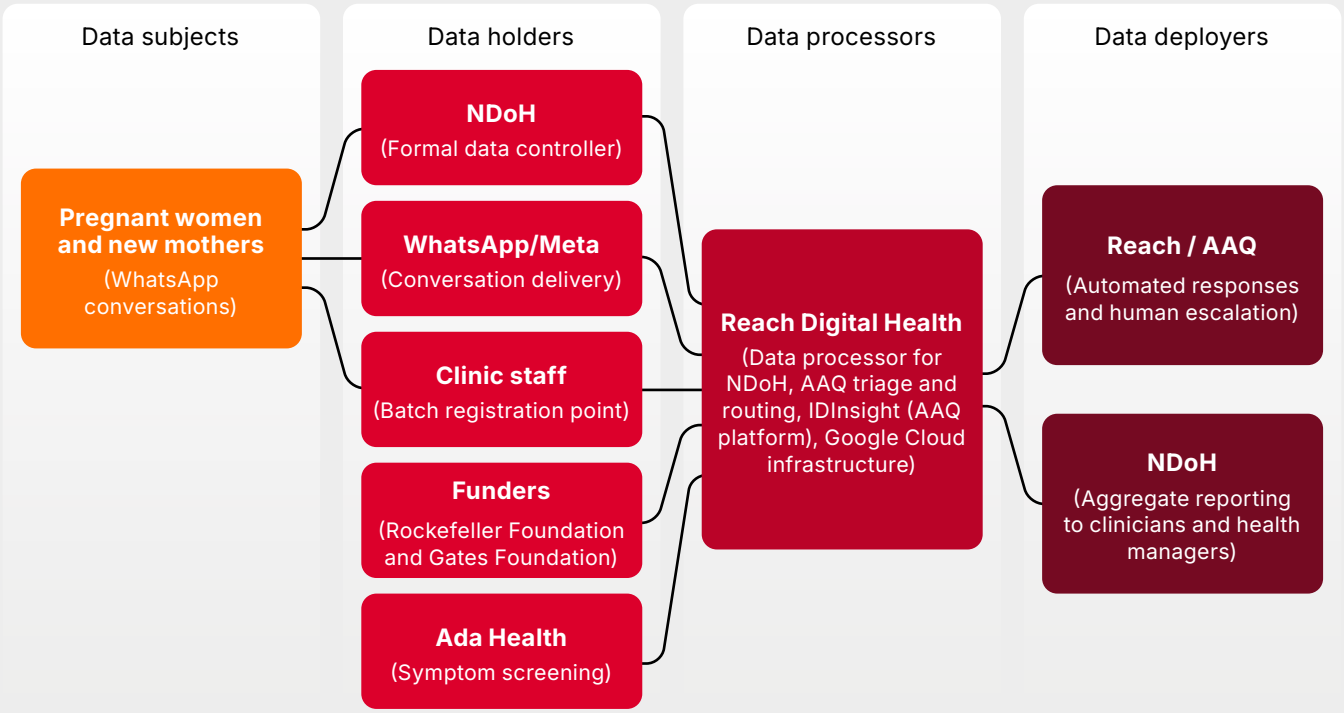
Data life cycle position:
Processor and deployer

Data sources used:
WhatsApp/SMS conversational data, NDoH databases, clinical annotations

MomConnect is a flagship maternal health programme of South Africa's National Department of Health (NDoH), launched in 2014 and operated by Reach Digital Health, a South African nonprofit technology organisation.[24] The programme provides stage-based health information via SMS and WhatsApp to pregnant women and new mothers across all 11 official languages, reaching close to 5 million registered users.[25] In 2018, Reach introduced its first AI capability, an NLP system to classify and triage helpdesk messages. This evolved into the Ask-A-Question (AAQ) feature, developed in collaboration with IDInsight's data science team while working with Reach Digital Health on the MomConnect platform, and subsequently open-sourced by IDInsight in 2024 as a stand-alone tool for the broader social sector.[26] The AAQ system is powered by Google's Gemma model and supported by Google through its AI accelerator programme.[27] By 2025, AAQ was processing up to 60,000 queries per month, achieving 83% accuracy in FAQ matching and 85% in urgency detection, and had directed approximately 100,000 urgent cases to human helpdesk operators.[28] MomConnect AAQ is now one of the most substantive national public health AI deployments in Africa.[29]

24 Reach Digital Health. (2025). MomConnect.
25 Ibid.
26 Verma, T. and You, S. (22 February 2024). "Breaking language barriers with IDInsight's generative AI tool – "Ask A Question"". IDInsight.
27 Reach Digital Health. (7 October 2025). "The power of dialogue: how AI is reshaping maternal health support in South Africa".
28 Ibid.
29 Chetty, R.L. (6 August 2025). "Google leverages its AI to assist local platform MomConnect". Hypertext.

How does the data flow?



Key considerations for MomConnect across the AI life cycle

Phase 1:
Data collection and generation

- Most users registered through clinic staff on shared devices, so individually confirmed consent under POPIA was never established.
- Valid consent was obtained only for health information delivery, not for the secondary use of data to train or improve AI models.
- Competing funder requirements on data standards from international donors place the reconciliation burden on Reach Digital Health.

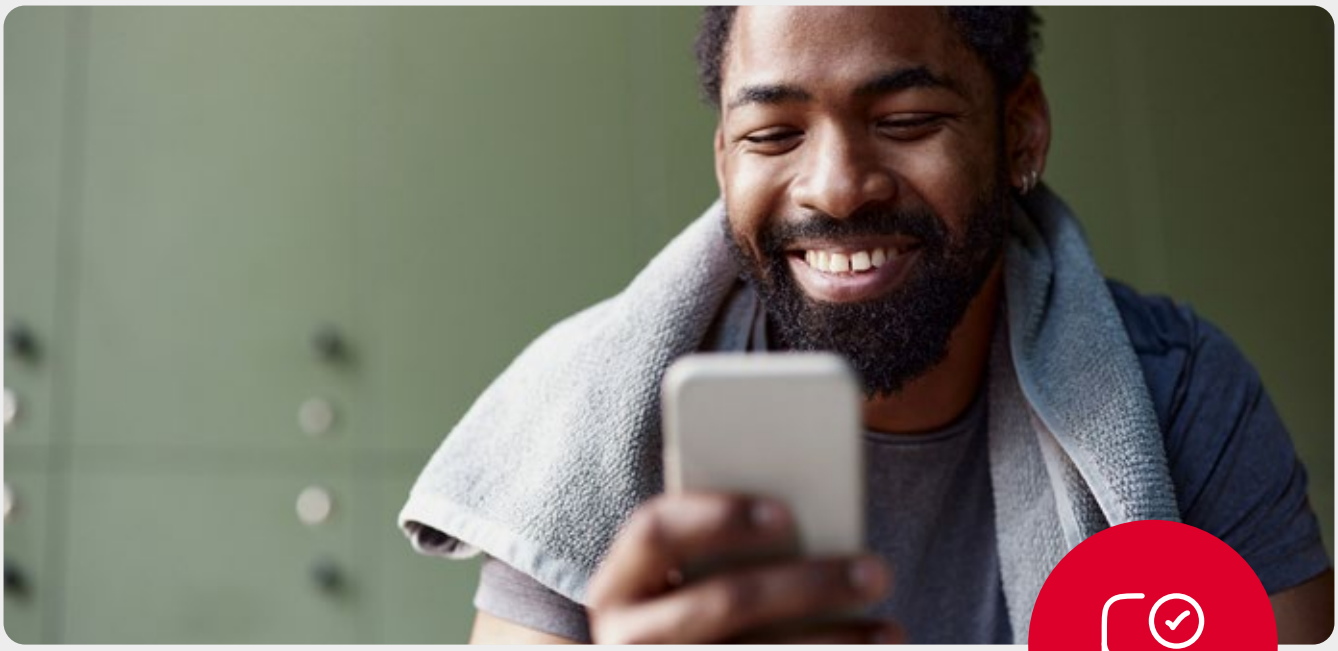
Phase 2:
Model design and evaluation

- No South African framework specifies what bias testing or representativeness standards apply to AI systems trained on maternal health conversation data.
- The Be Wise user population, who are digitally active WhatsApp users, may not represent the hardest-to-reach women the programme is designed to serve.

Phase 3:
Deployment and monitoring

- AAQ's urgency detection automatically determines whether a user receives an instant response or is escalated to a human operator, which has a direct impact on patient care.
- Responsibility for AI-assisted outcomes is shared by stakeholders with no clear accountability framework for liability.
- Government partners across African markets resist health data crossing national borders, limiting the ability to scale MomConnect's open-source tools internationally.

MomConnect represents the height of data governance in South Africa's health AI landscape. Reach Digital Health adopted POPIA compliance before it became law and has embedded data governance as a core organisational practice, adjusting its approach as new AI capabilities are introduced rather than treating compliance as a separate workstream. Still, structural governance challenges persist that are instructive for the broader sector. For the broader SME population building AI tools in partnership with public health institutions, this accountability vacuum creates significant liability exposure that requires governance frameworks.



CASE STUDY 2
Audere's Be Wise Health

A WhatsApp-based AI health companion providing sexual and reproductive health, mental health and chronic disease support to South Africans via conversational AI.

Key lesson

Building governance frameworks before funders or regulations require it to be strategic. However, internal frameworks are no substitute for national standards, which do not yet exist.

Sector:
Primary and community healthcare

AI maturity stage:
Growth

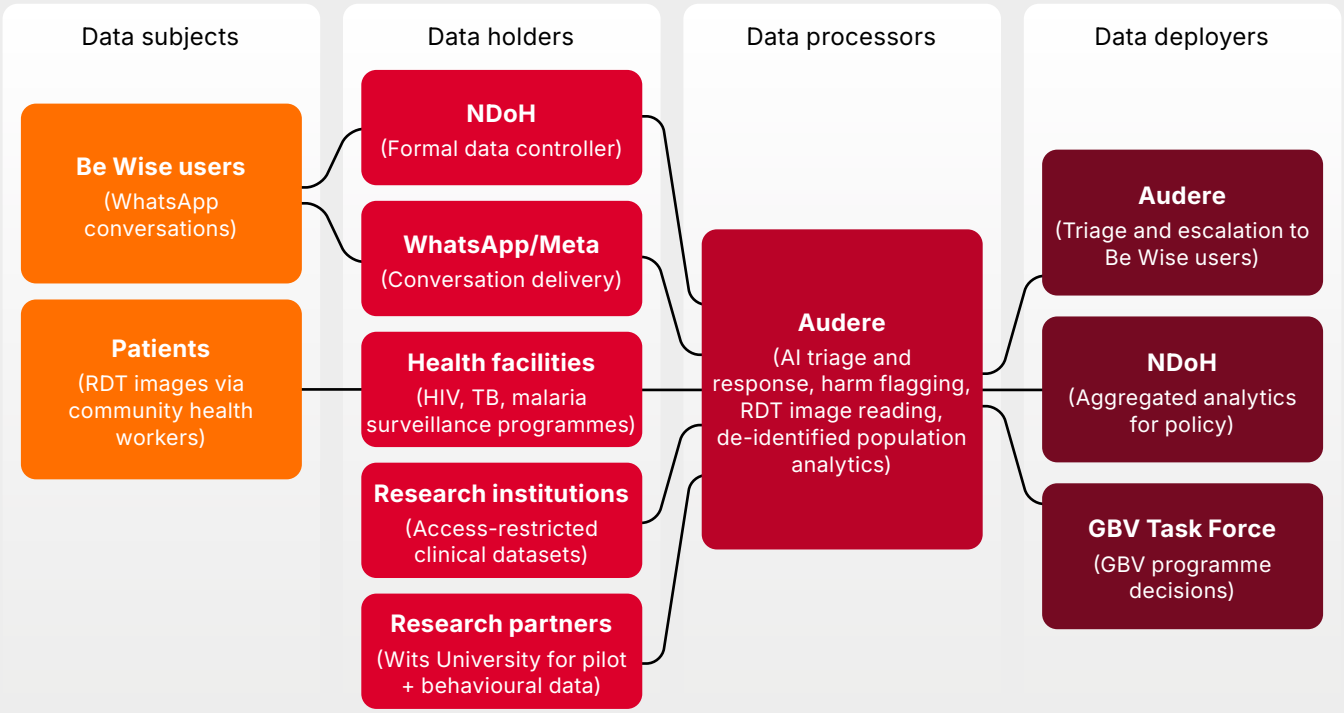
Data life cycle position:
Developer and processor

Data sources used:
WhatsApp conversation data, rapid diagnostic test (RDT) images, NDoH population-level analytics

Audere is a nonprofit AI health technology organisation co-headquartered in Seattle and Johannesburg, founded in 2018 with an initial grant from the Gates Foundation.[30] In South Africa, Audere operates primarily through Be Wise Health, a direct-to-consumer AI companion deployed via WhatsApp in partnership with the NDoH. By early 2025, Be Wise had more than 50,000 active users and had facilitated more than 1 million conversations covering sexual and reproductive health, mental health support and chronic disease management. Audere also operates a computer vision capability through an on-device software development kit that reads rapid diagnostic tests, with emerging applications in South Africa for HIV, tuberculosis (TB) and malaria surveillance.[31]

30 Business Wire. (18 April 2023). "Audere Receives $9.35M Grant for Its Flagship HealthPulse AI™ Technology Share".
31 See: Audere website.

How does the data flow?



Key considerations for Audere across the AI life cycle

Phase 1:
Data collection and generation

- Users consent via terms and conditions at their first WhatsApp interaction, with no coverage of secondary uses of conversation data for GBV risk analytics.
- The deliberate absence of a unique user identifier protects privacy but makes it virtually impossible to honour individual data access requests or provide redress if an AI interaction causes harm.

Phase 2:
Model design and evaluation

- Audere invested three years in contextual pilot testing across user groups, clinical settings and health topics, but the Be Wise user population skews to digitally active individuals and may not represent older women, rural users or those with low digital literacy.
- No South African framework defines what "sufficient explanation" means for a conversational AI health tool, leaving Audere to construct its own compliance standards from US Food and Drug Administration (FDA), Health Insurance Portability and Accountability Act (HIPAA) and GDPR guidance.

Phase 3:
Deployment and monitoring

- Audere's harm-flagging system automatically escalates mentions of suicide, gender-based violence (GBV) and mental health crisis to human operators, but has not been independently verified against any South African standard.
- Audere is incorporated in the United States with data infrastructure spanning multiple jurisdictions, while the governance of cross-border data flows under POPIA has not been fully documented.

Audere represents a governance-by-design approach that is less common among AI-active SMEs in South Africa. Harm monitoring, safety flagging and evaluation frameworks were built before funders or regulators required them. Still, the absence of sector-specific AI health standards means Audere is constructing compliance frameworks that should be provided by a recognised regulatory authority. For the broader SME population, most lack technical capacity, a mission-driven culture and funder requirements to motivate investment from Audere.

4.4.2 Alternative credit scoring

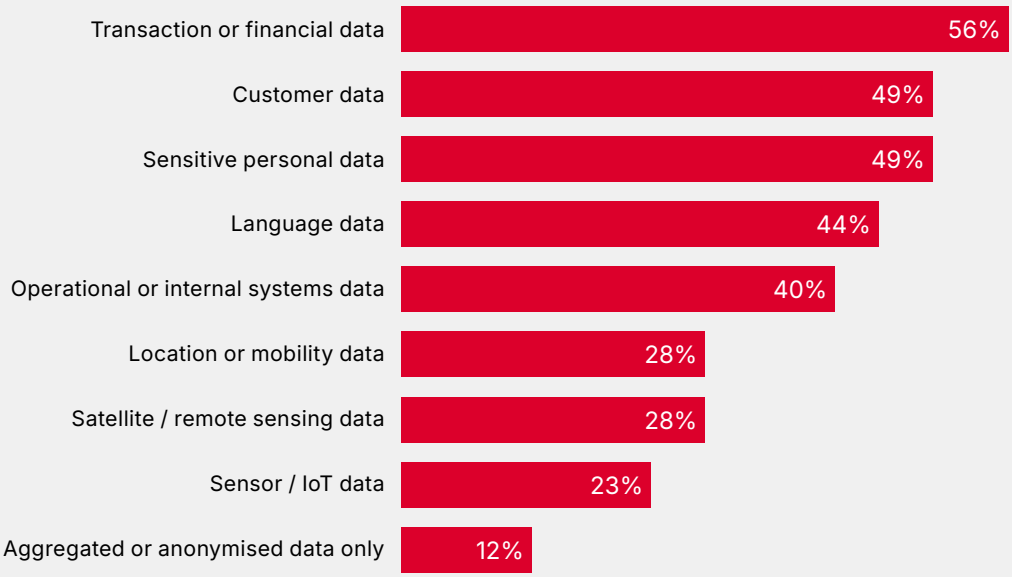
South Africa's SMEs have a significant and well-documented credit gap. The International Finance Corporation (IFC) estimates that only 5% of formal MSMEs in South Africa have access to credit, which is one of the lowest rates among middle-income countries with developed financial systems.[32] Traditional credit assessment models, which depend on formal banking histories, collateral and bureau data, exclude a substantial proportion of viable SME borrowers, particularly those in the informal sector, those in early stages of formalisation and women-owned businesses. AI-enabled alternative credit scoring addresses this gap by using untraditional data signals such as mobile transaction patterns, business account activity, payment behaviour and platform-generated behavioural data, to assess creditworthiness where conventional records are absent or insufficient. SMEs using AI in the financial services sector work with high amounts of financial transaction and customer data, as well as sensitive personal data (Figure 14).

For SMEs and startups operating in this space, AI data governance is central to whether their models produce fair outcomes, whether their data access arrangements are legally sustainable and whether the automated decisions they generate can survive regulatory scrutiny as enforcement matures. The intersection of POPIA, the National Credit Act (NCA) and the Financial Sector Conduct Authority's (FSCA) consumer protection mandate creates a layered regulatory environment that most alternative lenders navigate without AI-specific guidance.

32 IFC. (18 September 2025). "IFC and FirstRand Bank partner to widen access to finance for small businesses in South Africa". Press release.

Figure 14

Types of data used by AI-enabled SMEs in the financial services sector



N=200
Source: GSMA AI Governance 2026

Alternative credit scoring

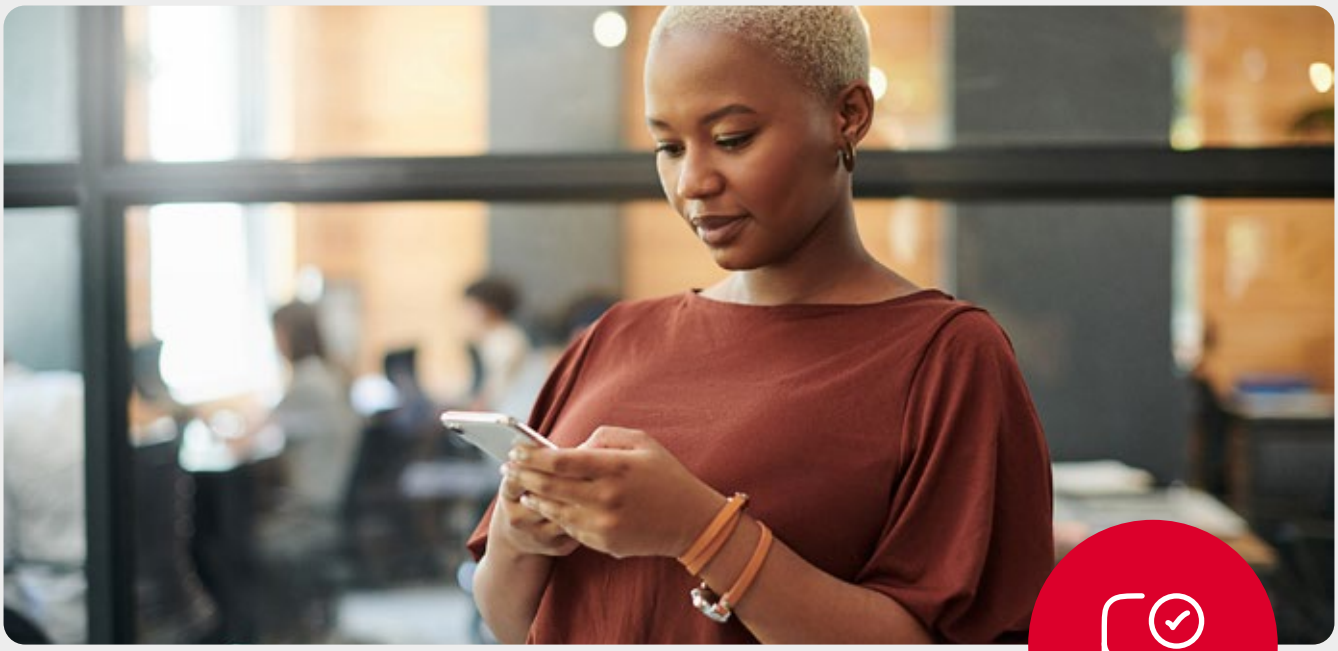
The two case studies examined are:

JUMO

A Cape Town-founded banking-as-a-service (BaaS) infrastructure provider operating AI credit scoring across Africa, illustrates the governance dynamics of a growth-stage platform operating in a multi-actor, multi-jurisdiction ecosystem.

Lula (formerly Lulalend)

A South African fintech that pioneered AI-enabled SME lending in the country, has evolved into an embedded credit and digital-only banking platform. Lula illustrates the governance journey of a domestic deployer building progressively richer data assets through its own financial infrastructure.



CASE STUDY 3
JUMO

A banking-as-a-service (BaaS) platform that uses AI to deliver credit and savings products to underserved individuals and MSMEs across Africa via mobile network operator (MNO) partnerships.

Key lesson

Compliance by design is achievable at scale, but distributing responsibility for governance across MNOs, AI engine providers and licensed lenders creates accountability gaps that have not been resolved by existing legal frameworks, and are difficult for affected customers to navigate.

Sector:
Alternative credit scoring and digital lending

AI maturity stage:
Scale

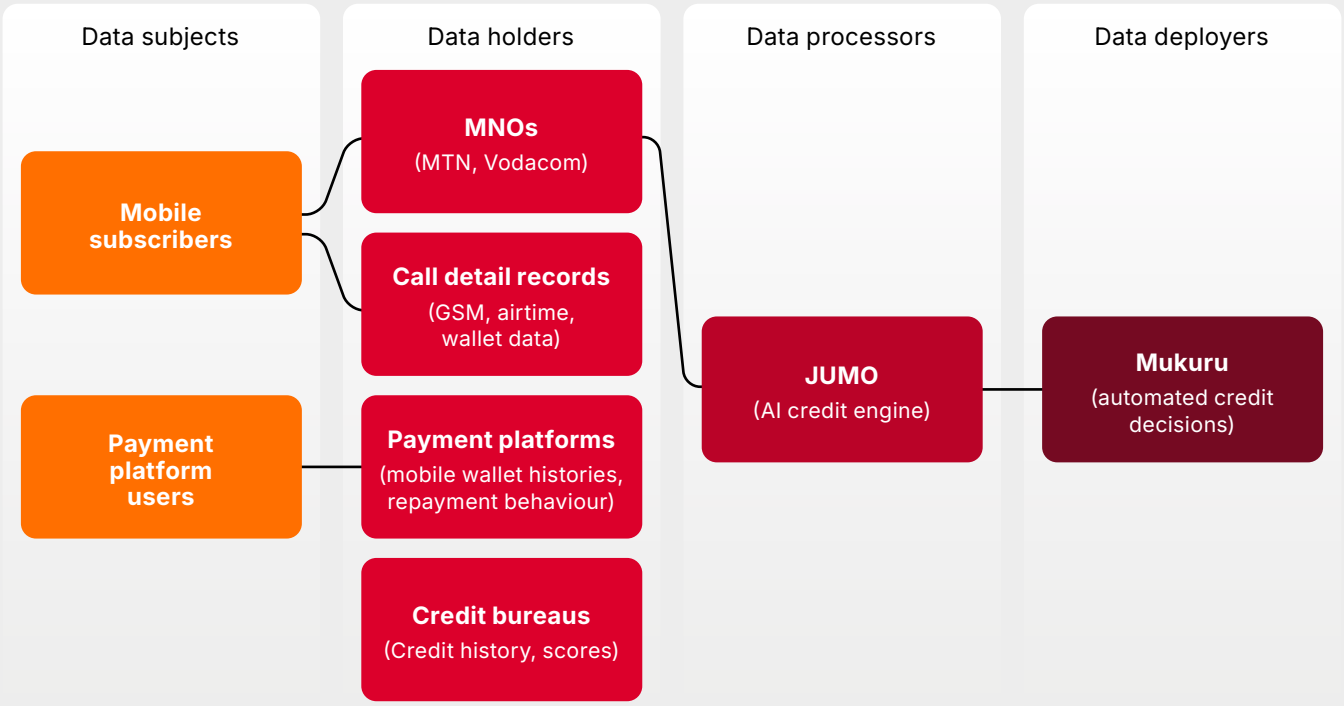
Data life cycle position:
Processor

Data sources used:
MNO call records, mobile wallet transaction histories, airtime usage patterns, repayment behaviour

JUMO is a fintech founded in Cape Town in 2015 by Andrew Watkins-Ball and now headquartered in London.[33] It operates as a BaaS infrastructure provider, developing AI-powered credit and savings products for underserved individuals and MSMEs across Africa. By 2025, JUMO had disbursed more than $7.9 billion in loans to more than 31 million customers across nine markets, maintaining default rates below 4%.[34] In South Africa, JUMO partnered with Mukuru to launch Fast Loan, targeting the estimated 16.8 million South Africans outside the formal credit system.[35] JUMO has driven down the cost of lending by 47% since 2015, and its 2026 survey found that 93% of MSMEs reported that access to credit had helped their business.[36] JUMO's ethical lending approach was independently verified in 2025, achieving a 92.2% score in the Cerise+SPTF Customer Protection Assessment.[37]

33 See: JUMO website.
34 Ibid.
35 Invest Africa. (23 October 2025). "Mukuru and JUMO partner to launch responsible AI powered credit solution in South Africa".
36 JUMO. (2025). "The impact of MSMEs". LinkedIn.
37 Invest Africa. (8 July 2025). "Fintech JUMO achieves landmark 92.2% score for customer protection standards".

How does the data flow?



Key considerations for JUMO across the AI life cycle

Phase 1:
Data collection and generation

- JUMO's credit engine relies on MNO behavioural data collected under telecommunications consent frameworks, which may not extend to credit scoring.
- The mechanisms through which consent is obtained, communicated and renewed at scale across platform-mediated ecosystems have not been publicly documented.

Phase 2:
Model design and evaluation

- JUMO's models are trained on mobile behavioural data that reflects existing patterns of digital access, which may be shaped by historical inequalities in access to technology.
- There are no verified compliance pathways for providing customers with meaningful explanations of AI-driven credit decisions under POPIA.

Phase 3:
Deployment and monitoring

- JUMO generates credit decisions with no human review of individual applications, yet there is no sector code of conduct for automated credit decisions in South Africa.
- No legal framework clearly allocates responsibility for consent, explainability or customer redress across MNOs, JUMO and licensed lending partners.
- JUMO's AI infrastructure spans multiple jurisdictions, yet POPIA has no adequacy determinations in place. This gap is growing as JUMO expands to other African markets.

JUMO is among the more governance-mature actors in South Africa's alternative lending ecosystem, with a compliance-by-design approach and independent third-party certification. Still, structural gaps in consent, explainability and cross-border accountability persist. For the broader population of SME-scale fintech lenders with far less capacity, these challenges are more acute and less managed.

CASE STUDY 4
Lula (formerly Lulalend)

Lula is a South African fintech using proprietary AI credit scoring to provide working capital, equipment financing and digital-only financial services to SMEs unable to access traditional bank lending.

Key lesson

Having direct accountability for automated credit decisions creates cleaner compliance obligations than multiparty models, but as Lula evolves from lender to embedded platform to neobank, its governance framework must expand in step with its growing data asset.

Sector:
Alternative credit scoring and digital lending

AI maturity stage:
Growth to scale

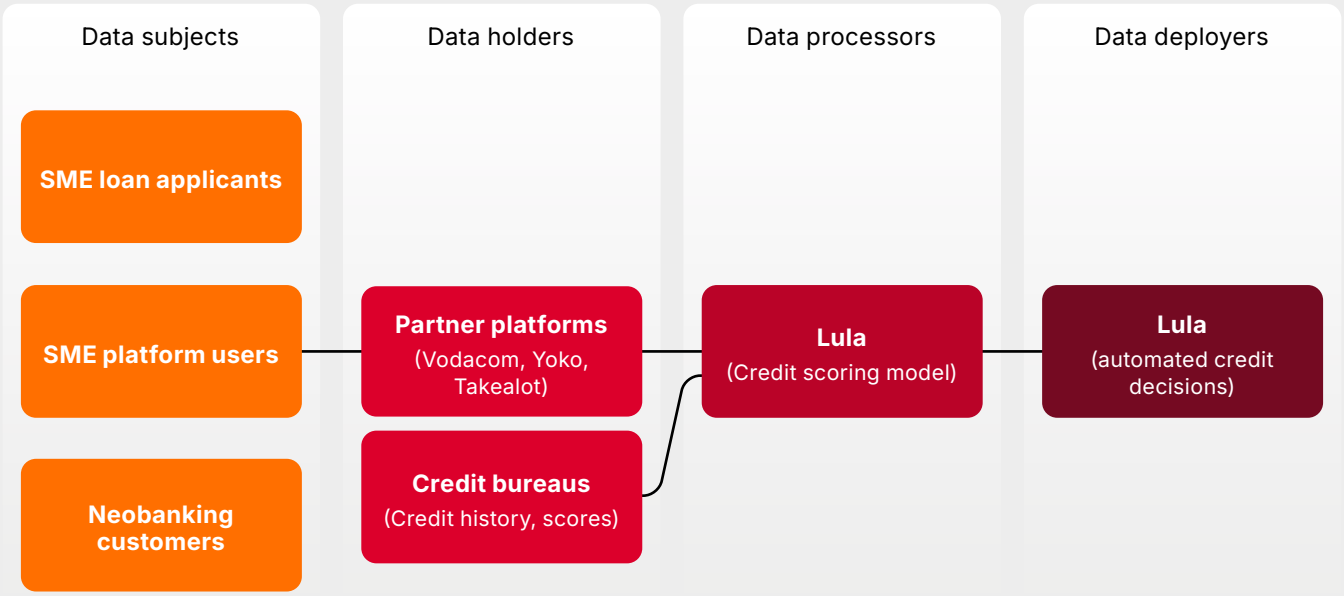
Data life cycle position:
Developer and deployer

Data sources used:
Business bank statements, identity documents (IDs), partner platform transactional data, digital-only account activity, credit bureau data

Founded in Cape Town in 2014 as Lulalend, Lula is South Africa's first, and one of its most established, AI-powered SME lenders.[38] It uses a proprietary credit-scoring algorithm to provide short-term working capital, equipment financing and revolving credit to businesses that cannot access traditional bank lending, disbursing funds in hours with loan sizes ranging from ZAR 10,000 to 5 million (up to USD 0.6 million).[39] Lula has evolved from a specialist lender into an embedded credit and digital-only banking platform, integrating with Vodacom, Yoco and Takealot, and launching a digital-only product with Access Bank that provides SMEs with business accounts, AI-led cash flow management and real-time credit access.[40] Backed by impact-oriented investors including IFC, Lightrock, DEG and Women's World Banking, with ZAR 170 million (USD 10 million) from IFC in 2025 and ZAR 340 million (USD 21 million) from FMO in 2026, Lula has a specific programme targeting women-owned SMEs.[41,42] Lula holds its own credit provider registration under the National Credit Act and has direct accountability for lending decisions, making it a particularly relevant reference case for the broader South African SME fintech population.[43]

38 Kene-Okafore, T. (1 February 2023). "South African digital lender Lulalend to launch banking product off the back of $35M Series B". Crunch.
39 See: Lula website.
40 Kene-Okafore, T. (1 February 2023). "South African digital lender Lulalend to launch banking product off the back of $35M Series B". TechCrunch.
41 Ibid.
42 IFC. (11 November 2025). "IFC partners with Lula to create jobs and boost small business growth in South Africa".
43 Lula. (2017). "About Lula".

How does the data flow?



Key considerations for Lula across the AI life cycle

Phase 1:
Data collection and generation

- Applicants using partner platforms such as Vodacom or Takealot may not have consented to their transactional data being used by Lula for credit decisions.
- Digital banking account holders generate continuous credit-relevant data signals, but it is has not been documented whether consent to open an account covers subsequent uses for model training or portfolio monitoring.

Phase 2:
Model design and evaluation

- Calibrating models for regional business payment cycles improved approval rates for informal retailers by 30% while reducing defaults, demonstrating both the risk of poorly calibrated models and the commercial case for contextual fairness.
- Lula's training data skews to digitally active, banked businesses, likely underrepresenting informal, cash-based and women-owned SMEs.

Phase 3:
Deployment and monitoring

- Credit decisions are fully automated with no human review, directly engaging POPIA's automated decision-making provisions, including the obligation to provide reasons for refusals.
- As the direct lender and registered NCA credit provider, Lula bears sole accountability for automated credit decisions, making governance failures more directly attributable and more exposed to regulatory action than in multiparty models.
- Lula's growing integration with partner ecosystems and cloud infrastructure means the extent of any cross-border data processing warrants ongoing scrutiny under POPIA.

Lula's domestic single-entity model, under which it owns the customer relationship, the credit decision and an expanding digital-only banking data asset, makes its governance obligations cleaner and more directly enforceable than multiparty BaaS models. Yet the same evolution that strengthens Lula's data advantage also expands its governance obligations. As richer, more continuous data is generated through its digital-only banking product, the governance framework indicating what that data can be used for must keep pace with the product strategy. This is a challenge the broader SME fintech population faces with far less institutional and investor support.

Cross-cutting insights from the case studies

MomConnect, Audere, JUMO and Lula differ substantially in sector, scale and data governance maturity. Yet across all four case studies, certain AI data governance dynamics are consistent.

Stakeholder interviews have confirmed that these dynamics are broadly characteristic of the wider AI-active SME population in South Africa.

Consent frameworks were not designed for AI.

Across all four cases, consent was obtained for a purpose that predated the AI application built on that data. MomConnect users consented to health information delivery, not model training. JUMO draws on MNO data collected under telecommunications frameworks. Lula's digital-only banking customers generate ongoing credit signals under account-opening terms that do not address this particular use. Audere's consent mechanism does not explicitly cover the secondary uses of conversation data for population-level analytics. In every case, the consent gap is structural and a result of AI being layered onto existing data relationships. This cannot be resolved through the efforts of individual organisations alone.

Bias documentation is absent from all four cases at the model level.

None of the organisations has publicly documented systematic bias testing for its South African-specific AI deployment. In each case, the populations served the least by the AI system are also least represented in training data, which is a governance failure that current frameworks do not require any actor to address.

Interpretation of POPIA provisions is unclear.

Every case study involves automated decisions with material consequences for individuals, such as health triage, crisis escalation, credit scoring or credit refusal. POPIA's requirement for sufficient information about the underlying logic of automated decisions has no published interpretation for any of these contexts.

Accountability for the AI model's performance and accuracy is distributed but not allocated.

In healthcare, responsibility sits jointly with the NDoH, implementing organisations, technology partners and third-party integrations with no legal framework clearly assigning liability when an AI-assisted interaction causes harm. In financial services, it is distributed across MNOs, AI engine providers and licensed lenders through private contractual arrangements that appear opaque to affected customers. In both sectors, the absence of a clear accountability framework means that if something goes wrong, the affected individual has no obvious route to redress, since no actor in the value chain has a clear obligation to provide it.

Post-deployment monitoring is only present when organisations have built it themselves.

Of the four organisations, only Audere invested in real-time harm monitoring before any funder or regulator required it. No equivalent infrastructure exists in the financial services cases. Model drift, real-world performance divergence and growing bias over time receive no attention from existing regulatory frameworks, investor due diligence processes or ecosystem intermediaries.

4.4.3 Lessons from international data standards

The AI data governance challenges identified in this report are not unique to South Africa. Across jurisdictions, SMEs and startups have struggled to navigate governance frameworks designed for larger enterprises while operating in increasingly complex, platform-mediated AI ecosystems. Table 3 shows the range of mechanisms applied by other countries to address data governance constraints around compliance and data access. The four regions in the table were selected because they respond directly to constraints identified in the South African use cases.

Table 3: Overview of relevant global data governance and AI policies

Region	Data governance policy	Relevance to AI use	Lessons for South Africa
 **European Union**	EU Artificial Intelligence Act[44]	• Explicitly names SMEs as a distinct category, with free regulatory sandboxes, simplified technical documentation and proportionate conformity fees. • Risk-tiered obligations are directly relevant to high-risk AI use cases, such as credit scoring and clinical decision support.	• Illustrates what systemic, sector-specific AI governance looks like when proportionality is embedded. • The burden of high-risk AI on SMEs must be addressed through explicit architectural choices.
 **India**	Digital Personal Data Protection Act[45]	• Phased compliance timelines (up to 18 months) reduce barriers for startups deploying data-intensive AI models. • Startup exemptions from Data Protection Office appointments and mandatory audits lower the fixed cost of AI-driven experimentation.	• Proportionality can be legislated explicitly, rather than left to regulatory discretion. • Offers a model for differentiating obligations by organisation size and maturity, which is currently absent from POPIA.
 **Kenya**	Data Protection Act,[46] Draft Guidance Notes[47]	• Fines capped at 1% of annual turnover link the severity of penalties to the scale of a business, making regulatory risk clearer for AI-deploying SMEs. • Sector-specific guidance notes (e.g. digital finance, telecoms) translate abstract obligations into context-aware expectations.	• Turnover-based penalties are more directly adaptable to South Africa's context than EU models, given comparable development conditions. • Demonstrates that proportionality can be achieved through sanctions architecture and interpretive infrastructure without rewriting baseline obligations.
 **United Kingdom**	Information Commissioner Office advice,[48] Digital Regulation Cooperation Forum (DRCF)[49]	• The ICO's plain-language compliance tools help SMEs engage with AI innovation without overengineering governance processes. • DRCF AI and Digital Hub coordinates across four regulators, reducing conflicting guidance where AI systems trigger overlapping regulatory domains.	• Proportionality can be delivered institutionally rather than legislatively, which is relevant to South Africa's fragmented regulatory architecture. • A single coordinating interface could address interpretive gaps that currently fall disproportionately on smaller enterprises.

44 EU. (2024). EU Artificial Intelligence Act.
45 India Ministry of Electronics and Information Technology. (2023). The Digital Personal Data Protection Act 2023.
46 Republic of Kenya. (2019). The Data Protection Act 2019.
47 Office of the Data Protection Commissioner. (2026). "Draft Guidance Notes".
48 ICO. (2026). "Advice for small and medium organisations".
49 Digital Regulation Cooperation Forum. (2026). "AI and Digital Hub".

Across the four case studies, several insights are directly relevant to the South African context.

Robust data regulation alone is insufficient to enable SME participation in AI-driven innovation

In each example, compliance burdens on SMEs were reduced by investing in institutional coordination, clarifying how the law can be interpreted in the case of SMEs, and establishing mechanisms that translate abstract obligations into practical pathways. South Africa's challenge is not that POPIA is too weak, but that the infrastructure for operationalising it in AI contexts does not yet exist.

Proportionality must be designed, not assumed

Where proportionality has been effective, it has been embedded explicitly through tiered obligations, phased compliance timelines, tailored guidance or enforcement architecture calibrated to the scale of an organisation. Tiered obligations would mean that data governance provisions would vary based on the size, scale and risk profile of an organisation, so smaller SMEs would still be able to meet compliance requirements. South Africa's uniform application of POPIA obligations, regardless of organisation size or risk profile, combined with the absence of SME-specific compliance pathways, leaves proportionality a stated intent rather than an operational reality.

Ecosystem-level mechanisms matter as much as an organisation's compliance tools

Regulatory sandboxes, trusted data environments, and coordinated regulatory interfaces reduce the governance burden on individual organisations by addressing constraints that lie outside any single organisation's control. This is particularly relevant in South Africa, where SMEs depend heavily on platforms, cloud providers and dominant data holders, and cannot resolve the most binding governance constraints through organisation-level compliance investment alone.

Mechanisms to support SMEs are not without limitations

Sandboxes are resource-intensive and difficult to scale. Tiered obligations require clear criteria that are difficult to define and subject to gaming. Sector guidance requires sustained regulatory capacity that South Africa's Information Regulator has acknowledged it does not yet have. These constraints do not reduce the relevance of the mechanisms, but they shape how ambitiously any specific intervention can be implemented.

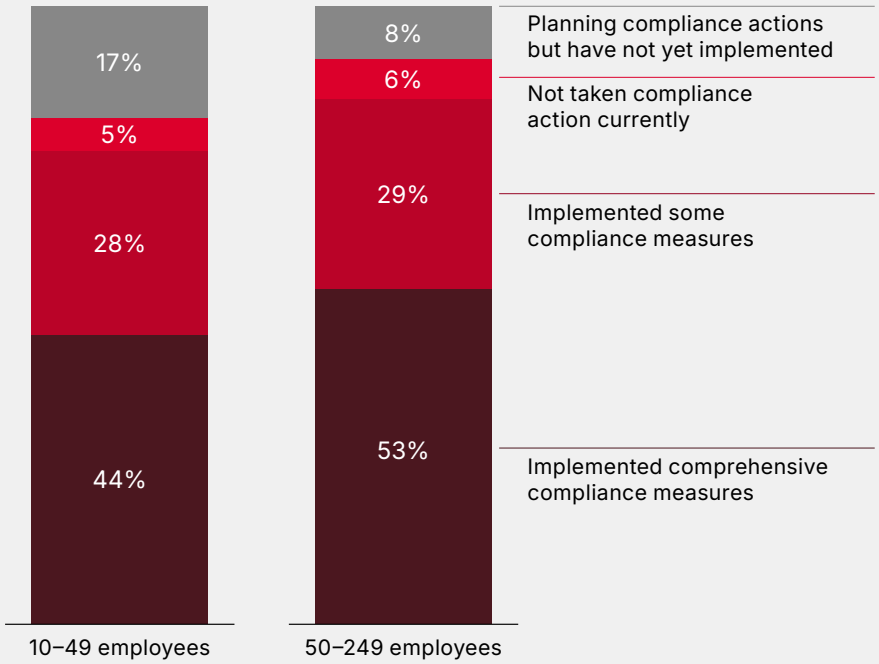
5. KEY CONSIDERATIONS FOR AI DATA GOVERNANCE

5.1 Drivers of data regulatory compliance

Despite South Africa's robust data protection policy, compliance by smaller players like SMEs and startups is uneven and not comprehensive. This is partly because POPIA's existing provisions are applied uniformly regardless of an organisation's size or maturity, resulting in SMEs having to meet compliance expectations such as documented consent management, data management, security safeguards and automated decision-making accountability, for which they do not have the capacity or funds to meet. According to our survey data, only 53% of medium-sized organisations have implemented comprehensive POPIA compliance measures, with smaller organisations exhibiting less comprehensive compliance measures at 44% (Figure 15).

Figure 15

Level of organisational compliance with POPIA based on size



N=200
Source: GSMA AI Governance 2026

The regulatory side also faces an important capacity constraint that affects SMEs and startups. Under POPIA, every public and private body is required to appoint an Information Officer, which automatically falls to the head of the organisation for private bodies unless formally delegated.[50] Despite this, the Information Regulator estimates that approximately 3 million Information Officers have not registered as required, which can lead to compliance assessment bottlenecks for SMEs in the future.[51] No government-supported SME compliance toolkit, subsidised advisory service or AI-specific guidance has been identified, which coupled with a lack of practical guidance on how provisions should be interpreted and low enforcement of provisions, makes it a struggle for SMEs to prioritise data governance compliance.

These findings highlight that although operationalising data governance policies to improve compliance is important, it must complement efforts to influence behaviour change among AI-enabled SMEs and startups. Effective drivers are more likely to come from market mechanisms than regulatory guidance alone. Including data governance standards embedded in investor due diligence, funder requirements and private sector procurement tailored to the stage of development of SMEs and startups, can provide an effective behavioural push towards data governance compliance.

50 Information Regulator South Africa. (2021). Guidance Note on Information Officers and Deputy Information Officers.
51 Michalsons. (5 March 2025). "Information regulator annual performance plan for 2025 to 2026".

5.2 Capital constraints

Stakeholder interviews conducted for this study consistently identified undercapitalisation as a primary barrier to investment in data governance among South African SMEs, which was more commonly cited than skills gaps, regulatory uncertainty or lack of awareness. SMEs and startups tend to redirect resources away from governance functions when operational upkeep is the priority. Investment in governance processes is typically deferred until a market trigger makes it unavoidable, such as an enterprise client requiring it as a condition of contracting or an investor including governance questions in due diligence.

Another persistent trend identified through stakeholder interviews is the tendency of organisations to represent themselves as AI-enabled in funding applications. This is in response to grant makers, development finance institutions (DFIs) and impact investors who signal a preference for AI-inclusive proposals. Where AI adoption is driven by funding availability, rather than whether AI is the right fit for the problem they want to solve, processes like assessing data readiness, risk or model design before deployment are adversely affected. This dynamic also creates reputational or regulatory exposure for organisations that exaggerate their AI capabilities.

Current due diligence practices in South Africa are limited to assessing team quality, commercial viability and market fit. Due diligence frameworks rarely probe AI-specific governance maturity in a way that creates formalised requirements, meaning that for most SMEs the trigger never comes. Most do not include AI-specific governance considerations, such as whether organisations are documenting consent for training data, undertaking model bias assessments, conducting post-market monitoring or validating AI models against real-world performance. Investors are largely assuming that AI models perform as intended, and few have AI-specific conditions embedded in their funding considerations.

Recommendations that rely on SMEs absorbing governance costs are likely to fail without financial support mechanisms such as grants, subsidised advisory services or development partner-funded capacity building. Existing SME support or development programmes can also be empowered to provide AI-specific compliance support to SMEs and startups.

5.3 The role of ecosystem actors

Findings from stakeholder interviews highlight that SMEs are currently not receiving specific AI or data governance support from ecosystem actors like development services providers, technology hubs, incubators and industry associations. SME support programmes can play an active role in efforts to comply with AI data governance regulations by translating governance frameworks into practical toolkits. However, organisations supporting SME growth currently lack the knowledge, funding or mandate to support AI adoption. There is a need to build specialised capacity in intermediary or SME support organisations for data protection and AI governance practices, enabling them to reach SMEs at scale.

Data access is a key barrier to AI adoption highlighted by SMEs and startups. SMEs operate in AI ecosystems where large platforms, cloud providers and data holders set access and governance terms unilaterally, leaving smaller enterprises to absorb compliance costs they did not design and cannot negotiate. Data is largely controlled by a small number of institutions including financial services providers, MNOs, retail platforms and public health systems, and formal data-sharing mechanisms are not always available to smaller players. Data-sharing agreements rely on strong relationships and networks that may be inaccessible to SMEs and some startups, creating a significant barrier to AI development.

In many cases, the data generated by South African communities and institutions is held or controlled by multinational platforms and foreign organisations operating under no domestic obligation to provide access. This adds a data sovereignty dimension to data access constraints, which is unlikely to be resolved through interoperability standards or data-sharing frameworks without accompanying policy intervention. 70% of surveyed SMEs indicated that they rely either entirely or partly on data from external sources (see Figure 16).

The terms of data access mechanisms and data-sharing partnerships should be re-examined to ensure that data holders have obligations to make locally generated data accessible for locally relevant AI development.

Figure 16

Data sources used by AI-enabled SMEs

N=200
Source: GSMA AI Governance 2026

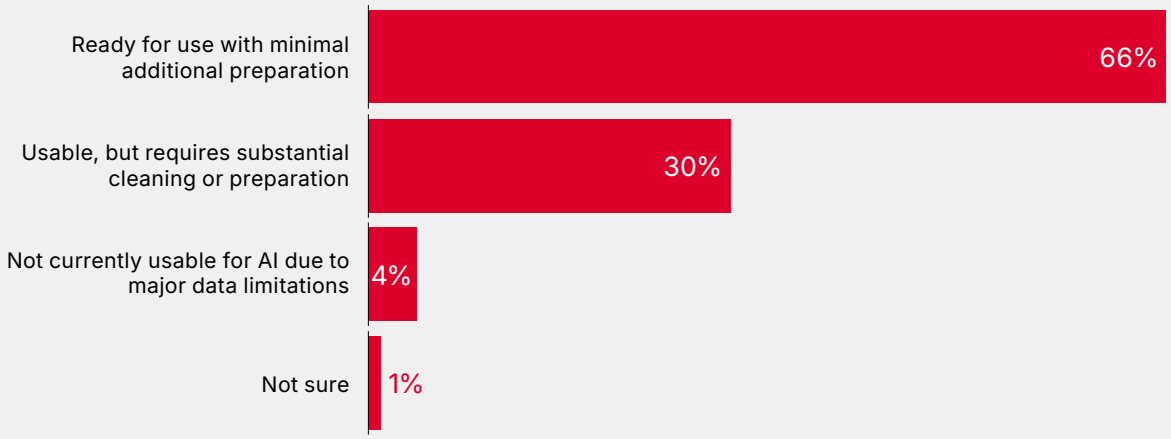
5.4 Data quality and local language data

Even where data is accessible to SMEs and startups, it can fail to meet the quality requirements for effective AI applications, as shown in Figure 17, where 34% of SMEs surveyed report that their data is not yet ready for use with minimal preparation. Inconsistent schemes, legacy systems without interoperability, incomplete public datasets and an absence of standardised identifiers are some of the key barriers to data quality. Another major constraint is the inaccessibility of local language datasets.

High-quality datasets for South African languages are either nonexistent, held in academic or research institutions at commercially prohibitive access prices, or available only through enterprise language technology providers whose pricing excludes SME-scale developers. For a country where more than 90% of the population speaks one or more African languages as a first language, this is a significant constraint, and a structural barrier to AI development that can genuinely serve its intended populations.[52]

Figure 17

Quality of data currently in use by SMEs



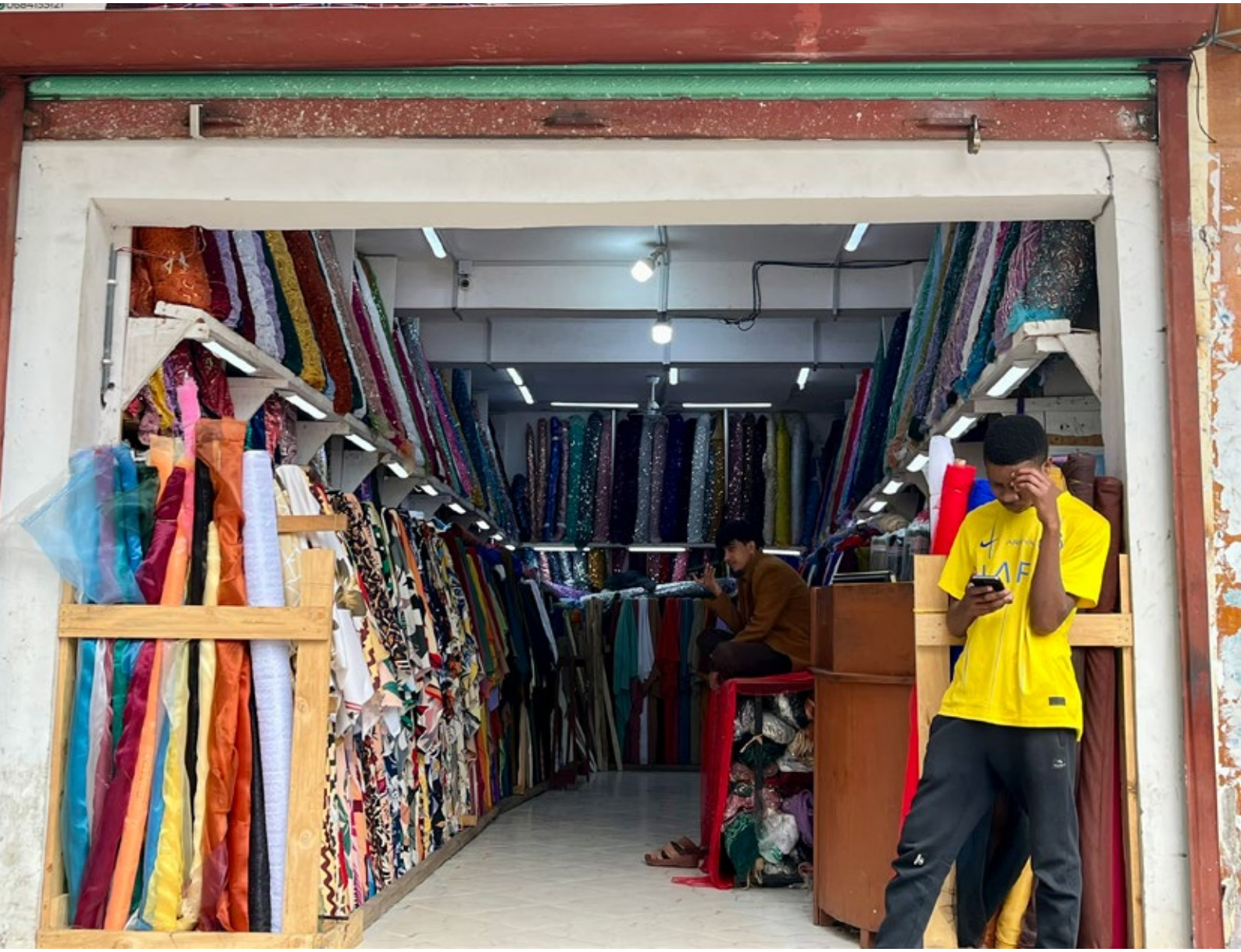
N=200
Source: GSMA AI Governance 2026

52 South Africa Gateway. (n.d.). The languages of South Africa.

AI systems deployed for South African populations that cannot adequately serve non-English speakers can introduce bias to the outcomes they produce. For a substantial portion of this population, including those with limited literacy, those living in rural areas and for whom English is a second or third language, voice interfaces and local language AI capabilities are a key enabler of digital access, especially in sectors such as health, finance and social delivery where the potential for social impact is greatest.

While a corpus of local language data is being built, local language AI tools are still not accessible or affordable for AI developers in SMEs and startups. Most available datasets require significant compute investment and specialist capability to train usable models. Commercial language tools are typically intended for enterprise clients and outside the price range of SMEs and startups, deterring them from building local language AI systems that are equipped to serve populations at the last mile. Most SMEs and startups opt instead for off-the-shelf multilingual models whose performance with South African languages has not been independently validated and is unlikely to reflect the diversity of the population being served.

Addressing the data quality gap requires dedicated efforts at the policy and ecosystem level, in the form of investments in open, publicly accessible corpora for South African languages, obligations by publicly funded research institutions to make linguistic data available for AI development and consideration of whether existing enterprise development mechanisms can be extended to incentivise private-sector data holders to make language datasets accessible to SME developers.



6. RECOMMENDATIONS

The following recommendations address the key AI governance and data challenges identified in this report, and require coordinated action from regulators, investors, development partners and local ecosystem actors in the AI value chain. Recommendations are categorised as short, medium and long term based on the ease of implementation and whether they are institutional or ecosystem driven.

RECOMMENDATION 1:
Operationalise existing policies to improve compliance by SMEs

Establish sector-specific regulatory sandboxes with priority SME access

Short term

South Africa currently has no AI-specific regulatory sandbox. Such sandboxes could build on sector-specific AI frameworks that are in the works. They can be designed to provide dedicated access to SMEs, ensuring they do not get sidelined by larger ecosystem players. Priority entry could be provided for SMEs using AI to create socio-economic impact, especially in the priority sectors of health and financial services described in this report, but also in sectors that are most likely to benefit from a structured environment to assess compliance approaches before full AI deployment.

Develop practical, sector-specific governance toolkits

Medium term

Toolkits structured by sector and use case can help SMEs and startups comply more fully with data governance requirements. A key challenge identified in this report is that SMEs and startups approach governance as a compliance checklist rather than a core function of their business operations. This is due to a lack of guidance on how data governance requirements translate into their day-to-day operations. A useful comparative guide is the UK Information Commissioner's Office (ICO) Data Essentials Programme, which provides a training programme for SMEs through plain-language, practical tools for data protection compliance. This recommendation is supported by the SME survey, which found that technical toolkits, standard templates and practical guideline were the most frequently requested types of support by AI-enabled SMEs.

Relevant stakeholders: Government departments, institutional regulators, enterprise development organisations, SME associations

RECOMMENDATION 2:
Issue AI-specific regulatory guidance under POPIA

Issue AI-specific guidance on how POPIA applies to automated decision-making

Short term

POPIA's provision of "sufficient information about the underlying logic" should be clarified to state what it means in practice for sectors such as health and financial services, where the presence of highly sensitive data presents significant risk for AI-driven automated decision-making. This would also pave the way for other sector-specific policy frameworks dealing with the handling and processing of such data.

Consider post-deployment monitoring standards a regulatory priority

Medium term

The AI data governance conversation in South Africa, as in most comparable jurisdictions, concentrates almost entirely on pre-deployment obligations. Post-deployment monitoring, which includes the detection of model drift, real-world performance reassessment and bias monitoring over time falls outside every existing obligation and due diligence framework examined in this study. Post-deployment monitoring standards should be included as a priority in any AI-specific data regulation, with a focus on thoughtful design to ensure regulatory requirements are not technically demanding or financially unattainable for smaller businesses.

Relevant stakeholders: Government departments, institutional regulators, technology hubs, academic and research institutions

RECOMMENDATION 3:
Establish formal data access and sharing mechanisms

Establish sector-specific data-sharing mechanisms

Medium term

There is a need for formal data-sharing mechanisms at the sectoral level. This would include appointing sector-specific data intermediaries and putting in place conditions for data access, to reduce SMEs' dependence on informal bilateral brokerage and personal relationships. Voluntary mechanisms to share data can be coupled with mandated access or interoperability depending on the nature of the data, enabling SMEs to address one of their most challenging constraints to data use.

Relevant stakeholders: Industry associations, big tech companies, institutional regulators, sector-specific government bodies

RECOMMENDATION 4:
Invest in the governance capacity of ecosystem intermediaries

Provide AI governance capacity to intermediary organisations

Short term

Organisations that support SMEs to grow and scale do not always have the necessary capabilities to support SMEs in adopting AI. Dedicated investment by international development partners and SME industry associations can help intermediary organisations build the knowledge and capacity that does not currently exist in the ecosystem. Technology hubs and incubators that support AI-active SMEs with products, funding and talent can also be equipped and incentivised to offer AI governance support as a core service.

Relevant stakeholders: International development organisations, technology hubs, startup incubators and accelerators, SME associations, industry associations

RECOMMENDATION 5:
Activate market mechanisms to support data governance compliance

Update due diligence processes to include AI-specific governance

Medium term

Current due diligence practice in the South African impact and venture investment community assesses team quality, commercial viability and market fit, but does not systematically probe AI-specific governance maturity, such as whether training data consent, bias assessments and post-deployment monitoring have been established, nor whether the model's real-world performance has been validated against its lab results. Investor diligence processes should be updated to include AI governance maturity as a criterion, with questions on data access and processing, model accountability and post-deployment monitoring. Making these changes at the due diligence level would help trigger behavioural change among SMEs and startups that have reached sufficient scale but are not yet complying with data protection requirements. International development partners can play a key role in convening the investor community to agree on what constitutes AI-specific governance maturity.

Review public and large private-sector procurement requirements

Long term

Partnerships with large enterprise clients, including public sector bodies and major corporates, are one of the triggers for larger SMEs to comply with AI governance standards. South Africa's public procurement framework should be updated to require evidence of AI-specific governance practices, including bias assessment, explainability documentation and post-deployment monitoring for AI-enabled services procured above a defined contract threshold. This can be done by updating the application of existing procurement standards to AI-enabled services.

Relevant stakeholders: Institutional regulators, development finance institutions, capital providers, industry associations, startup accelerators

ANNEXES

Annex 1:
Stakeholders interviewed for the research study

Key informants and roundtable participants in South Africa

Africa CDC
AI Collective
AUDA NEPAD
Audere
Baobab Network
Department of Communications and Digital Technologies (DCDT)
Envisionit Deep AI
Global Center on AI Governance
Google
HSRC
JUMO
Lula
Musa Ventures
Onafriq
OpenUp
Primus Tech Hub
Quantium Health South Africa
Reach Digital
Research ICT Africa
Simple Capital
SM Digital
SMART Africa
UVU Africa
Viamo
Vula Medical

Annex 2:
Selection of priority sectors

AI maturity among SMEs refers to the extent to which a small or medium-sized enterprise can strategically adopt, govern and operationalise AI systems in a way that is sustainable, compliant and value-generating, given its organisational capabilities, data assets and ecosystem dependencies.

For SMEs, AI maturity is defined as their ability to:

- Access and govern data across the AI life cycle
- Integrate AI in core business processes and decision-making
- Manage legal, ethical and operational risks associated with AI use
- Adapt organisational structures, skills and partnerships to support responsible AI at scale

AI maturity, therefore, is not just the deployment of technology but rather a combination of technical, organisational and governance capabilities.

Dimensions of AI maturity

Drawing from the referenced frameworks, AI maturity among SMEs can be assessed across five interrelated dimensions:

- **Strategic integration** – whether AI is peripheral or embedded in core business functions
- **Data readiness and governance** – the quality, accessibility and governance of data used for AI
- **Organisational capacity** – skills, processes and accountability structures supporting AI use
- **Risk and compliance management** – ability to interpret and operationalise data protection and AI-related requirements
- **Ecosystem integration** – how SMEs manage dependencies on platforms, cloud providers and data holders

Selection of priority sectors

Priority sectors were selected using a 2×2 matrix plotting AI maturity against data governance complexity, with impact potential used as an additional lens to assess the development and economic value of each quadrant.

Table 4: Dimensions and scoring criteria for priority sector selection

	 **Impact potential** The expected development and economic value of scaling responsible AI adoption in a sector, considering social outcomes, SME productivity and ecosystem-wide effects.	 **AI maturity** SMEs in a sector are already developing or deploying AI in core processes, supported by data, skills and infrastructure.	 **Data governance complexity** The governance burden and risk associated with data used in AI systems, driven by data type, decision stakes and pipeline complexity.
High	Essential services or major economic constraints with strong relevance for SMEs	Active AI use with observable practices and examples	Personal or sensitive data with high-stakes decision-making
Medium	Clear productivity gains but narrower social or system-wide effects	Emerging or partial AI adoption, often tool- or vendor-led	Mixed personal and nonpersonal data with moderate risk
Low	Limited relevance for SMEs or low development value	Minimal AI use beyond basic digital tools	Predominantly nonpersonal or operational data

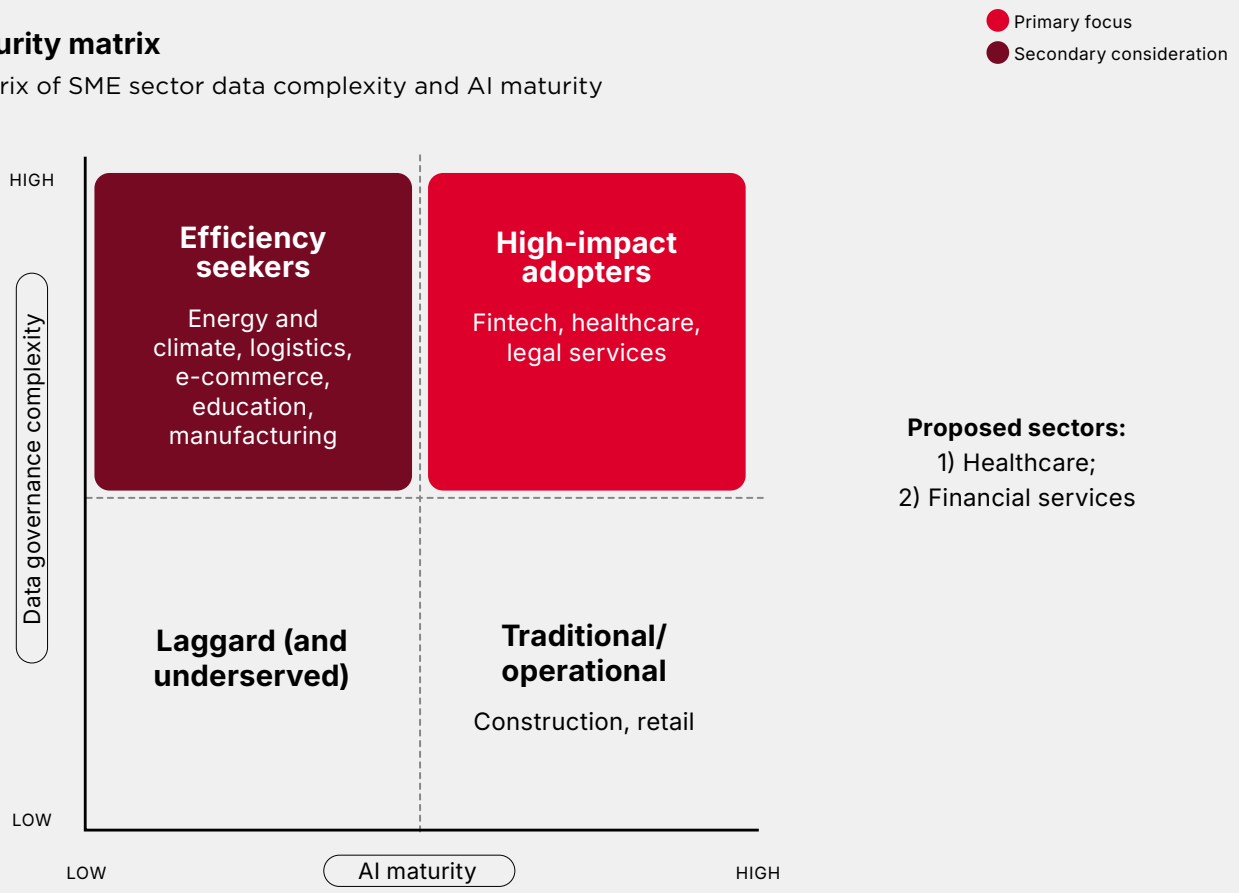
Sectors were mapped across four quadrants within the matrix.

- **High-impact adopters:** Sectors with both high AI maturity and high data governance complexity were identified as the primary focus, as they combine the greatest potential for AI-driven development impact with the most acute governance challenges.
- **Efficiency seekers:** Sectors with high data governance complexity but lower AI maturity, including energy, logistics, e-commerce and education, were identified as secondary considerations.
- **Traditional/operational:** Sectors such as construction and retail, which show high AI maturity but lower data governance complexity.
- **Laggard:** Sectors with low scores on both dimensions, which were excluded from the primary scope.

Figure 18

AI maturity matrix

2x2 matrix of SME sector data complexity and AI maturity



Healthcare and financial services were selected as the two priority sectors because both fall clearly within the high-impact adopter quadrant and because, despite sharing high data sensitivity and active AI adoption, they present distinct governance dynamics that make them complementary rather than duplicative as case studies.

In healthcare, the primary governance tensions centre on consent for secondary data use, accountability across multi-actor public-private delivery chains and the absence of sector-specific AI standards for clinical applications. Data flows are largely internal to health systems, and the populations affected are defined by clinical need.

In financial services, the primary tensions centre on data access asymmetry, automated decision-making accountability and explainability in credit decisions. Data flows are platform-mediated and cross-institutional, involving MNOs, banks and third-party AI providers in chains where accountability is distributed but unallocated. The populations affected are defined by financial exclusion, when the same groups that AI alternative credit scoring is designed to serve are also those most exposed to its governance failures.

Examining both sectors provides a more complete picture of the AI data governance challenge than either would alone, and generates findings and recommendations that are applicable across South Africa's AI-active SMEs.

Annex 3:
Existing laws, policies and frameworks on AI data governance for SMEs

Protection of Personal Information Act 4 of 2013 (POPIA)

Geographic reach
National

Legal status
Binding act

Target businesses
Almost all entities that process personal information, including micro- and small enterprises and high-growth startups

Impact of data life cycle on SMEs/startups

- Data collection must be limited to what is necessary for a defined purpose, requiring SMEs to be intentional from the outset about what data they collect and why.
- Obligations around secure storage, access controls and breach notification shape SME decisions on cloud provider selection and internal data management.
- Rules on profiling, analytics and automated decision-making mean that startups using scoring, recommendation or targeting tools must ensure their data uses are explainable and respect rights.
- Restrictions on data sharing and cross-border transfers affect how SMEs structure contracts with third-party vendors, foreign partners and multinational platforms.
- Retention and deletion requirements, including secure de-identification, have direct implications for backup policies and data infrastructure design.

Draft National AI Policy Framework (2024)

Geographic reach
National

Legal status
Nonbinding draft policy

Target businesses
AI developers, data-rich platforms and SMEs that use or build AI systems

Impact of data life cycle on SMEs/startups

- Signals that training and test datasets should be representative, lawful and ethically sourced, affecting how startups design data-collection and labelling pipelines.
- Emphasises transparency and accountability, encouraging better documentation and logging across the AI data life cycle.
- Likely to inform future binding rules and funding criteria, rewarding SMEs that already have strong data governance practices in place.

National Data and Cloud Policy (2024)

Geographic reach
National

Legal status
Nonbinding policy

Target businesses
Cloud and data infrastructure providers; large data holders; SMEs that rely on cloud or work with public sector data

Impact of data life cycle on SMEs/startups

- Signals that certain data categories may need to be stored within South Africa or on designated infrastructure, influencing SME hosting decisions.
- Aims to increase availability of public sector data, expanding the datasets startups can use for analytics and AI.
- May lead to new cross-border transfer and data localisation rules that affect where SME customer and transaction data can be stored or mirrored.

National Cybersecurity Policy Framework for South Africa

Geographic reach
National

Legal status
Nonbinding policy

Target businesses
All organisations that hold or process data, with an emphasis on critical sectors

Impact of data life cycle on SMEs/startups

- Sets expectations for protecting critical information infrastructure and reporting serious incidents, influencing how SMEs in key sectors design security and data resilience.
- Supports development of national capabilities such as cybersecurity incident response teams, which SMEs may need to engage when sharing incidents and log data.
- Provides a policy basis for future sector-specific cybersecurity requirements that will shape SME data practices.

Consumer Protection Act 68 of 2008 (CPA)

Geographic reach
National

Legal status
Binding act

Target businesses
All enterprises dealing with consumers, including micro- and small enterprises where the consumer threshold is met

Impact of data life cycle on SMEs/startups

- Regulates direct marketing and fair contract terms, affecting how SMEs build and use marketing databases and loyalty programme data.
- Requires clear disclosure of key information, so customer and product data must be accurate and well-structured at the point of sale.
- Works alongside POPIA to govern when contact details can be reused, updated or deleted when consumers withdraw consent.

Open Data South Africa / data.gov.za

Geographic reach
National

Legal status
Nonbinding portal and policy practice

Target businesses
Startups and SMEs that reuse government data; public bodies publishing data

Impact of data life cycle on SMEs/startups

- Increases availability of machine-readable public datasets that SMEs can use for analytics and AI.
- Sets expectations on formats and licences, governing how SMEs can store, integrate and reuse public data in products and services.
- Can reduce data acquisition costs for SMEs, enabling more experimentation in analytics and AI services that rely on government data.

African Union Data Policy Framework (2022)

Geographic reach
International

Legal status
Nonbinding framework

Target businesses
Governments, regulators and enterprises engaged in cross-border digital trade in Africa

Impact of data life cycle on SMEs/startups

- Encourages countries to open and share more data across borders, potentially expanding datasets available to startups operating regionally.
- Envisages trusted data spaces and interoperability frameworks, shaping technical and governance standards for data exchange.
- Underlines the need to balance innovation with rights, informing future regulatory conditions for regional SME data use.

Annex 4:
The AI Impact Index

AI Impact Index: methodology and scoring

The AI Impact Index was developed by GSMA Intelligence to provide a standardised measure of AI adoption and maturity among surveyed SMEs. Rather than relying on self-reported assessments of AI capability, the index draws on observable variables such as investment capacity, breadth of adoption and deployment status to produce a composite score that allows comparison across organisations and sectors. The index is used throughout this report to contextualise findings on governance readiness, compliance investment and the relationship between AI maturity and data governance capacity.

Table 5: AI Impact Index definitions and scoring framework

Component	Description
Definition	AI impact measures the effect of AI on an organisation, including its capacity to invest in AI, the extent to which AI is adopted and applied and the deployment status of AI use cases within the organisation.
Index variables	I. Organisational capacity to invest in AI II. AI adoption and application III. Status of AI use case deployments. Based on survey questions Q9–Q11.
Scoring range	0–100%. Higher scores indicate greater AI impact and adoption maturity.

Category	Score range	Description
Low AI impact	Below 33%	Organisations that have either not deployed or do not plan to deploy the majority of assessed use cases.
Medium AI impact	33% to below 66%	Organisations that have deployed or plan to deploy most use cases, typically in prototype, proof of concept or pilot phases. Higher scores within this range indicate more advanced implementation stages and deployment status.
High AI impact	66% and above	Organisations that have already deployed most AI use cases, with many in large-scale implementation or early adoption phases. Higher scores within this category reflect greater maturity in both deployment progress and operational status.

Benchmark guidance	
Score interpretation	Higher scores indicate greater AI activity and adoption. Scores should be interpreted in context: a 65% score may represent significant progress in one sector, while a 45% score may be equally substantial in a sector that traditionally lags in technology adoption, such as agriculture.

GSMA Head Office
1 Angel Lane
London
EC4R 3AB
United Kingdom

gsma.com

GSMA
Mobile for Development

M4D