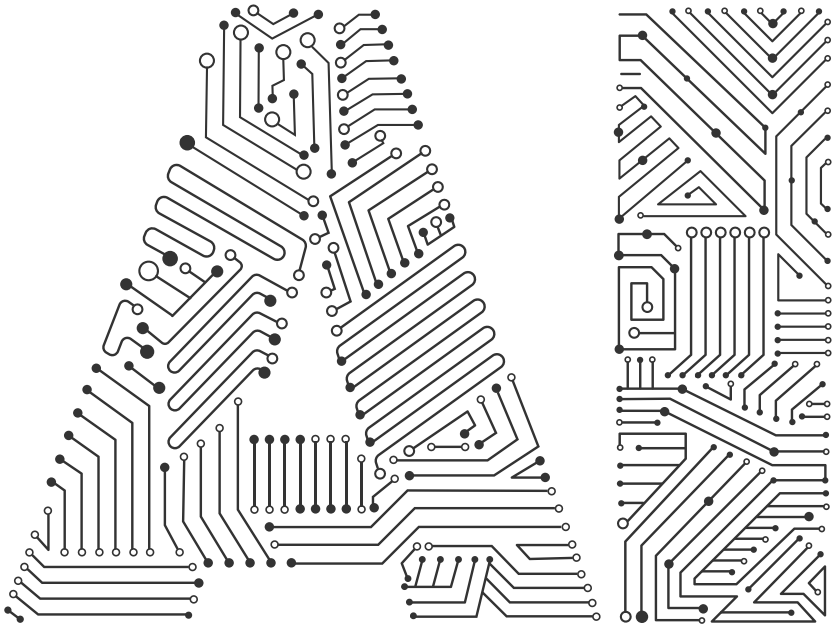


# Artificial Intelligence in the South African Financial Sector

by Nolwazi Hlophe and Lebogang Mabetha

November 2025

FSCA
Financial Sector Conduct Authority

South African Reserve Bank
Prudential Authority

# CONTENTS

## FIGURES

# EXECUTIVE SUMMARY

In the South African financial sector, artificial intelligence (AI) began with basic automation and data analysis tools. Over time, machine learning, natural language processing and advanced analytics were integrated into AI, thus enabling local financial institutions to process vast amounts of data with speed and accuracy.

This evolution paved the way for applications such as predictive analytics, algorithmic trading, and hyper-personalised financial services. Adopting AI in finance offers benefits such as enhanced efficiency, improved accuracy, personalisation and scalability. Automation of routine tasks reduces operational costs and frees up human resources for strategic activities. AI-driven data analysis minimises human errors, leading to more accurate financial forecasting and improved decision-making.

AI enables financial institutions to offer tailored products and services, thus improving customer satisfaction and loyalty. AI solutions can be scaled up to support business growth, for example by increasing the capacity to process higher volumes of data and transactions.

However, despite its present benefits and future potential, the use of AI in the financial sector presents challenges.

The challenges lie particularly in the areas of consumer protection, market conduct, ethics and financial stability. Ensuring the protection of sensitive financial data requires robust cybersecurity measures. Navigating the complex regulatory landscape and ensuring AI systems comply with legal requirements is critical. Addressing ethical issues such as bias in algorithms and the impact on employment is essential for responsible AI adoption.

The joint survey conducted by the Financial Sector Conduct Authority (FSCA) and the Prudential Authority (PA) in 2024, as part of the market research into AI within the South African financial sector, uncovered the following insights:

1. **AI Adoption:** Banking institutions are the leading adopters of AI in South Africa's financial sector, with approximately over 50%, followed by payments institutions with 50% adoption.
2. **Investment Intentions:** Most institutions plan to invest less than R1 million in AI, indicating cautious spending and a focus on incremental implementation for 2024. However, the banking sector demonstrates a strong inclination towards substantial AI investment, with over 50% planning to invest more than R20 million in the same period.
3. **Leading Applications:** Operations and IT are the primary areas for traditional AI applications, while sales and marketing lead for Generative AI (GenAI) applications.
4. **Benefits:** Data and analytical insights are the highest perceived benefit of traditional AI usage, followed by productivity.
5. **Risks:** Data privacy and protection are the most significant risks across all sectors, followed by cybersecurity.
6. **Constraints to adoption:** Data privacy and protection laws are the most significant regulatory constraints, while insufficient talent and appropriate transparency and explainability also are significant non-regulatory constraints.
7. **Governance Frameworks:** Established governance frameworks applicable to AI include data governance, model development, validation, and ongoing monitoring.
8. **Ethical Considerations:** Data privacy is the leading ethical consideration. It is followed by security and accuracy.
9. **Use Cases:** Fraud detection and product/service promotion are the leading use cases for machine learning (ML) and GenAI respectively.

In addition, the research considered both microprudential and macroprudential risks, highlighting how AI can introduce vulnerabilities at the institutional level, such as model risk and operational dependencies, while also posing systemic risks through interconnectedness, procyclicality, and market concentration. While AI holds disruptive potential for South Africa's financial sector, its successful integration hinges on a balanced approach that prioritises innovation, ethical considerations, and regulatory compliance to ensure sustainable growth and enhanced customer trust.

As AI technologies continue to evolve, their integration into South Africa's financial sector will require a coordinated and forward-looking approach. The FSCA and PA have identified several key learnings from the data analysis:

- **Explainability and transparency:** Recognised explainability methods – such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) – can promote AI-driven decisions being more understandable and auditable.
- **Governance frameworks:** Comprehensive governance structures, including strong data governance, model risk management and board-level oversight, can promote ethical and effective AI deployment.
- **Regulatory coordination:** It will be important for the PA and FSCA to consider how to support the Information Regulator and alignment with the Protection of Personal Information Act 4 of 2013 (POPIA), particularly in relation to data privacy and consumer protection.
- **Prioritisation of use cases:** To benefit consumers, the sector may consider prioritising positive-impact use cases, whilst taking suitable steps to manage any associated risks.
- **Ethical standards and oversight:** The development of sector-wide guidance for ethical, fair and responsible AI is contemplated, alongside enhanced oversight to mitigate bias, inaccuracies and consumer harm.
- **Efficient and effective disclosure:** Clear disclosure when AI is used in consumer-impacting decisions, such as credit assessments or insurance pricing, can assist consumers in understanding how decisions are made, fostering trust, and enabling informed recourse or appeal if needed.
- **Digital and AI literacy:** Promoting consumer education and awareness can help individuals understand how AI affects their financial decisions and rights.

These insights promote a more inclusive, transparent and resilient financial system that leverages AI responsibly to benefit both institutions and consumers.

# PART 1: INTRODUCTION

## 1.1 ABOUT THIS REPORT

This report presents the findings of research that was conducted jointly by the FSCA and PA on the current use of AI in the South African financial sector.

This report addresses a critical knowledge gap by providing the first comprehensive view of how AI is being adopted across South Africa's financial sector. While global discussions on AI in finance are well advanced, there has been limited local evidence to inform regulatory and industry responses.

This report fills that gap by presenting data-driven insights on AI adoption trends, investment intentions, use cases, and associated risks, alongside governance and ethical considerations. It also outlines international regulatory approaches and shares the overarching lessons learnt from the financial sector of South Africa, making it a key resource for policymakers, regulators, and industry stakeholders navigating the safe and responsible integration of AI.

## 1.2 THE SOUTH AFRICAN CONTEXT

The South African financial sector is consistently at the forefront of adopting new technologies to enhance efficiency, security and customer satisfaction. In recent years, AI has emerged as a significant force within the sector, with the potential to innovate various aspects of the financial sector.

AI's journey in the financial sector began with the adoption of basic automation and data analysis tools. Over time, the integration of machine learning (ML), natural language processing (NLP), and advanced analytics has enabled financial institutions to process vast amounts of data with speed and accuracy. This evolution has paved the way for more sophisticated applications, such as predictive analytics, algorithmic trading, and hyper-personalised financial services.

AI's applications in the financial sector are diverse and far-reaching. The adoption of AI in finance offers numerous benefits, including enhanced efficiency, improved accuracy, hyper-personalisation and scalability. Automation of routine tasks reduces operational costs and frees up human resources for more strategic activities. AI-driven data analysis minimises human errors, leading to more accurate financial forecasting and decision-making.

Despite its potential, the integration of AI in the financial sector is not without challenges. Ensuring data governance and management is paramount, through requiring robust cybersecurity measures. Addressing ethical issues, such as bias in AI algorithms and the impact on employment is essential for responsible AI adoption. The successful implementation of AI requires skilled professionals that can develop, manage, and maintain AI systems.

**SA Financial Sector Fast Facts**

- South Africa's financial sector operates under a twin-peaks model with the FSCA and Prudential Authority.
- Consists of 67 banks, 158 registered insurance institutions, approximately 315 lenders, around 27 major payments institutions, over 5000 pension funds and over 200 fintechs.
- Contributes approximately 20% to GDP in South Africa.

## 1.3 DEFINITION OF ARTIFICIAL INTELLIGENCE

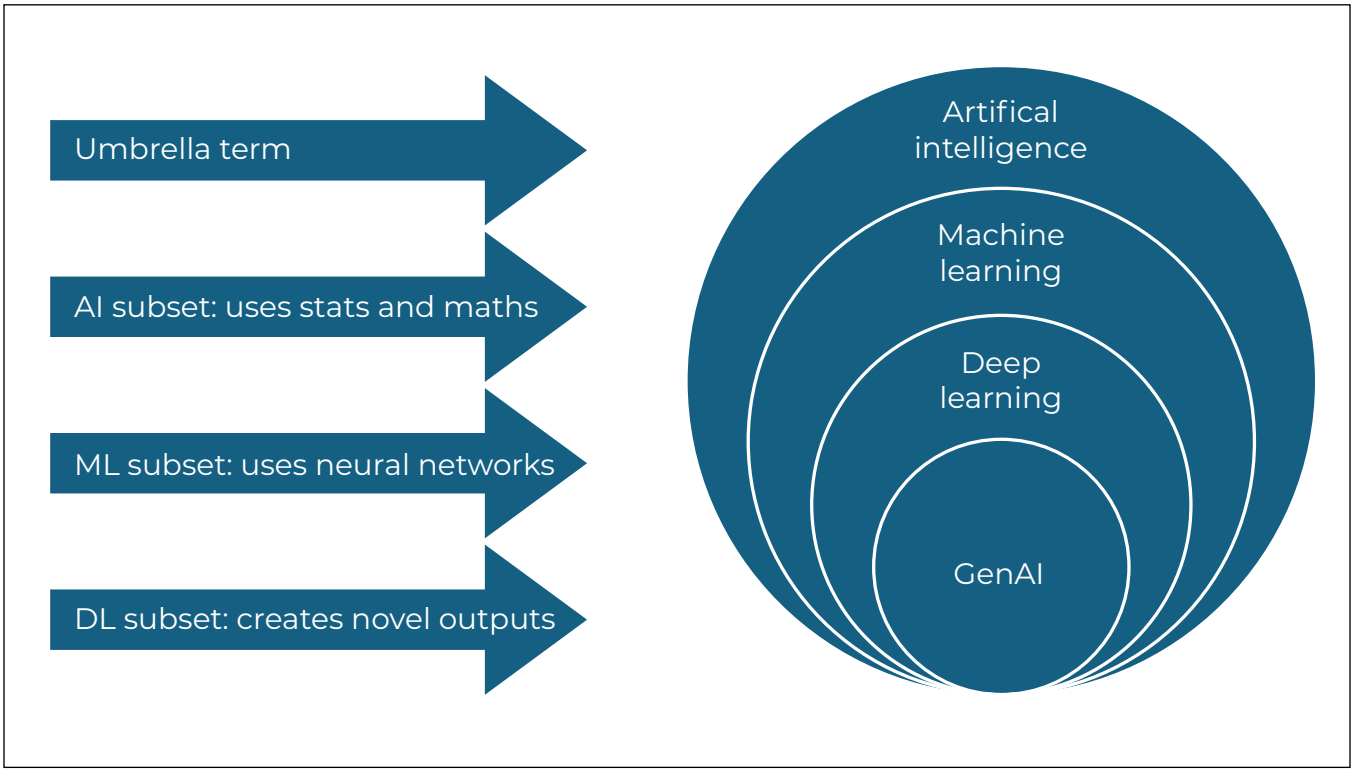
As there is no universally accepted definition of AI, for the purposes of this report, we adopt the Organisation for Economic Co-operation and Development (OECD) definition for AI, which states the following:

"An AI system is a machine-based system that, for explicit or implicit objectives, infers from the input it receives how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment."[1]

The OECD definition is extremely broad and wide ranging, which emphasises how complex and challenging AI is for adoption as well as for regulation and oversight in the financial sector. The definition is particularly relevant because it has been adopted by multiple jurisdictions as the foundation for their AI policy and regulatory frameworks.[2]

The widespread adoption of the OECD definition of AI systems – such as its integration into the European Union (EU) AI Act – supports regulators and financial institutions in developing a shared understanding of what constitutes an AI system. This common framework is essential because it enables consistent identification and classification of AI technologies across jurisdictions, which is particularly important in a globalised financial system. With a harmonised definition, regulators can more effectively coordinate cross-border supervision, align regulatory expectations, and foster international cooperation. For financial institutions, this clarity helps streamline compliance efforts and ensures that risk assessments are based on a universally accepted understanding of AI systems, thereby improving the consistency and reliability of regulatory oversight.

Figure 1: Relationship between AI, ML, and GenAI




SOURCE: IAIS (2024)

[1] *OECD (2024), "Regulatory Approaches to Artificial Intelligence in Finance"*, available here.
[2] *OECD (2024), "Regulatory Approaches to Artificial Intelligence in Finance"*, available here.

## 1.4 PURPOSE AND OBJECTIVES OF THE REPORT

This report reflects the findings of a comprehensive study that was conducted across South Africa's financial sector to determine how AI is being adopted in the financial sector. The study focused on the scope and impact of AI technologies, the current level of investment and use of AI, the specific applications of ML, and the use cases for AI and GenAI. Additionally, the study explored the following:

- Associated risks and benefits
- Regulatory and non-regulatory constraints
- Ethical considerations
- Consumer protection issues
- The implications for financial stability.

The finding from the study as reflected in this report is aimed at bridging a critical information gap that will support and inform regulatory decision-making.

## 1.5 RESEARCH APPROACH AND METHODOLOGY

On 21 October 2024, the PA and FSCA launched a survey aimed at understanding the AI landscape in South African financial services. This voluntary survey received approximately 2 100 responses, covering segments such as banking, insurance, investments, payments, pensions, fintechs, and lending. Across all sectors, 220 respondents (10.6%) use AI in the South African financial sector. This analysis focuses on the banking, insurance and investment segments, as they are the largest segments within the financial sector by asset level. However, the data for the complete survey is available in the databook.

The results presented in this report are anonymised and aggregated with respondents grouped into the sectors listed in Table 1.

Table 1: Sector classification

| Segment | Types of institutions included |
|---|---|
| **Banks** | Deposit-takers, SA and international retail and commercial banks |
| **Insurance** | General insurers, health insurers, life insurers, personal and commercial lines insurers, insurance brokers |
| **Investments** | Asset managers, fund managers, hedge funds |
| **Payments** | System operators, third-party payment providers and payment services providers |
| **Non-bank lending** | Non-bank lenders, credit brokers |
| **Pensions** | Pension funds |

Information submitted by the different financial institutions was augmented by interviews with experts in the field and desk-based research. The desk-based research drew information from various sources, including policy and academic texts, national and international datasets from both public and private stakeholders.

## 1.6 LIMITATIONS OF THE REPORT

This report provides important insights into the adoption and use of AI in South Africa's financial sector; however, several limitations should be acknowledged. First, while the survey was distributed across the entire financial sector – including banking, insurance, investments, payments, pensions, and lending segments – participation was voluntary, and response rates varied across segments. As a result, the findings may not fully represent the entire sector, particularly smaller institutions or niche players. Second, the survey was open for only two months (October to December 2024), which may have limited participation, especially for institutions requiring internal approvals before responding. Consequently, some perspectives, particularly from larger or more complex organisations, may not have been captured.

Third, the report relies on self-reported data, which, despite assurances of confidentiality, may be subject to interpretation bias, optimism bias, or incomplete disclosure, especially in sensitive areas such as investment intentions and risk management practices. Additionally, the survey focused primarily on the current state of AI adoption and short-term investment plans for 2024, without capturing long-term strategic initiatives or emerging use cases that could significantly influence the sector in the future.

Furthermore, while the survey included some open-ended questions, it did not incorporate extensive qualitative interviews or case studies, limiting the depth of analysis on organisational strategies, governance practices, and consumer outcomes. The classification of segments also presents a limitation: fintech respondents were merged into the payments segment for reporting purposes, which may obscure nuances in AI adoption patterns specific to fintechs. Finally, the research did not comprehensively assess external factors such as macroeconomic conditions, infrastructure readiness, or broader digital literacy levels, all of which may influence AI adoption and its impact on the financial sector.

# PART 2: APPROACHES TO AI AROUND THE WORLD

Given that AI is characteristically global and not limited by national borders or jurisdictional boundaries, this section delves briefly into different legislative and regulatory approaches to AI around the world.

We recognise that in helping to shape a regulatory approach to AI that is tailormade to SA and its local financial sector, it is crucial for the FSCA and the PA to draw on the approaches of other countries and jurisdictions, and, importantly, on the lessons learnt. Fortunately, these have been and continue to be documented and shared.

The OECD AI policy observatory, for example, provides detailed data and information on AI policies, strategies, and initiatives across countries and regions, with real-time updates on AI governance. The online platform aims to facilitate international collaboration and to offer insights for developing AI policies aligned with ethical standards and human rights.

Note also that the OECD AI Principles were adopted by the Group of 20 (G20) in 2019 and presently serve as a global benchmark for helping governments and organisations shape a human-centric approach to trustworthy AI. Although South Africa is not a member of the OECD, our country has been a Key Partner since 2007, cooperating on areas such as macro-economic policy and structural reform, among others.

## 2.1 RELYING ON EXISTING REGULATION VS INTRODUCING NEW REGULATION

Around the world, countries are adopting various approaches to govern the use of AI and data. Our desktop research found that there are two approaches which are predominantly considered – and both hold significant implications for financial sectors in their respective jurisdictions.

The first approach sees regulators relying on existing regulatory requirements that do not necessarily distinguish between different types of technologies but instead impose general guidance that may apply when technology is used.

The second approach involves developing technology-specific regulatory requirements that are applied alongside general regulations that also govern the use of technology. A relevant case report is the European Union's approach to regulating artificial intelligence in the financial sector.

While general financial regulations such as the Markets in Financial Instruments Directive (MiFID II) and the General Data Protection Regulation (GDPR) continue to apply, the EU has introduced the AI Act, which imposes additional AI-specific requirements. For example, high-risk AI systems used in credit scoring or fraud detection must meet strict transparency, data governance and human oversight standards.

The EU's layered regulatory model ensures that AI applications in finance are both innovative and accountable, addressing risks that general financial laws alone may not fully capture.[3]

[3] *European Commission. (2025) European Artificial Intelligence Act (AI Act). Available at: https://transition-pathways.europa.eu/policy/european-artificial-intelligence-act-ai-act (Accessed: 14 July 2025).*

## 2.2 AREAS OF COHESION

Global professional services firm EY analysed the regulatory approaches of eight jurisdictions (Canada, China, the European Union (EU), Japan, Korea, Singapore, the United Kingdom (UK), and the United States (US) and found that there is wide agreement on the foundational principles to govern the use of AI.

Although the analysed jurisdictions have adopted legislative and regulatory approaches that align with their unique cultural norms and legal frameworks, there are six areas of cohesion or commonality. The areas of cohesion share the overarching goal of minimising potential harm from AI, while promoting its use for the betterment of society and the economy. These areas include the following:

- Core principles consistent with the OECD AI Principles.
- A risk-based approach tailoring AI regulations to perceived risks.
- A combination of sector-agnostic and sector-specific requirements addressing AI's varying use cases.
- Policy alignment integrating AI-related frameworks within other digital policy priorities.
- Private-sector collaboration using regulatory sandboxes.
- Supervisory approaches conducting thematic reviews and workshops to gather information and identify good practices for AI regulation.

EY elaborates on these areas of cohesion in its 2024 paper entitled *The Artificial Intelligence (AI) Global Regulatory Landscape: Policy Trends and Considerations to Build Confidence In AI.*[4]

Cohesion is beneficial as it helps ensure that AI technologies are developed and used responsibly, balancing innovation with ethical considerations and protecting individuals from potential harm.

Financial institutions worldwide must navigate these evolving regulatory landscapes to leverage AI effectively while maintaining compliance and safeguarding customer trust. Figure 2 illustrates some of the approaches around the world.


[4] *EY (2023), "The Artificial Intelligence (AI) Global Regulatory Landscape: Policy Trends and Considerations to Build Confidence In AI", available here.*

Figure 2: Legislative and regulatory approaches

**European Central Bank (ECB)**
The ECB initiated its opening to AI in 2018, with the creation of a dedicated team within its Directorate General Statistics. The ECB has since developed several AI-based applications, including a chatbot for internal use and an AI-based system for analyzing large volumes of text data. The ECB is also exploring the use of machine learning techniques for improving its economic forecasting models.

**European Union (EU)**
The EU has set up various initiatives to promote the development and adoption of AI technologies across member states. These include funding research projects, creating regulatory frameworks, and fostering collaboration between academia, industry, and government agencies.

**Canada**
Canada is recognised as a global leader in AI research and development, with several world-renowned institutions and companies based in the country. The Canadian government has invested heavily in AI through initiatives such as the Pan-Canadian Artificial Intelligence Strategy.

**United States**
In 2019, the Federal Reserve launched an initiative to explore how artificial intelligence could be used to improve its operations and services. This includes using machine learning algorithms to analyse large datasets and enhance decision-making processes.

**China**
China's State Administration for Market Regulation (SAMR) oversees product quality standards across various industries including electronics manufacturing where many components are produced using automated systems powered by artificial intelligence technology.

**United Arab Emirates (UAE)**
The UAE's National Innovation Strategy aims at positioning the UAE among top countries globally when it comes to an innovation-driven economy leveraging advanced technologies like Artificial Intelligence (AI). Several smart city projects have been launched incorporating these innovations into urban planning, infrastructure management, public services delivery, etc., making life easier and more efficient for residents and visitors alike.

**Singapore Government**
The Singapore Government has implemented policies to support growth and adoption within the local ecosystem, encouraging businesses to integrate AI solutions into their operations, thereby boosting productivity and competitiveness of the nation overall.

**Monetary Authority of Singapore (MAS)**
MAS established a FinTech Regulatory Sandbox to provide an environment for financial institutions to test innovative products and services without immediately incurring all normal regulatory consequences of failure. This allows experimentation with new ideas while ensuring consumer protection and market integrity are maintained throughout the process. Additionally, MAS actively collaborates with international counterparts to share knowledge and best practices related to regulation and supervision of emerging technologies.

**Spanish Agency for Data Protection (AEPD)**
The AEPD was created in December 2020 to oversee the implementation of a national strategy promoting ethical and responsible use of personal data, ensuring compliance with relevant laws and regulations, and protecting citizens' privacy rights in the digital age.

**South African Reserve Bank (SARB)**
SARB introduced a Conceptual Framework to regulate fintech innovations, focusing primarily on areas like crowdfunding, virtual currencies, and blockchain technology, aiming to create a comprehensive and robust governance structure.

## Box 1: South Africa's regulatory approach to AI automated advice

The EU's General Data Protection Regulation (GDPR) in Article 22 establishes a right for data subjects not to be subjected to decisions (including profiling) based solely on automated processing that produce legal or significant effects – unless specific conditions are met, such as contractual necessity, legal authorisation, or explicit consent.

South Africa's Protection of Personal Information Act (POPIA) includes similar provisions aiming to protect individuals from automated decisions with legal or substantial effects, particularly those based on profiling. While POPIA does not explicitly define 'automated processing' or 'profile,' it aligns with GDPR's intent to safeguard data subjects from potentially harmful automated decisions.

POPIA Section 71 allows for automated decision-making under certain circumstances, such as when it is necessary for a contract and the data subject's request has been met, or when appropriate measures are in place to protect their legitimate interests.

These measures include allowing data subjects to make representations about the decision and providing them with sufficient information about the underlying logic of the automated process. For instance, in loan applications, automated approval is permitted, but rejections require additional safeguards like human review and an appeal mechanism to ensure compliance and fairness. Automated decisions are also permissible if governed by a law or code that specifies appropriate protective measures.

South Africa's financial sector provides a practical example of these regulations in action. The Determination of Fit and Proper Requirements, 2017, issued under the Financial Advisory and Intermediary Services Act, 2002 (Act No. 37 of 2002) (FAIS Act), recognise and regulate 'automated advice.' Financial services providers offering such advice must adhere to certain standards, which include ensuring human resources and competence, oversight (including monitor, review and testing of algorithms), implementing robust internal controls, maintaining system governance, and possessing adequate technological resources/ infrastructure. In addition, the General Code of Conduct for Authorised Financial Services Providers and their Representatives, 2003 issued under the FAIS Act contains various general principles-based requirements that promote consumer protection and addresses themes such as conflicts of interest, advertising, disclosure, risk management, and the like. These requirements are technology-neutral, meaning that it will equally apply to, for example, 'robo-advisors'. Looking ahead, the sector-wide Conduct of Financial Institutions Bill (CoFI Bill) aims to further regulate digital innovations in the financial sector based on outcomes, reinforcing consumer protection.

The PA plays a critical role in ensuring that AI-driven processes, including automated advice, do not compromise the safety and soundness of financial institutions. Its regulatory framework emphasises robust risk management, governance, and operational resilience, requiring institutions to maintain adequate controls when deploying AI technologies. In addition, the PA's prudential standards mandate that financial institutions assess and mitigate model risk, third-party dependencies, and systemic vulnerabilities introduced by automation and advanced analytics.

# PART 3: THE AI ECOSYSTEM IN FINANCIAL SERVICES

The AI ecosystem in financial services in SA is complex and features a variety of applications, technologies and key players that collectively drive innovation and efficiency.

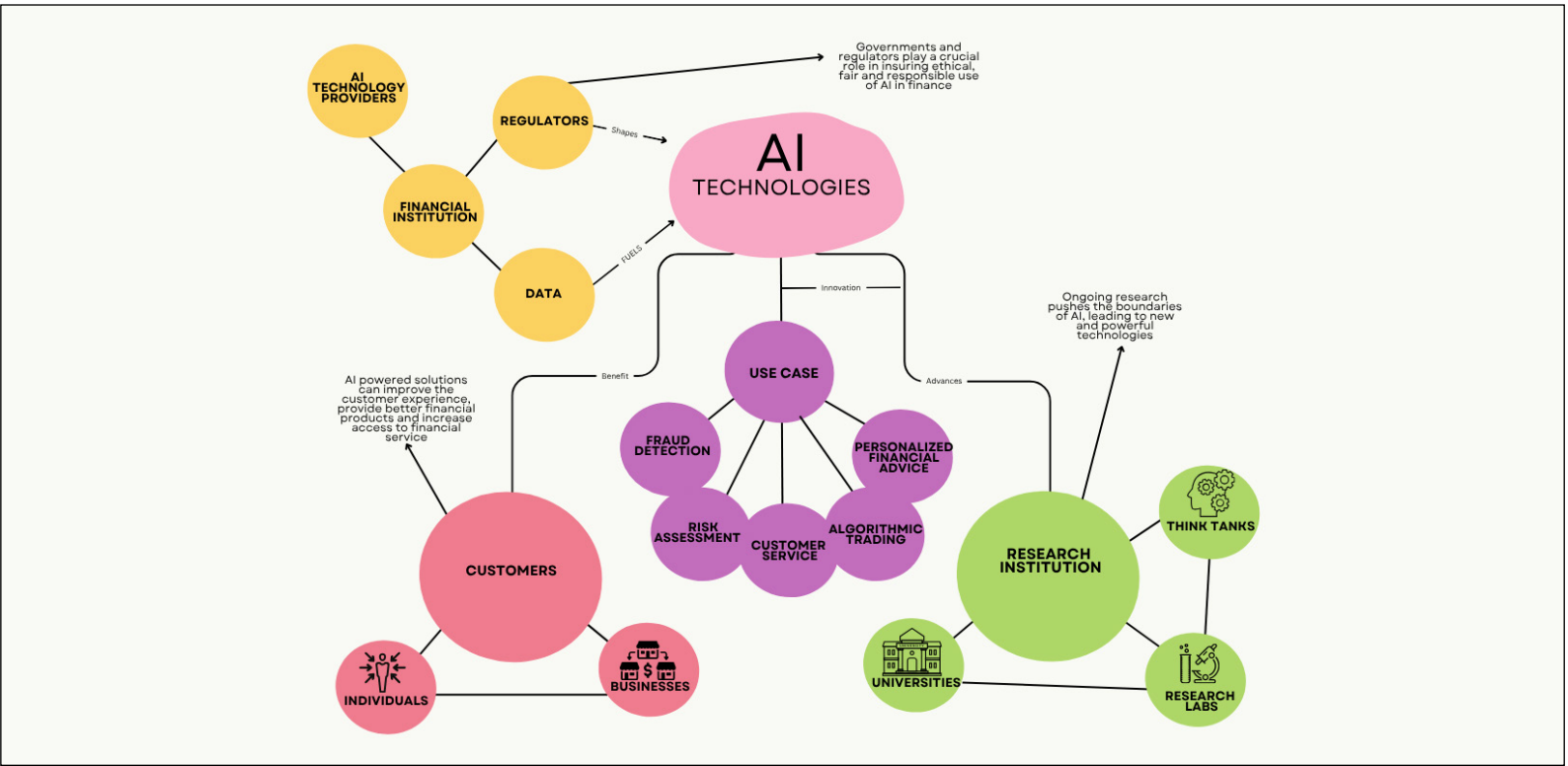
Financial institutions, technology providers, and fintech startups are central to this transformation, using AI to enhance decision-making, improve customer experiences, and manage risks more effectively. Major banks and financial institutions use AI for tasks such as fraud detection and personalised customer service, while technology providers supply the necessary infrastructure and AI tools, including cloud computing and machine learning platforms, to enable the implementation of AI solutions. Fintech startups play a significant role, often leading the development of innovative AI-driven financial products and services.

The technologies and techniques used in AI for finance are diverse and advanced. ML algorithms are essential for predictive analytics, risk assessment, and customer segmentation, analysing historical data to identify patterns and make accurate predictions about future trends. NLP allows AI systems to understand and process human language, which is crucial for applications like sentiment analysis, automated reporting, and chatbots.

Anomaly detection techniques help identify unusual patterns in data, aiding in the detection of potential fraud or operational issues. Additionally, recommendation systems offer personalised suggestions for financial products and services based on customer behaviour and preferences.

Ideally, regulatory bodies should ensure that AI applications in the financial sector comply with legal and ethical standards, while data providers supply the large datasets needed for training AI models and ensure the accuracy and reliability of data used in AI applications. Together, these components can contribute to the dynamic and multifaceted nature of the AI ecosystem in the financial sector, enabling institutions to enhance decision-making, improve customer experiences, and manage risks more effectively. Figure 3 illustrates the AI ecosystem in the financial system as described above.

Figure 3: AI ecosystem in finance



## 3.1 RELYING ON EXISTING REGULATION VS INTRODUCING NEW REGULATION

The adoption of AI in finance is driven by targeting increased efficiency, cost reduction, improved decision-making, and enhanced customer experience. These factors drive financial institutions to leverage AI technologies to stay competitive and meet market demands:

- **Efficiency and cost reduction**: AI-powered systems can enhance efficiencies by automating routine tasks like data entry, transaction processing and customer service, cutting down on manual labour and operational expenses. This automation can speed up processes, minimise errors, and lead to more accurate financial operations. For instance, robotic process automation (RPA) reduces manual labour and errors in tasks such as invoice processing and reconciliation, freeing resources for strategic initiatives.[5]

- **Improved decision-making**: Technologies such as ML and predictive analytics enable financial institutions to make better-informed decisions. By analysing vast amounts of data in real-time, these AI technologies can identify patterns, trends, and anomalies that would take much longer to process by human analysts or be missed entirely. This capability is particularly valuable in areas like risk management, fraud detection, and investment strategies, where timely and accurate insights can lead to more effective and profitable outcomes.[6]

- **Enhanced Customer Experience**: AI is also transforming the way financial services are delivered to customers. Through personalised recommendations, chatbots, and virtual assistants, AI can provide tailored financial advice and support, improving customer satisfaction and engagement. Additionally, AI-driven tools can offer 24/7 assistance, ensuring that customers have access to financial services whenever they need them, thus enhancing overall user experience. Figure 4 further shows other market drivers for AI adoption in financial services.

Figure 4: Market drivers of AI in finance

**Advancements in new technologies**

Increased computing power, storage capabilities and new algorithms are enabling programmers to grow AI applications exponentially. Industries are utilising these advancements to automate business processes, gain insight into product behaviour and customer actions and improve decision-making.

**Ongoing growth of cloud computing**

The major cloud providers continue to invest in building out their cloud infrastructure, even when it has been economically risky to do so (think rising interest rates). The major cloud providers have made it easier and more cost-effective for other companies to jump in.

**Big data and data availability**

Until recently, data storage was limited to what could be loaded on a local magnetic disk or cartridge. Now cloud-based storage gives companies easy access to huge volume of structured and unstructured data. This access provides significant opportunities for AI-powered data analysis to detect patterns, identify trends and make predictions.

**Increased competition**

Companies are constantly seeking new technologies that will give them advantages over their competitors. AI/ML technologies enhance product development, optimise operations and enable strategic, data-inspired decisions, providing many businesses with improved outcomes, better efficiency and cost savings.

[5] *Finance Alliance (2024), "AI and data analytics-driven finance transformation", available here*
[6] *KPMG (2021), "Artificial Intelligence in Risk Management", available here*

## 3.2 USE CASES OF AI IN FINANCE

AI is increasingly being integrated into the financial sector, offering a range of applications. Table 2 provides an overview of various AI use cases in finance, explaining how these technologies are implemented and the potential value that can be derived from the identified use cases.

By understanding these use cases, financial institutions can better leverage AI to improve their services and manage risks more effectively.

Table 2: The most promising AI use case examples across the financial sector[7]

| Domain | AI Use Case | Purpose/Outcome |
|---|---|---|
| **Retail banking** | Personalised product recommendations | Based on transaction history, life events, and behavioural patterns |
| | Chatbots and virtual assistants | 24/7 customer service, KYC verification,FAQs |
| | Behavioural credit scoring | Use of alternative data (e.g. mobile usage, utility payments) for inclusion |
| **Insurance** | Dynamic pricing and underwriting | Adjusting premiums in real-time based on behavioural or IoT data |
| | Claims triage and fraud detection | Prioritising and verifying claims using NLP and image recognition |
| **Wealth & asset management** | Robo-advisors | Robo-advisors use AI to automatically provide personalised investment advice and portfolio management based on an individual's financial goals, risk tolerance, and market data |
| | Sentiment analysis | Analysing market trends from news and social media feeds |
| | ESG screening | AI tools identifying companies' environmental/social risks |
| **Capital markets** | Algorithmic and high frequency trading | AI models executing trades based on real-time data and predictive signals |
| | Market surveillance | Detecting insider trading and market manipulation using anomaly detection |
| **Lending & credit** | AI-driven credit assessments | Using ML to model default probabilities and expand access to credit |
| | Loan servicing automation | Monitoring repayment behaviours and nudging interventions |
| **Compliance & regulation (regtech)** | AML/CFT transaction Monitoring | Identifying suspicious patterns across transaction networks |
| | Regulatory reporting Automation | NLP systems extracting and structuring data for compliance filings |
| | Risk model validation | AI models cross-checking and validating traditional risk frameworks |

[7] *World Economic Forum (2025), "Artificial Intelligence in Financial Services", available here.*

| | | |
|---|---|---|
| **Cybersecurity & fraud** | Biometric authentication | Face/voice recognition for secure logins |
| | Real-time fraud prediction | Adaptive learning to detect new fraud typologies |
| **Operations & infrastructure** | Process automation (RPA + AI) | Automating repetitive back-office tasks (reconciliations, KYC) |
| | IT systems optimisation | Predictive maintenance, capacity forecasting |
| **Human resources & strategy** | Talent analytics | Predicting attrition, optimising recruitment pipelines |
| | Strategic scenario planning | Simulating macro/credit/economic scenarios using AI |

In retail banking, AI is used to improve customer service interactions through virtual assistants. AI is also being leveraged to deliver hyper-personalised product recommendations, behavioural credit scoring using alternative data sources (such as mobile phone usage or utility payments), and intelligent risk profiling. These tools can expand access to credit for underbanked populations, supporting financial inclusion objectives.[8]

In the insurance sector, AI enables dynamic pricing and underwriting, drawing on telematics, IoT data, and behavioural analytics to offer more accurate and risk-sensitive premiums.[9] AI models are also used to streamline complex claims processes and detect fraudulent submissions using natural language processing and computer vision tools.[10]

In wealth and asset management, AI is central to the rise of robo-advisory platforms, which provide automated investment strategies based on client risk profiles and market conditions.[11] Asset managers use AI for real-time portfolio optimisation, sentiment analysis derived from financial news and social media, and increasingly, for screening investments against environmental, social, and governance (ESG) criteria.[12]

Capital markets continue to be a leading domain for advanced AI use. Algorithmic and high-frequency trading models rely on machine learning to identify arbitrage opportunities and execute trades within milliseconds.[13] Market surveillance systems powered by AI help regulators and trading firms detect market abuse, insider trading, and other anomalous behaviour.[14]

AI also plays a pivotal role in compliance and regulatory technology (regtech). Financial institutions employ AI to enhance anti-money laundering (AML) and combating the financing of terrorism (CFT) monitoring systems, automate regulatory reporting, and validate internal risk models.[15] These applications reduce compliance costs while improving detection and oversight. As supervisory bodies develop "suptech" tools to monitor systemic risks, AI is becoming an indispensable part of financial oversight infrastructures.[16]

In the area of cybersecurity, AI is used for behavioural biometric authentication and anomaly detection, helping institutions mitigate sophisticated fraud and cyber threats.[17] Similarly, operational domains see widespread adoption of AI-enhanced robotic process automation (RPA) for back-office functions, predictive IT systems management, and data reconciliation.[18]

Finally, governance and strategic decision-making are also being transformed by AI. Financial institutions use advanced analytics for workforce planning, scenario simulation, and macroeconomic stress testing.[19]

In sum, the integration of AI into financial services extends far beyond customer-facing roles. It permeates decision-making, operational resilience, regulatory compliance, risk modelling, and strategic planning. A comprehensive understanding of AI's role in financial services requires recognition of these diverse use cases, including both their potential benefits and associated risks.

---

[8] *World Bank Group, Data-Driven Credit: Risks and Opportunities of Alternative Data in Financial Services (2022) https://documents.worldbank.org accessed 20 May 2025.*
[9] *McKinsey & Company, AI in Insurance: A Quiet Revolution (2021) https://www.mckinsey.com accessed 20 May 2025.*
[10] *Accenture, AI in Insurance Claims Management (2020) https://www.accenture.com accessed 20 May 2025.*
[11] *Financial Conduct Authority (FCA), Evaluation of Automated Investment Services (2023) https://www.fca.org.uk accessed 20 May 2025.*
[12] *OECD, AI and Responsible Investment in ESG (2021) https://www.oecd.org accessed 20 May 2025.*
[13] *European Securities and Markets Authority (ESMA), Report on Algorithmic Trading (ESMA50-157-2394, 2022).*
[14] *European Central Bank, AI in Market Surveillance: Emerging Practices (2023).*
[15] *Financial Stability Board, The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions (2020) https://www.fsb.org accessed 20 May 2025.*
[16] *Bank for International Settlements (BIS), SupTech Applications in Financial Supervision (2021) https://www.bis.org accessed 20 May 2025.*
[17] *ENISA, AI Threat Landscape (2023) https://www.enisa.europa.eu accessed 20 May 2025.*
[18] *Deloitte, AI and RPA in Financial Operations: Integrating Intelligence (2022).*
[19] *IMF, Artificial Intelligence and Macroeconomic Forecasting: The New Frontier (2023) https://www.imf.org accessed 20 May 2025*

# PART 4: OPPORTUNITIES AND RISKS OF AI IN FINANCE

AI use and adoption by financial institutions worldwide has been notable in both back and front office operations to increase efficiency and productivity, while simultaneously benefitting financial consumers. The inherent risks of AI, however, also warrant careful consideration and mitigation.

## 4.1 BENEFITS AND OPPORTUNITIES OF AI

The benefits of AI can be categorised into four main areas: operational efficiency; hyper-personalised customer experiences; risk management and compliance; and data-driven decision-making. Tables 3, 4, 5 and 6 discuss the benefits of AI in the four main areas, in real-life scenarios.

Table 3: Operational efficiency

| | Description |
|---|---|
| **Automation of repetitive tasks** | RPA, which incorporates ML or NLP i.e., "intelligent automation", can automate routine tasks, reduce manual intervention and minimise errors. For example, AI can handle loan application processing, speeding up approvals. |
| **Faster processing** | Automation leads to quicker processing of transactions and applications, enhancing overall efficiency. |
| **Cost reduction** | While automation reduces costs, the initial investment in AI systems (e.g. training, integration, vendor solutions) can be significant. Overemphasising savings may risk oversimplification. |

Table 4: Personalised customer experience

| | Description |
|---|---|
| **AI-powered tools** | Chatbots and virtual assistants provide quick and accurate responses, enhancing customer satisfaction. For example, a chatbot can help customers to reset their passwords at any time of day. |
| **Hyper-personalised services** | AI analyses customer data to offer tailored financial products and services, ensuring personalised experiences. Chatbots and virtual assistants are able to provide investment advice based on individual financial goals. |
| **Seamless interactions** | AI tools handle a wide range of customer interactions, from basic queries to complex transactions. This means customers can get help with everything from checking their balance to applying for a mortgage, all through a single interface. |

Table 5: Risk management and compliance

| | Description |
|---|---|
| **Fraud detection** | AI analyses data in real-time to detect patterns and anomalies, identifying potential fraud quickly and accurately. For example, it can flag suspicious transactions that deviate from a customer's usual spending habits. |
| **Credit risk assessment** | AI evaluates the creditworthiness of loan applicants through comprehensive analysis, improving risk management. This means more accurate assessments and fewer defaults. |
| **Real-time monitoring** | Continuous monitoring of transactions and activities helps in early detection and prevention of risks. Imagine an AI system that constantly watches for signs of fraud, providing peace of mind for both the bank and its customers. |
| **Operational risk management** | AI systems can monitor operational processes in real-time, identifying inefficiencies and potential risks. This helps financial institutions improve their operational resilience and reduce the impact of disruptions. AI-powered tools can automate repetitive tasks, freeing up human resources to focus on more strategic activities. |

Table 6: Data-driven decision-making

| | Description |
|---|---|
| **Data analysis** | AI systems enable the rapid processing and analysis of large, complex datasets to identify trends, correlations, and anomalies. These insights support evidence-based decision-making in areas such as portfolio management, credit modelling, and product development. |
| **Market trend assessment** | "Better decisions and higher returns" may overstate outcomes. AI enhances the basis for decision-making, but outcomes depend on broader factors (e.g., governance, market volatility, regulatory constraints). |
| **Enhanced competitiveness** | Data-driven decisions help financial institutions stay competitive in the market. By leveraging AI insights, banks can stay ahead of the curve and offer innovative products and services. |

From a regulatory perspective, artificial intelligence offers significant potential to support both market conduct and prudential oversight objectives. RegTech encompasses AI-enabled tools used by financial institutions to comply with regulatory obligations, such as real-time transaction monitoring for AML and CFT, automated regulatory reporting, stress testing, and credit risk modelling. At the same time, regulators themselves are increasingly exploring supervisory technology applications, which enable the analysis of high-frequency market data, anomaly detection in trading behaviour, and predictive risk analytics to enhance supervisory responsiveness.[20]

[20] *BIS (2024), "Regulating AI in the financial sector: recent developments and main challenges", available here*

## 4.2 RISKS OF AI IN FINANCE

These risks can be grouped into four main clusters: consumer protection and market conduct risks; microprudential risks; macroprudential risks; and other risks. Table 7 provides a non-exhaustive list of these risks as highlighted by the adoption of AI in the financial sector.[21]

Table 7: Risks of AI in finance

| Risk Category | Risk Type | Description/example |
|---|---|---|
| **Consumer protection and market conduct risk** | Price collusion | AI systems designed to optimise pricing may learn to replicate or anticipate competitors' pricing strategies, resulting in tacit collusion without direct human intent. This undermines competitive market dynamics and may breach competition law. |
| | Limited disclosure | Limited disclosure of AI usage in finance can lead to misinformed decisions due to a lack of transparency about how AI influences financial products and services. This can result in unexpected outcomes, such as unanticipated fees, biased lending practices, or privacy concerns. |
| | Cyber risk | AI-driven cyber risks to consumer protection stem from how the technology is used to target individuals. Cybercriminals can leverage generative AI to create highly realistic and convincing phishing emails, deepfake audio, and video, making it easier to execute elaborate scams and impersonation fraud. The use of AI also requires financial institutions to collect and analyse vast amounts of customer data, creating a larger attack surface that, if breached, could lead to widespread data theft and identity fraud. Furthermore, if AI systems are trained on biased data, they can make discriminatory decisions regarding credit, loans, and other financial products, unfairly impacting consumers. |
| | Unfair treatment and poor consumer outcomes | • Unfair treatment of customers: Use of AI models trained on biased or incomplete datasets may result in discriminatory outcomes, such as higher insurance premiums for certain demographic groups.<br>  • Automated decision-making may reinforce exclusionary practices by profiling consumers as high-risk based on proxies in alternative data (e.g., location, device type).<br>• AI-enabled pricing models may exploit consumer vulnerabilities (e.g., behavioral patterns) to personalise and increase pricing unfairly. |

[21] *BIS (2024), "Regulating AI in the financial sector: recent developments and main challenges", available here.*

| | | |
|---|---|---|
| **Macroprudential/ financial stability risk** | Interconnectedness and concentration | Increased interconnectivity amongst firms from highly concentrated AI third-party providers could result in systemic risk if those third parties suffer from cyber-attacks or operational failures, affecting multiple financial institutions and markets simultaneously. |
| | Opacity and complexity | Limits to the explainability of certain complex AI models can result in risk management challenges, as well as lesser financial institution and supervisory insight into the build-up of systemic risks. |
| | Pro-cyclicality | AI models trained on past data may reinforce past cycles. AI models trained on historical trends may reinforce existing cycles of risk-taking (e.g. credit booms), and widespread use of similar models can contribute to correlated behaviour, amplifying market volatility. |
| | Herding behaviour | Common use of similar models across firms may amplify market volatility. |
| | Fire-sale dynamics | Algorithmic trading models may trigger simultaneous market exits. |
| **Microprudential risk** | Credit risk | Underestimation of probability of default or risk of loss due to inaccurate data inputs. |
| | Model risk | · Inaccurate model output due to the model not capturing changes to the nature of the data input.<br>· Lack of model explainability hinders the ability to assess its conceptual/technical soundness.<br>· Inaccurate model output due to overfitting or underfitting; that is, the model output cannot generalise to other conditions or circumstances, or it is too simplistic and hence fails to capture the underlying patterns in the data.<br>· Hallucination, inconsistent responses and dependency on data quality.<br>· Overestimation of the capabilities of AI models, leading to misuse of such models beyond their capabilities.<br>· AI models may not produce reliable predictions if they are not trained with the most latest information available. |
| | Other operational risk | · Institutions with legacy IT systems may add complexity to their IT architecture, thus increasing potential operational risks arising from IT failures.<br>· Increased use of third-party services (data providers, AI model providers) could lead to dependency, disruption of critical services and lack of control of processes, which may be exacerbated by vendor lock-in risk and increased market concentration.<br>· Quick obsolescence of risk controls due to rapid updates by AI systems. |

| | | |
|---|---|---|
| | Cyber risk | AI's impact on prudential risks relates to the stability and resilience of the financial system itself. A key concern is the potential for systemic risk if financial institutions become overly reliant on a small number of third-party AI providers. A cyberattack on one of these providers could lead to a cascading failure across the sector. Additionally, AI contributes to a more complex and rapidly evolving cyber threat landscape, as malicious actors use the same technology to create more sophisticated and automated attacks. This forces financial institutions into an escalating "arms race" to maintain their defences. Finally, the integrity of the AI models themselves poses a risk, as a malicious actor could "poison" a model's training data, causing it to make flawed financial decisions that could lead to significant losses for the institution. |
| **Other risks** | Data privacy risk | AI models, particularly large language models (LLMs), may inadvertently reveal or infer personal or sensitive information present in training datasets. This can result in regulatory breaches (e.g. under GDPR or POPIA), reputational damage, or legal liability. |
| | Market competition risk | The high cost of developing and maintaining AI technologies may limit their adoption for smaller financial institutions, potentially increasing the market power and systemic importance of larger firms, while making it difficult for smaller firms to compete. |

AI can be a double-edged sword for cyber resilience. AI can enhance cybersecurity by detecting threats and identifying vulnerabilities through data analysis. It can forecast potential cyber-attacks and improve security measures. However, cyber criminals can also use AI to conduct sophisticated attacks, making them harder to detect and prevent. The increased use of AI in the financial sector necessitates that regulators enhance cybersecurity requirements in their regulatory frameworks. Since AI introduces new vulnerabilities and risks, it requires a more robust and adaptable security framework.

# PART 5: AI GOVERNANCE AND ETHICAL CONSIDERATIONS

AI governance in finance is crucial for harnessing the benefits of AI while managing its risks. The Authorities maintain their role as outcomes-focused regulators, applying a technology-agnostic and principle-based approach where appropriate.

## 5.1 AI GOVERNANCE CONSIDERATIONS

Several international frameworks provide comprehensive non-binding principles for AI governance. The OECD AI Principles, for instance, emphasise the importance of inclusive growth, sustainable development, and well-being. They advocate for human-centred values and fairness, transparency and explainability, robustness, security, and safety, as well as accountability.[22] Similarly, the EU AI Act categorises AI systems based on their risk levels and imposes specific obligations on high-risk AI systems to ensure transparency, accountability, and human oversight.[23]

The United Nations has also proposed guidelines that focus on the ethical use of AI, promoting human rights and fundamental freedoms.[24] These frameworks collectively aim to foster innovation while ensuring that AI technologies are developed and deployed responsibly and ethically across various sectors, including finance. Figure 5 illustrates key general non-binding AI governance efforts from around the world, which are designed to apply to AI in general.

Figure 5: Prominent non-binding AI Governance efforts

| | |
|---|---|
| **Principles** | • OECD AI Principles<br>• European Commission's Ethics Guidelines for Trustworthy AI<br>• UNESCO Recommendation on the Ethics of AI<br>• The White House Blueprint for an AI Bill of Rights<br>• G7 Hiroshima Principles |
| **Laws and regulations** | • EU AI Act<br>• EU Product Liability Directive, proposed<br>• EU General Data Protection Regulation<br>• Canada – AI and Data Act, proposed<br>• U.S. AI Executive Order 14110<br>• Sectoral U.S. legislation for employment, housing and consumer finance<br>• U.S. state laws, such as Colorado AI Act, Senate Bill 24-205<br>• China's Interim Measures for the Management of Generative AI Services<br>• The United Arab Emirates Amendment to Regulation 10 to include new rules on Processing Personal Data through Autonomous and Semi-autonomous Systems<br>• Digital India Act |

[22] *OECD (2025), "AI Principles", available here.*
[23] *EU Artificial Intelligence Act (2024), "The Explorer", available here.*
[25] *United Nations (2022), "Principles for the Ethical Use of Artificial Intelligence in the United Nations System", available here.*

| | |
|---|---|
| **AI frameworks** | • OECD Framework for the classification of AI Systems<br>• NIST AI RMF<br>• NIST Special Publication 1270: Towards a Standard for Identifying and Managing Bias in AI<br>• Singapore AI Verify<br>• The Council of Europe's Human Rights, Democracy, and the Rule of Law Assurance Framework for AI systems |
| **Declarations and voluntary commitments** | • Bletchley Declaration<br>• The Biden-Harris Administration's voluntary commitments from leading AI companies<br>• Canada's guide on the use of generative AI |
| **Standards efforts** | • ISO/IEC JTC 1 SC 42<br>• The Institute of Electrical and Electronics Engineers Standards Association P7000<br>• The European Committee for Electrotechnical Standardization AI standards for EU AI Act<br>• The VDE Association's AI Quality and Testing Hub<br>• The British Standards Institution and Alan Turing Institute AI Standards Hub<br>• Canada's AI and Data Standards Collaborative |

Source: iapp & FTI Consulting (2024)

Ethical use of AI is paramount, ensuring fairness, transparency, and accountability.[25] AI systems should be trustworthy-by-design to avoid biases that could lead to discriminatory practices, particularly in lending, credit scoring, and insurance underwriting. AI systems should be designed and implemented in ways that promote fairness, transparency, and accountability.

Bias in AI algorithms can lead to discriminatory practices, particularly in areas such as lending, credit scoring, and insurance underwriting.[26] Regular audits and bias detection mechanisms are essential. Transparency in AI decision-making processes is also vital. Financial institutions should provide clear explanations of AI-driven decisions, enabling customers and regulators to understand and trust these systems. Disclosure requirements play a significant role in maintaining transparency, with institutions encouraged to inform customers when AI is used in decision-making processes that affect them.[27]

Box 2 considers key consumer protection and financial soundness implications related to the ethical use of AI.

## Box 2: Consumer protection and financial soundness key considerations

Consumer protection and financial soundness are key considerations. The FSCA Treating Customers Fairly (TCF) principles (FSCA, 2025) guide financial institutions to ensure fair treatment of customers through clear communication, product suitability, and effective claims handling. AI can enhance these principles by improving customer communication and personalising product offerings. However, AI introduces new risks, including model risk, operational risk, and cybersecurity threats. Comprehensive risk management frameworks should be developed to identify, assess, and mitigate these risks.

AI in finance introduces potential vulnerabilities that could impact financial soundness (FSB, 2024). These include dependencies on third-party providers, increased market correlations, cyber risks, and issues related to model risk, data quality, and governance. Additionally, GenAI raises concerns about financial fraud and disinformation. While current regulatory frameworks address many of these vulnerabilities, further enhancements in monitoring, policy frameworks, and supervisory capabilities may be necessary to mitigate risks effectively.

[25] *World Bank (2024), "Global Trends in AI Governance: Evolving Country Approaches", available here.*
[26] *University of Cambridge (n.d.), "Foundations: Trustworthy by Design", available here.*
[27] *World Bank (2024), "Global Trends in AI Governance: Evolving Country Approaches", available here.*

Transparency and explainability in AI decision-making processes are crucial. It is key to provide clear explanations of how AI-driven decisions are made, enabling customers and regulators to understand and trust these systems.[28] Furthermore, disclosure requirements play a vital role in maintaining transparency. Financial institutions are encouraged to disclose the use of AI in their operations, particularly in decision-making processes that affect customers.[29] This includes informing customers when AI is used to evaluate creditworthiness or determine insurance premiums, ensuring they are aware of and can understand the factors influencing these decisions.[30]

Moreover, financial institutions are encouraged to provide mechanisms for consumers to challenge decisions made by AI systems. However, there exists the trade-off between high performance models vs. explainable model as explainability only works for simpler models.

Human autonomy and oversight are important components of AI governance. Organisations are encouraged to ensure that AI systems respect human freedom and autonomy, guaranteeing human oversight.[31] Throughout the entire life cycle of an AI system, human-centric design practices should be followed, allowing for meaningful human decision-making. Human oversight means that humans must have the ability to manage the overall activity of an AI system, decide when and how to use it, and override decisions made by the system. Decisions involving life or death, or fundamental human rights must always involve human intervention and should not be left to AI systems. AI systems should be consumer-centric by design, prioritising the needs and interests of consumers at every stage.[32] Box 3 illustrates the application of these components of AI governance in the United Kingdom.

## Box 3: United Kingdom case report

In the UK, the application of human oversight in AI within the financial sector is guided by a robust legal and regulatory framework. The Financial Conduct Authority (FCA) and the Prudential Regulation Authority (PRA) play pivotal roles in ensuring that AI systems are deployed responsibly. These regulatory bodies mandate that financial institutions maintain accountability for AI-driven decisions, emphasising the importance of human oversight (human in the loop) to mitigate risks associated with automated processes (Bank of England, 2022).

The FCA's guidelines on AI and machine learning stress the necessity for transparency and explainability in AI models. Financial institutions are required to implement governance structures that include human review and intervention points, particularly in high-stakes decision-making scenarios such as credit scoring and fraud detection. This approach helps to ensure that AI systems operate within ethical boundaries and comply with existing regulations, thereby safeguarding consumer interests and maintaining trust in the financial system (Financial Conduct Authority, 2024).

Overall, the UK's regulatory framework underscores the critical role of human oversight in the responsible adoption of AI in finance, balancing innovation with the need for accountability and ethical considerations.

Data privacy is a paramount concern in the governance of AI in finance. AI systems rely on vast amounts of data to function effectively, and this data often includes sensitive personal and financial information.[33] It is key to implement stringent data protection measures and privacy-by-design principles to safeguard customer data from breaches and misuse.[34] This includes adopting advanced encryption techniques, ensuring secure data storage, and establishing clear data governance policies. Moreover, institutions might find it advantageous to be transparent with customers about how their data is used and provide them with control over their personal information. Compliance with data protection regulations, such as POPIA, is essential to maintaining customer trust and avoiding legal penalties.[35] Box 4 displays the EU case report for governing data privacy in AI.

[28] *Vinayak Pillai (2024), "Enhancing Transparency and Understanding in AI Decision-Making Processes", available here.*
[29] *Harvard Kennedy School (2023), "Disclosure Dilemmas: AI Transparency is No Quick Fix", available here.*
[30] *United Nations System (2022), "Principles for the ethical use of artificial intelligence in the United Nations system", available here.*
[31] *Inter-Parliamentary Union (2025), "Ethical principles: Human autonomy and oversight", available here.*
[32] *United Nations System (2022), "Principles for the ethical use of artificial intelligence in the United Nations system", available here.*
[33] *GDPR Local (2024), "The Future of Finance: Adapting to AI and Data Privacy Laws", available here.*
[34] *OneTrust (n.d.), "The 7 principles of privacy by design", available here.*
[35] *CFA Institute (2024), "Navigating the Risks of AI in Finance: Data Governance and Management Are Critical", available here..*

## Box 4: European Union case report

In the EU, the governance of AI in the financial sector, particularly concerning data privacy, is shaped by comprehensive legal and regulatory frameworks. The GDPR serves as the cornerstone of data privacy laws. Financial institutions deploying AI systems must adhere to GDPR requirements, which include obtaining explicit consent from individuals, implementing data protection by design and by default, and conducting Data Protection Impact Assessments (DPIAs) for high-risk AI applications (EY, 2024).

The EU Artificial Intelligence Act, which came into effect in August 2024, further complements GDPR by categorising AI systems based on their risk levels and imposing specific obligations on high-risk AI systems (OECD, 2024). This regulation mandates that financial institutions ensure human oversight, transparency, and accountability in their AI operations. The AI Act also requires regular monitoring and auditing of AI systems to ensure compliance with data privacy standards (BIS, 2024).

Overall, the EU's regulatory approach emphasises the importance of safeguarding personal data while fostering innovation in AI. By integrating GDPR and the AI Act, the EU aims to create a balanced framework that protects individuals' privacy rights and promotes the ethical use of AI in the financial sector.

Data governance is equally important in the context of AI in finance. It involves the management of data availability, usability, integrity, and security within an organisation. Effective data governance ensures that data used by AI systems is accurate, consistent, and reliable. This is crucial for making informed decisions and maintaining the trust of consumers. It is key to establish comprehensive data governance frameworks that include data quality management, data stewardship, and data lifecycle management. By ensuring that data is properly managed and protected, institutions can enhance the performance of their AI systems and reduce the risk of errors and biases. Additionally, strong data governance practices can help institutions comply with regulatory requirements and demonstrate accountability in their data handling processes.

Risk management is another vital governance consideration when deploying AI in finance. AI systems can introduce new risks, such as model risk, operational risk, and cybersecurity threats. Financial institutions could benefit from developing comprehensive risk management frameworks to identify, assess, and mitigate these risks. This includes conducting thorough testing and validation of AI models to ensure their accuracy and reliability. Additionally, institutions might consider establishing robust incident response plans to address potential AI-related failures or breaches. Cybersecurity measures are particularly important, as AI systems can be vulnerable to attacks that compromise their integrity and functionality.[36] Furthermore, establishing clear liability frameworks is essential to delineate responsibility in the event of AI system failures or malfunctions. These frameworks can help determine accountability and provide a basis for legal recourse, ensuring that both the institution and its customers are protected. The Japan case report in Box 5 illustrates how risk management in AI is governed.

## Box 5: Japan case report

In Japan, the application of AI governance in the financial sector, particularly concerning risk management, is guided by a comprehensive legal and regulatory framework. The Financial Services Agency (FSA) of Japan plays a crucial role in overseeing the deployment of AI technologies within financial institutions. The FSA has issued guidelines that emphasise the importance of robust risk management practices to mitigate potential risks associated with AI systems.

[36] *Cybersecurity measures are particularly important, as AI systems can be vulnerable to attacks that compromise their integrity and functionality. Examples of these measures include threat detection and prevention, where AI continuously monitors network traffic and user behaviour for anomalies, and automated incident response platforms, which streamline threat investigation and recovery processes. AI-driven identity and access management (IAM) systems ensure only authorised users access sensitive data, while AI-powered tools scan for vulnerabilities and apply patches proactively. Additionally, AI assists in proactive threat hunting, malware analysis, and reverse engineering to mitigate new threats. Penetration testing and ethical hacking simulations help identify system weaknesses, and AI-driven risk assessment tools evaluate an organisation's security posture to prioritise risks effectively. Implementing these measures enhances the protection of AI systems against evolving cyber threats.*

These guidelines require financial institutions to implement comprehensive risk management frameworks that include regular monitoring and evaluation of AI models. Institutions are also encouraged to conduct thorough risk assessments and ensure that AI systems are transparent and explainable. This approach helps to identify and address potential risks such as model bias, data privacy issues, and operational disruptions.

Overall, Japan's regulatory framework underscores the importance of effective risk management in the responsible adoption of AI in finance, ensuring that technological advancements are aligned with ethical standards and regulatory compliance. This balanced approach helps to safeguard consumer interests and maintain the stability of the financial system (Bank of Japan, 2024).

Accountability is a cornerstone of ethical AI use**.** Financial institutions may consider establishing clear internal processes that define:

- Who is responsible for AI system outcomes?
- How fairness, accuracy, and legality of AI decisions are verified.
- Mechanisms for consumer redress in cases of AI-related harm.

## 5.2 ETHICAL AND RESPONSIBLE AI CONSIDERATIONS

In the finance sector, the ethical use of AI is paramount, as financial institutions need to ensure that AI systems are designed and deployed to benefit consumers without causing harm. This involves several practical steps. Firstly, avoiding bias by training AI models on diverse datasets to prevent discrimination against any group is essential. Regularly auditing these datasets and the AI models can help identify and mitigate any biases that may arise. Secondly, ensuring fairness is crucial; decisions made by AI, such as pricing for insurance policies, should be fair and justifiable. This can be achieved by implementing fairness metrics and continuously monitoring AI decisions to ensure they align with ethical standards. Additionally, promoting inclusiveness is essential, as AI should be used to enhance financial inclusion, providing services to underserved communities.

Several ethical standards and guidelines have been established to guide the ethical use of AI. For instance, the European Union's guidelines on ethics in AI emphasise a human-centric approach, focusing on principles such as respect for human autonomy, prevention of harm, fairness, and explicability.[37] Similarly, the U.S. Department of Defence has adopted ethical principles for AI, which include responsibility, equitability, traceability, reliability, and governability. Additionally, the OECD AI Principles promote the use of AI that is innovative and trustworthy, respecting human rights and democratic values.[38] These principles emphasise inclusive growth, sustainable development, and well-being, ensuring that AI benefits all of society.

Trust in AI systems is crucial for their acceptance and effectiveness. Financial institutions need to build AI systems that are reliable, performing consistently and accurately. According to the OECD AI Principles, trustworthy AI should also be robust, secure, and safe. Robust AI systems can withstand various challenges and continue to operate effectively under different conditions, which reassures users of their dependability. Security measures, such as protecting AI systems from cyber threats and ensuring data integrity, further enhance user trust by safeguarding sensitive information and maintaining the confidentiality of user data. Transparency is another key aspect; the decision-making process of AI should be understandable to users. Furthermore, AI should treat all users equitably, without favouritism or prejudice, ensuring fairness in its operations. Box 6 discusses the level of trust that consumers have regarding South African financial institutions highlighting the importance of trust in AI systems in finance.

[37] *OECD (2025), "AI Principles", available here.*
[38] *OECD (2025), "AI Principles", available here.*

## Box 6: Level of Trust in South African Financial Institutions

The financial sector has been founded on trust, and in South Africa, this trust is generally high, with customers feeling confident that institutions will safeguard their money. However, this sentiment is conditional. The FSCA's *South African Financial Customer Behaviour and Sentiment Study* (published in 2023) reveals a deep-seated scepticism toward certain providers, such as those offering cash loans, due to unclear terms and a perception that institutions prioritise their own interests. The report highlights that trust is not just about security; it's about a feeling of fairness, transparency, and reciprocal benefit. This fragile trust is now being challenged and reshaped by the rapid adoption of artificial intelligence. Negative experiences with digital platforms can quickly erode this trust, making it a critical area for financial institutions to address. Access the report here: https://www.fsca.co.za/Documents/South%20African%20Retail%20Financial%20Customer%20Behaviour%20and%20Sentiment%20Report.pdf.

The World Economic Forum (WEF) Digital Trust framework offers a guide for building trust in this new digital landscape. The framework is built on three core pillars: security and reliability, accountability and oversight, and inclusive and ethical use. These principles are supported by eight key dimensions, including transparency, fairness, privacy, and redressability. For South African financial institutions, this framework provides a roadmap for integrating AI ethically. By focusing on transparency, institutions can make their AI-driven decisions, such as those in credit scoring, more understandable and fairer. The framework's emphasis on redressability directly addresses a key finding of the South African report—that customers face significant barriers when trying to lodge complaints. By adopting these principles, financial institutions can demonstrate a commitment to customer well-being in the AI era.

The *South African Financial Customer Behaviour and Sentiment Study* provides crucial evidence that validates the importance of the WEF framework. The report's findings on customer sentiment directly highlight the need for the framework's dimensions. Negative feedback about "billing irregularities", "channel downtime", and a perceived lack of fairness all reinforce the necessity of the WEF's principles of transparency and fairness. Similarly, the report's identification of significant barriers to lodging complaints and a general sense of powerlessness among some customer groups is a powerful argument for the framework's focus on accountability and redressability. Ultimately, the South African report serves as a powerful case report, demonstrating that for AI to be successfully and ethically integrated into the financial sector, a comprehensive and deliberate approach to digital trust is essential for meeting customer expectations and building lasting relationships.

Responsibility and accountability are key to maintaining trust in AI. Financial institutions need to consider assigning clear roles, defining who is responsible for the development, deployment, and monitoring of AI systems. Practical steps to place accountability include maintaining detailed records of AI design, training data, testing data, and decision-making processes. This documentation provides a clear audit trail and helps in tracing back any issues to their source.[39] Real-time monitoring of AI systems' performance and flagging anomalies as they arise is another crucial step. Implementing feedback mechanisms allows users to report issues or challenge AI decisions, ensuring continuous improvement. Establishing ethical review boards to oversee AI development and deployment can also help ensure that AI systems align with ethical standards. For example, a financial institution might set up a dedicated AI ethics committee responsible for reviewing and approving all AI-related projects, ensuring they meet ethical guidelines and regulatory requirements.

Data privacy and protection are critical in the finance sector, where sensitive information is handled. In South Africa, institutions must comply with the POPIA, which sets out strict guidelines for the collection, processing, and storage of personal data. Financial institutions must secure data by implementing robust security measures to protect it from breaches. According to the OECD, this includes encryption, access controls, and regular security audits. Ensuring compliance with POPIA involves not only adhering to its regulations but also fostering a culture of privacy within the organisation. Respecting user privacy by using data responsibly and only for the purposes for which it was collected is another important consideration. Institutions should also provide clear communication to users about their data rights and how their information is being used.

[39] *OECD (2023), "How countries are implementing the OECD Principles for Trustworthy AI", available here.*

Transparency and disclosure are essential for building trust in AI systems. Financial institutions are encouraged to communicate clearly with consumers about how their data is being used and how AI decisions are made. Providing understandable explanations for AI-driven decisions helps build trust and confidence. Additionally, being open about the limitations and potential risks of AI systems is important for maintaining transparency and trust. One significant challenge in this area is the “black box” problem, where the internal workings of AI systems, particularly those using deep learning, are not easily understood even by their creators. This opacity can hinder trust and accountability, making it crucial for institutions to invest in explainable AI tools and methods that can demystify these complex systems. Explainable AI (XAI) helps users understand how AI models make decisions, which can improve trust and allow for better oversight. By providing clear, understandable explanations, XAI can help ensure that AI systems are transparent and accountable. This allows for a clearer view of associated risks and informed decision-making to minimise risks for the financial institutions.

By focusing on these considerations, the finance sector can harness the power of AI responsibly and ethically, ensuring that it benefits all stakeholders while maintaining trust and integrity.

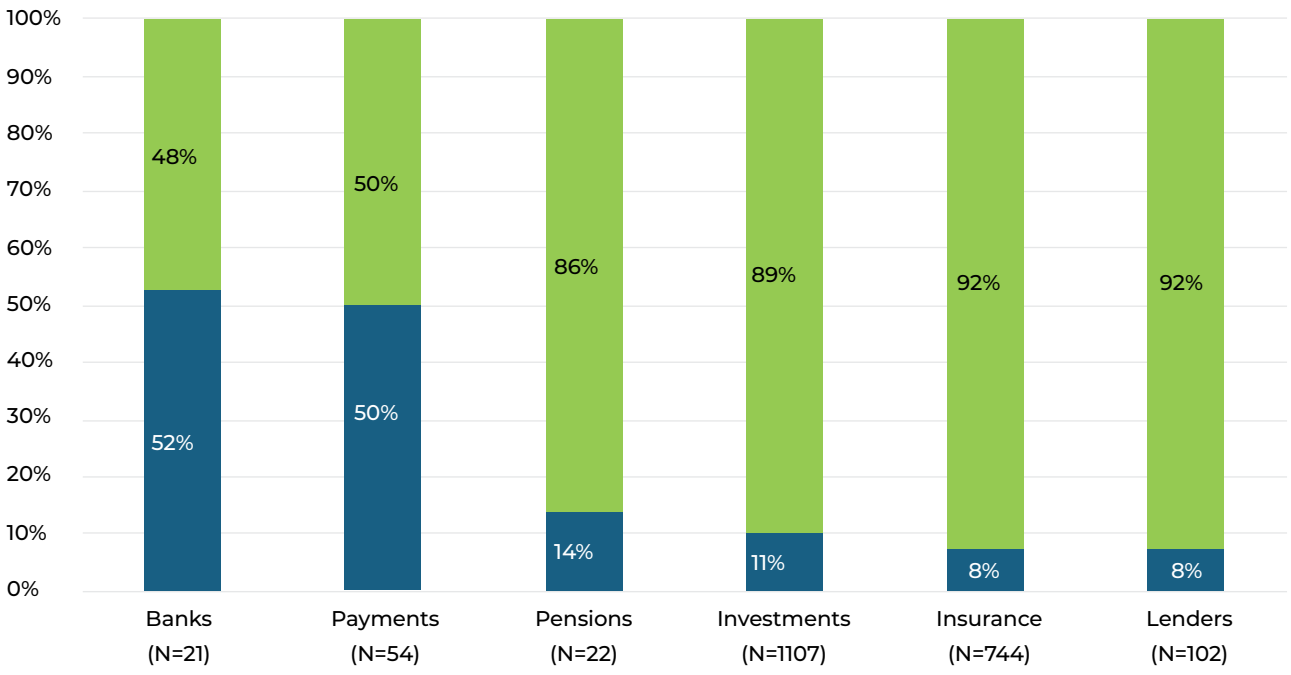
# PART 6: FINDINGS FROM THE INDUSTRY SURVEY

This section summarises the results of the 2024 industry survey on AI adoption in South Africa's financial sector. It examines current and planned use of traditional AI, rule-based and machine learning systems for prediction, automation, and decision support, and GenAIZ, which creates new content such as text or code using advanced models like large language models. The findings highlight adoption trends, investment intentions, key use cases, and the benefits, risks, and constraints shaping the sector's approach to AI.

## 6.1 AI ADOPTION AND INTENDED INVESTMENT

Figure 6 illustrates that banking institutions lead in AI adoption with approximately 52% adoption, this is followed by payment providers which also show significant adoption at 50%. Pensions follow thirdly in AI adoption at 14%. In contrast, the insurance (8%) and lending (8%) sectors have shown the lowest levels of AI integration indicating a cautious approach.

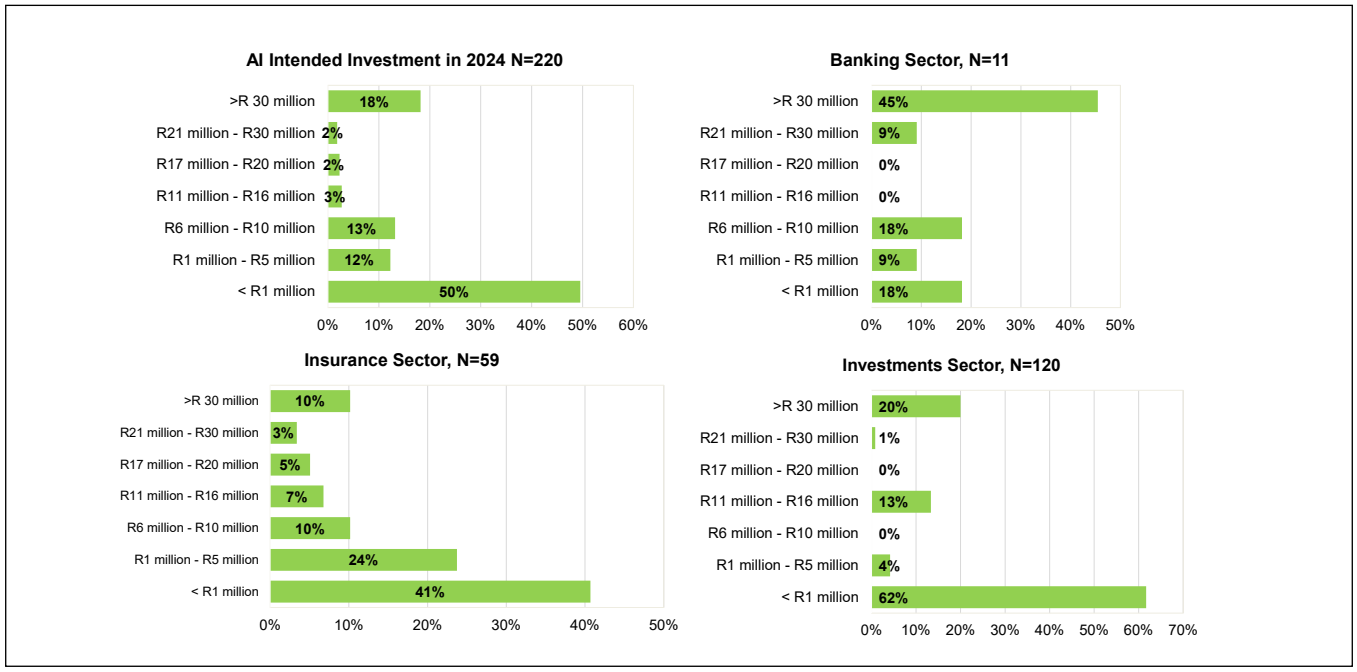
Figure 6: AI adoption by financial segment



As illustrated in Figure 7, planned AI investments vary significantly across SA financial sectors for institutions that have adopted traditional AI and/or GenAI. Banking institutions in SA demonstrate a strong inclination towards substantial AI investment. Of the 21 banks that have adopted AI, 45% planned to invest more than R30 million over 2024. This is the most significant overall investment per sector, indicating a strong focus on leveraging AI for technological advancements. In contrast, of the relatively few insurers and investment providers that have already adopted AI, planned investment for 2024 was also comparatively lower – about 62% of investment providers and 41% of insurers would spend less than R1 million.

When read against the primary use cases reflected below, significant investment in AI by banks may highlight a commitment to leveraging technology for operational efficiency, cyber security, sales and marketing, and productivity gains. It may necessitate strong oversight to manage the associated risks. Lower investments in the insurance and investment sectors may limit their ability to innovate over the longer term.

Figure 7: Intended AI investment by sector



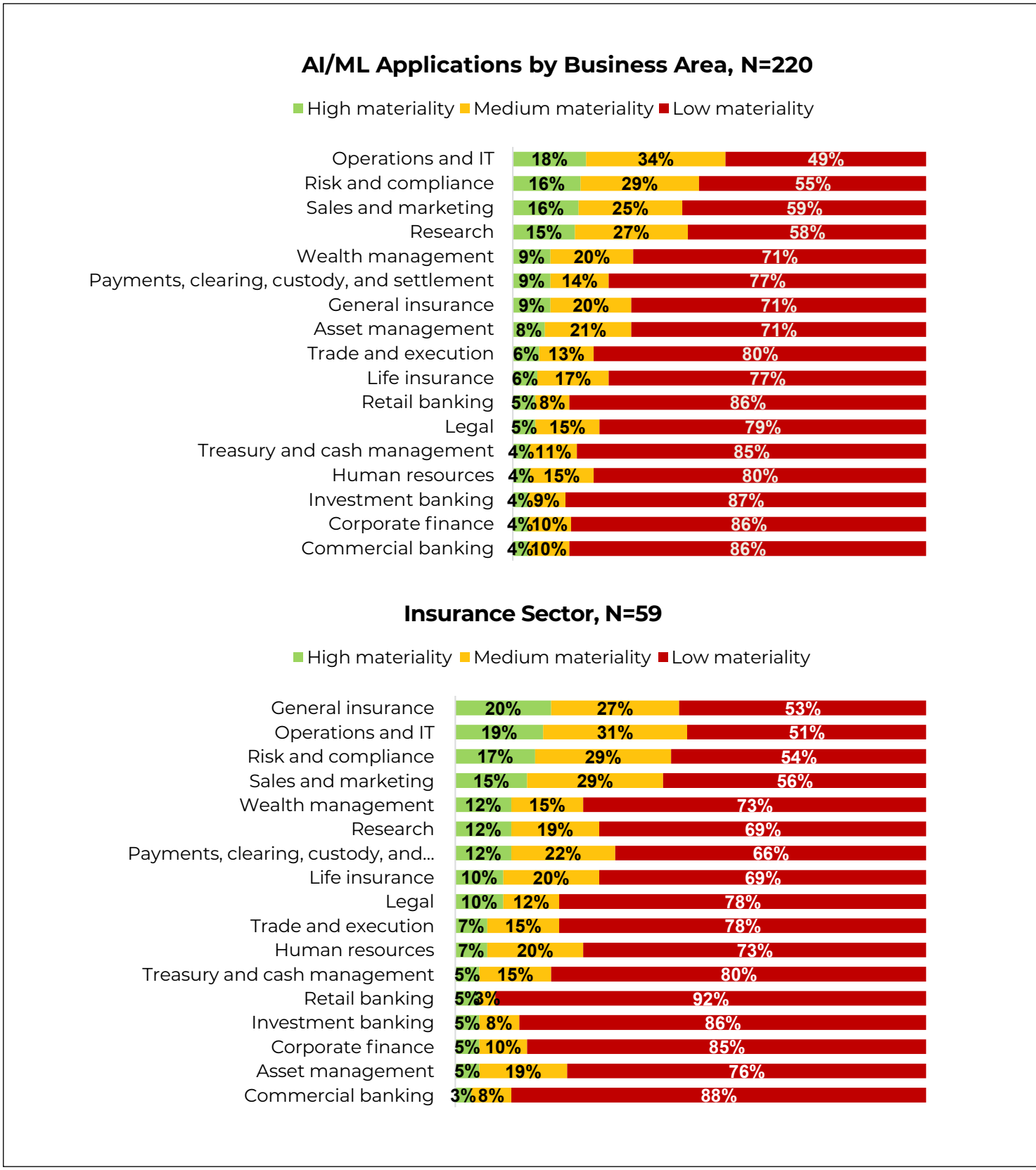

## 6.2 TRADITIONAL AI CURRENT AND PLANNED USE CASES

An emerging trend to monitor is the extent that AI is being employed to the advantage of the provider rather than to the advantage of the customer. While there may nonetheless be associated benefits for the consumer, for example enhanced cybersecurity can reduce risk of fraud and scams, and operational efficiencies may support competition and bring down prices, optimising these benefits will likely require planning by the financial institution for targeted consumer outcomes, underpinned by strong strategic leadership.

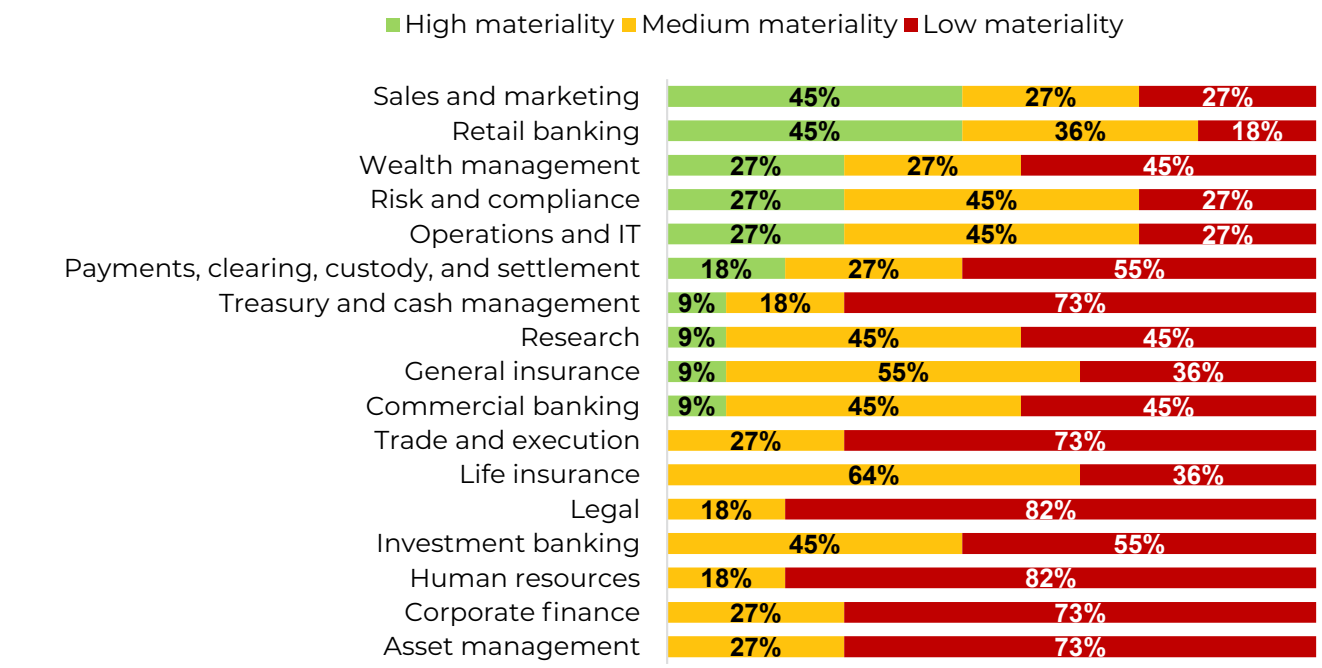
Figure 8 shows that traditional AI applications are mostly used in operations and IT (18%). Risk and compliance, along with sales and marketing, follow closely at 16%, leveraging these technologies to enhance operational efficiency and reduce costs.

Within the banking sector, the report found that roughly 45% of the banking institutions use traditional AI technology in their sales and marketing and retail banking business areas, while roughly two out of ten insurers use traditional AI applications in their general insurance and operations and IT business areas. The investment sector utilises traditional AI technology mostly within the research, risk and compliance, and operations and IT business areas of the operations.

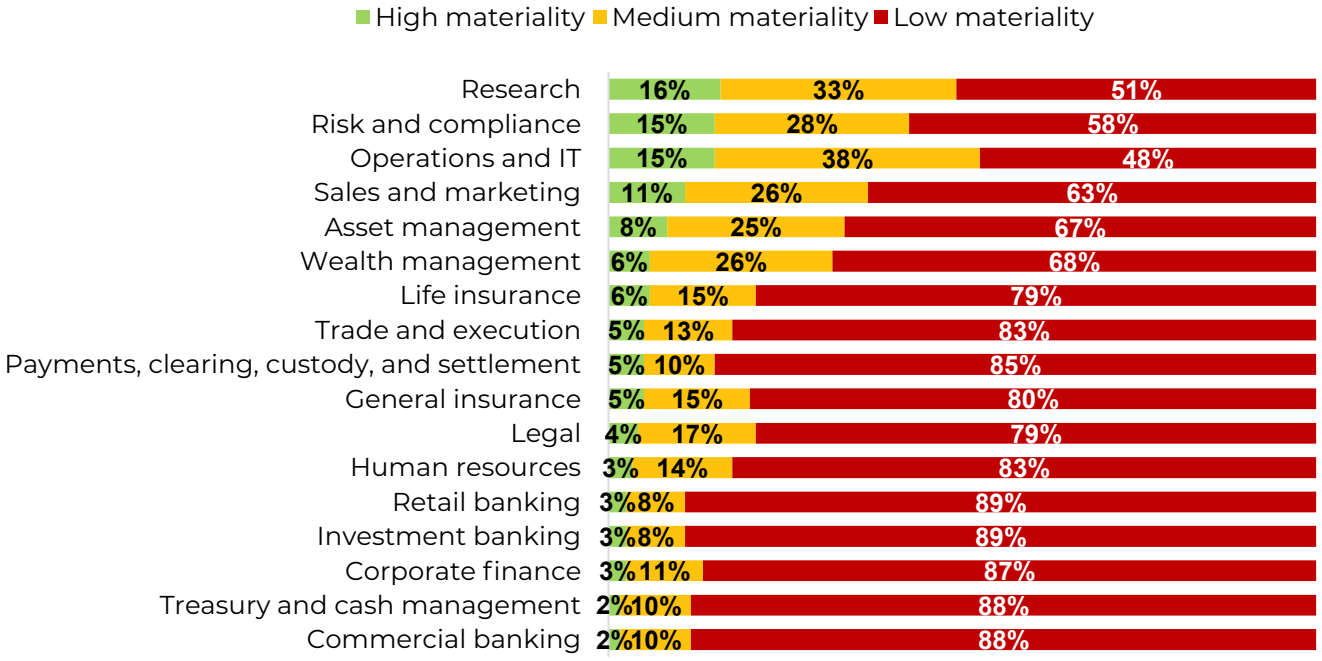
Figure 8: Traditional AI application by business area and materiality



## Banking Sector, N=11




## Investments Sector, N=120

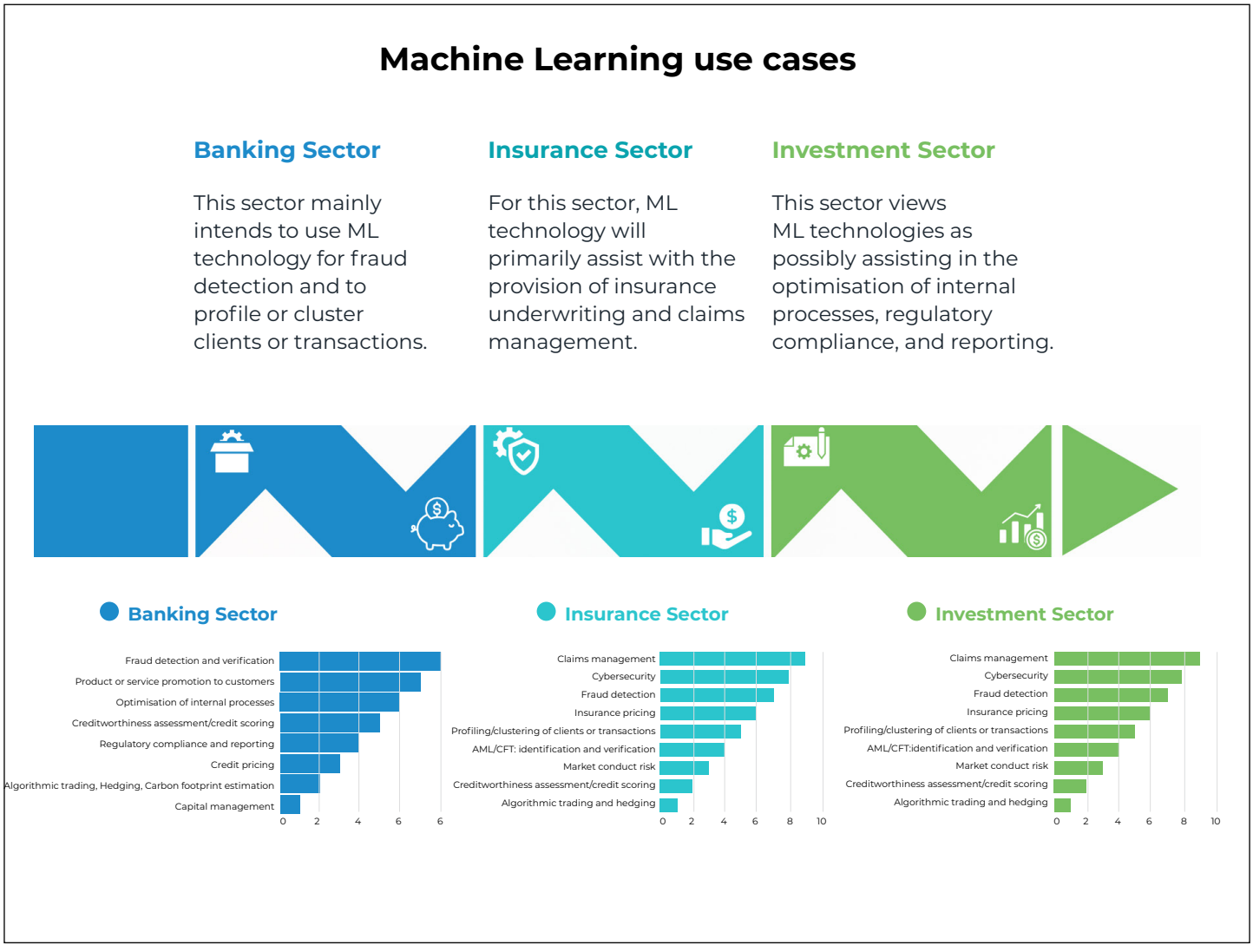


When assessing the application of traditional AI technology across their value chains, banking institutions primarily utilise it for chatbot services, productivity enhancement, fraud detection, and product development. The insurance sector leverages AI/ML predominantly for fraud detection, claims assessment, and security improvement. In contrast, the investment industry employs these technologies to optimise processes, manage risks, and facilitate research.

From a conduct perspective, the focus on operations and IT suggests a drive towards improving efficiency and reducing costs. The use of AI in risk and compliance underscores the importance of maintaining regulatory standards and managing risks effectively. From a prudential perspective, AI applications in risk and compliance can enhance the ability to detect and mitigate risks, contributing to financial stability. However, the reliance on AI for critical functions requires robust governance and continuous monitoring to prevent systemic risks.

Overall, the most prevalent planned use of ML applications is to optimise internal processes, to bolster cybersecurity and to provide improved customer support. The banking sector mainly intends to use ML technology for fraud detection and to profile or cluster clients or transactions, as well as identifying and verifying transactions to combat AML/CFT activities. For the insurance sector, ML technology will primarily assist in the provision of insurance underwriting and claims management, with the intention to optimise internal processes. The investment sector views ML technology as possibly an assistant in the optimisation of internal processes, regulatory compliance and reporting, and to provide customer support through chatbots (see Figure 9).

Figure 9: Traditional AI use cases across the value chain



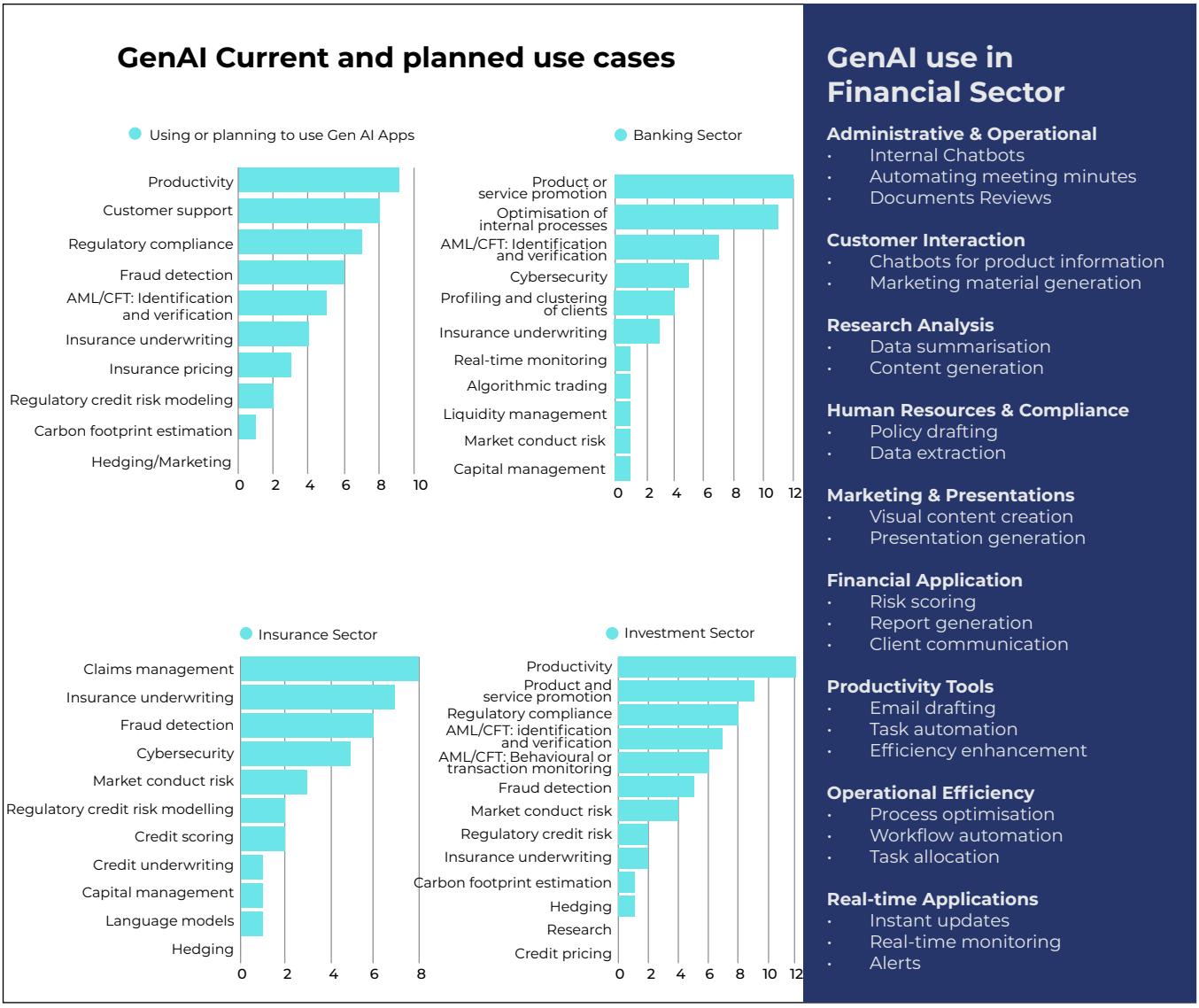
## 6.3 GenAI CURRENT AND PLANNED USE CASES

The respondents highlighted these various use cases of GenAI across the value chain of a financial institution.[40] In administrative and operational areas, GenAI is used for internal chatbots, automating meeting minutes, and document reviews. For customer interaction, chatbots inform members about products, and GenAI generates marketing material; this is illustrated in Figure 10.

Research and analysis benefit from data summarisation and content generation. Human resources and compliance use GenAI for policy drafting and data extraction. Marketing and presentations leverage AI for creating visuals and presentations. Financial applications include risk scoring, report generation, and client communication. Productivity tools like Microsoft Copilot and ZebraGPT enhance efficiency by drafting emails and reducing manual tasks. Experimentation and proofs of concept involve generative AI for various applications.

Operational efficiency is improved with AI tools. Data extraction and analysis are more accurate with AI compared to traditional OCR tools. Marketing and education use AI for personalised advisory services and generating collateral. Security and fraud detection involve real-time analysis of transaction patterns. Document compilation, customer experience enhancement, and scaling efficiency are other key areas where GenAI is applied.

Figure 10: GenAI use cases in the value chain

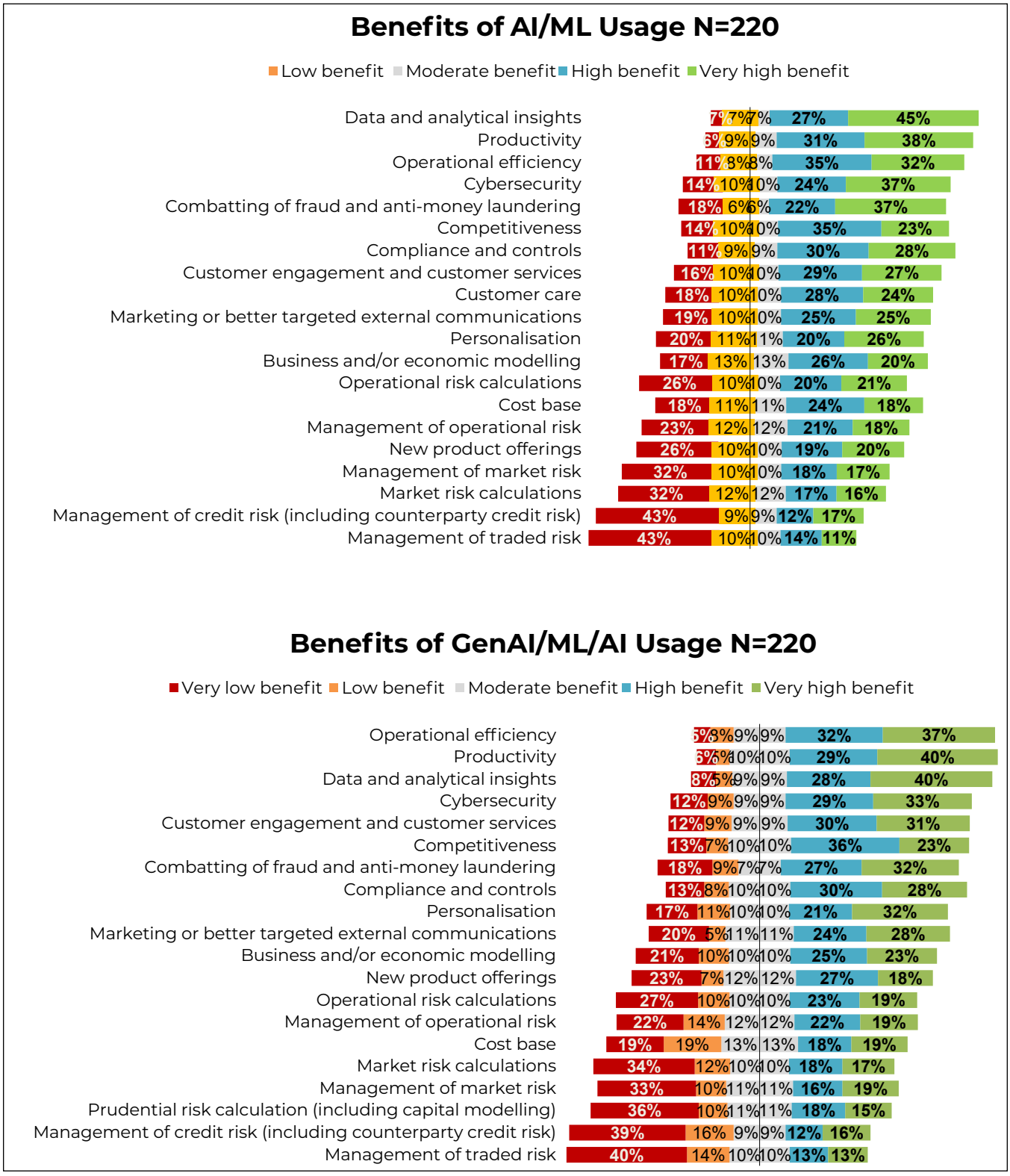



[40] *It is important to recognise that the interpretation of AI-related concepts, use cases, and terminology by market participants may vary significantly. This variation has implications for how institutions report on, implement, and govern AI systems. As such, further work may be required to better understand how these aspects are being interpreted across the financial sector, and to ensure that regulatory and supervisory responses are appropriately calibrated to reflect actual market practices and risks.*

## 6.4 TRADITIONAL AI/GENAI BENEFITS AND RISKS

Figure 11 shows that data and analytical insights; improved operational efficiency and productivity; improved cybersecurity and combating fraud and money laundering are the most impactful benefits introduced by traditional AI applications within the financial sector.

Figure 11: Benefits of traditional AI and GenAI

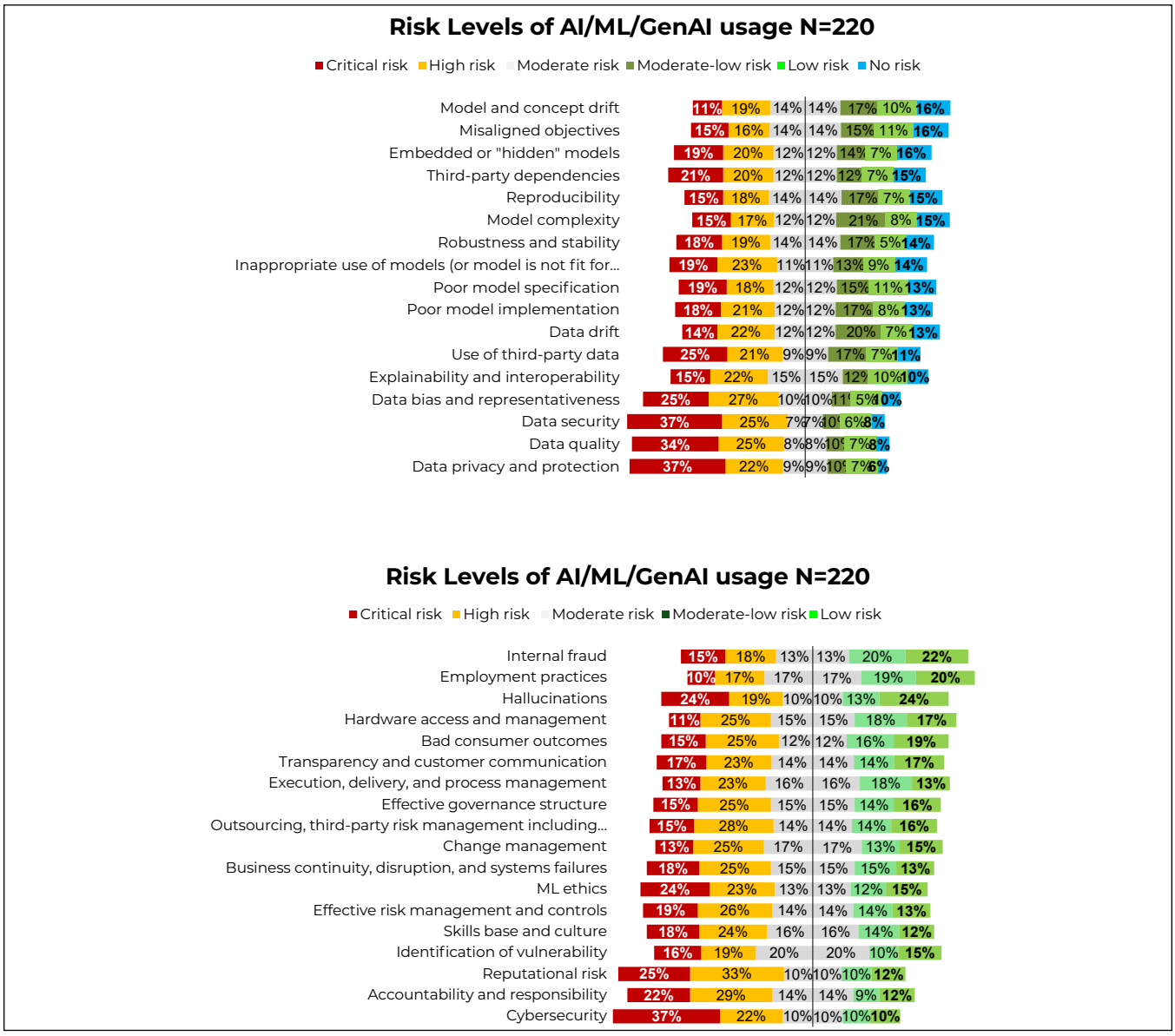


Within the specific financial segments, the research shows that for the banking sector, the highest perceived benefits of traditional AI are combating fraud and AML, while GenAI benefits customer engagement and services; within the insurance and investment segments the highest benefit is that traditional AI and GenAI enable more accurate data and analytical insights, while productivity is the most considerable benefit of traditional AI/GenAI.

From a conduct perspective, the ability of AI to provide data-driven insights and enhance productivity can lead to better customer service and more personalised financial products. Improved cybersecurity measures can protect consumers from fraud and data breaches. From a prudential perspective, enhanced operational efficiency and fraud detection capabilities contribute to the overall stability and resilience of financial institutions. However, the reliance on AI for these functions necessitates strong oversight to ensure accuracy and reliability.

Figure 12 underscores the most pressing traditional AI/GenAI risks confronting the financial sector, foremost among these are cybersecurity (37%) and reputational risk (25%), followed by the critical concerns of hallucinations and ML ethics, both cited by 24% of respondents. However, the banking and investment sectors are mostly concerned with data security; data privacy and protection and data bias and representativeness, while for the insurance sector cybersecurity is a major concern along with reputational risk and accountability and responsibility surrounding AI/ML applications.

Figure 12: Risks of traditional AI and GenAI



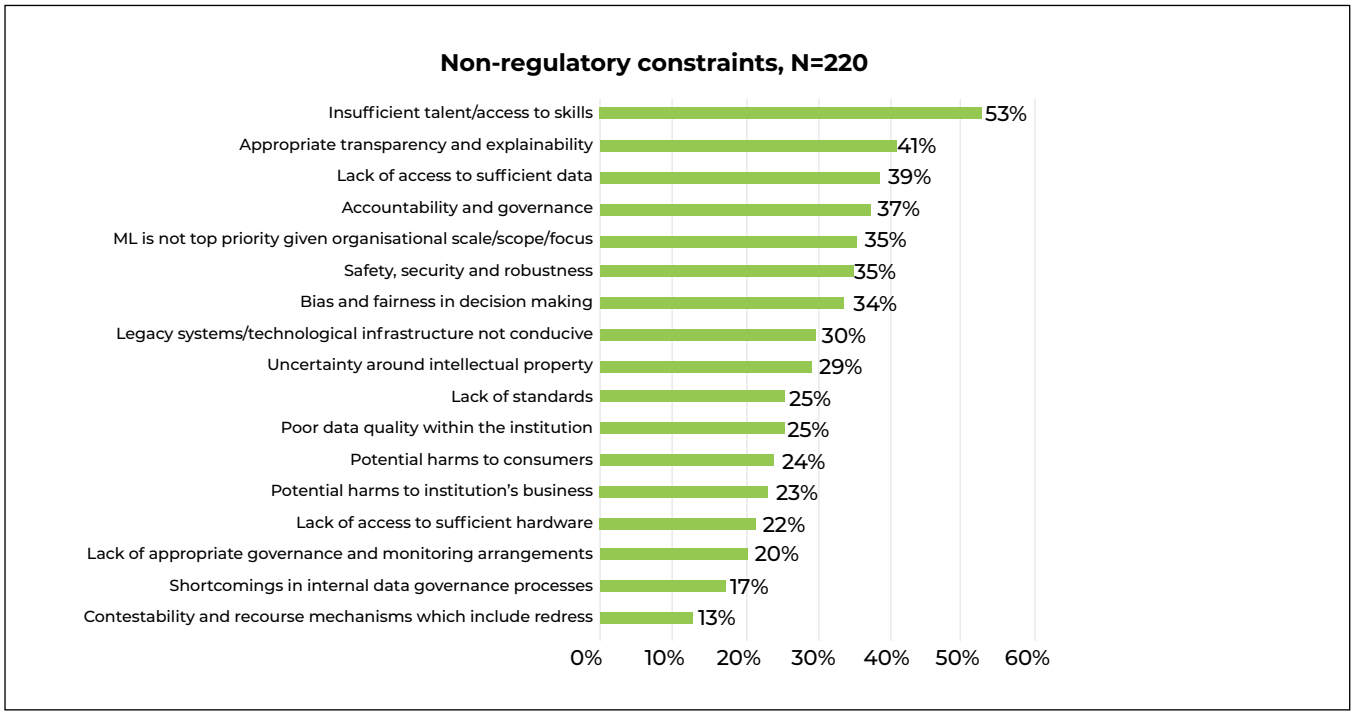
Cybersecurity risks highlight the need for robust security measures to protect consumer data and maintain trust. Reputational risks underscore the importance of ethical AI practices and transparency in AI decision-making. Financial institutions need to consider implementing a comprehensive risk management framework to mitigate these risks effectively.

## 6.5 CONSTRAINTS TO THE ADOPTION OF TRADITIONAL AI AND GENAI

The use of AI in finance introduces several important challenges that institutions must address to ensure responsible and effective implementation. Some non-regulatory and regulatory challenges include data privacy and protection laws, bias and fairness in AI algorithms, transparency and explainability of AI systems.

The effective implementation of AI requires a workforce with the necessary skills to leverage its potential. However, there is a notable shortage of professionals with expertise in AI technologies, machine learning, and data analytics as illustrated in Figure 13. Addressing this challenge requires strategic investments in comprehensive upskilling and reskilling programs. Collaborations with reputable academic institutions are essential to bridge the talent gap. Promoting a culture of continuous learning and innovation is crucial to overcoming this obstacle. Furthermore, the inappropriate transparency and explainability, alongside a lack of access to sufficient data, are also recognised as substantial contributors to the modest adoption rates of traditional AI technologies in South Africa.

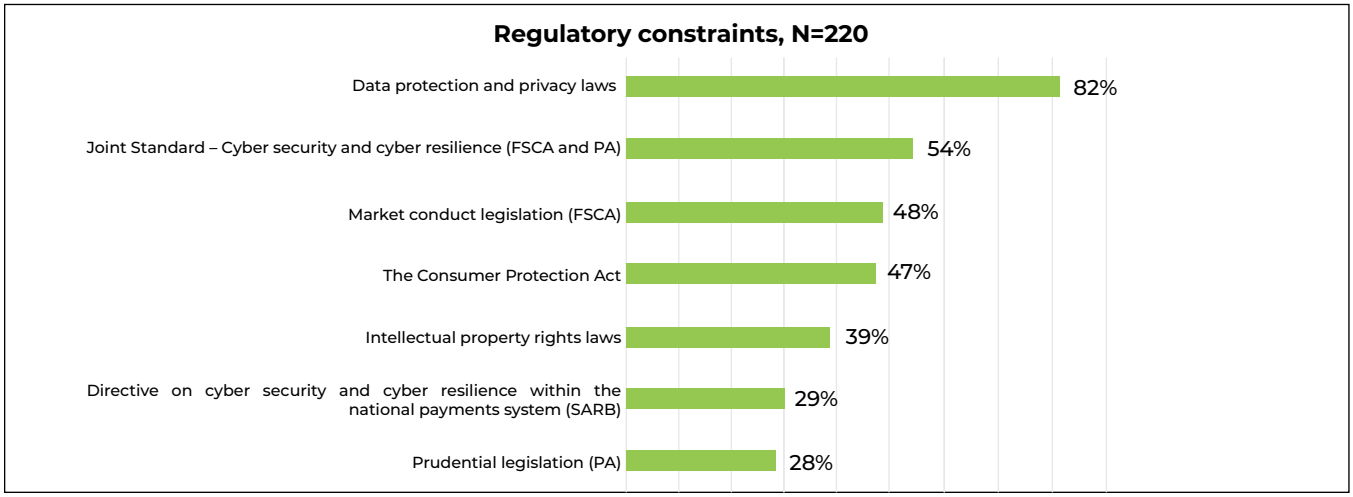
Figure 13: Constraints to AI usage




From a conduct perspective, compliance with data protection laws is essential for safeguarding consumer data and maintaining trust. Addressing talent shortages and improving transparency can enhance the ethical use of AI and ensure fair treatment of consumers. From a prudential perspective, regulatory constraints highlight the need for clear guidelines and support from regulatory bodies to facilitate AI adoption while ensuring compliance. Developing talent and improving explainability can help institutions manage risks and maintain financial stability.

There is general consensus across the surveyed financial sector participants that data protection and privacy laws are the most significant regulatory challenge to AI adoption as illustrated in Figure 14. However, this is not so much a constraint as a necessary requirement to ensure that consumers' data is effectively protected in the use of AI.

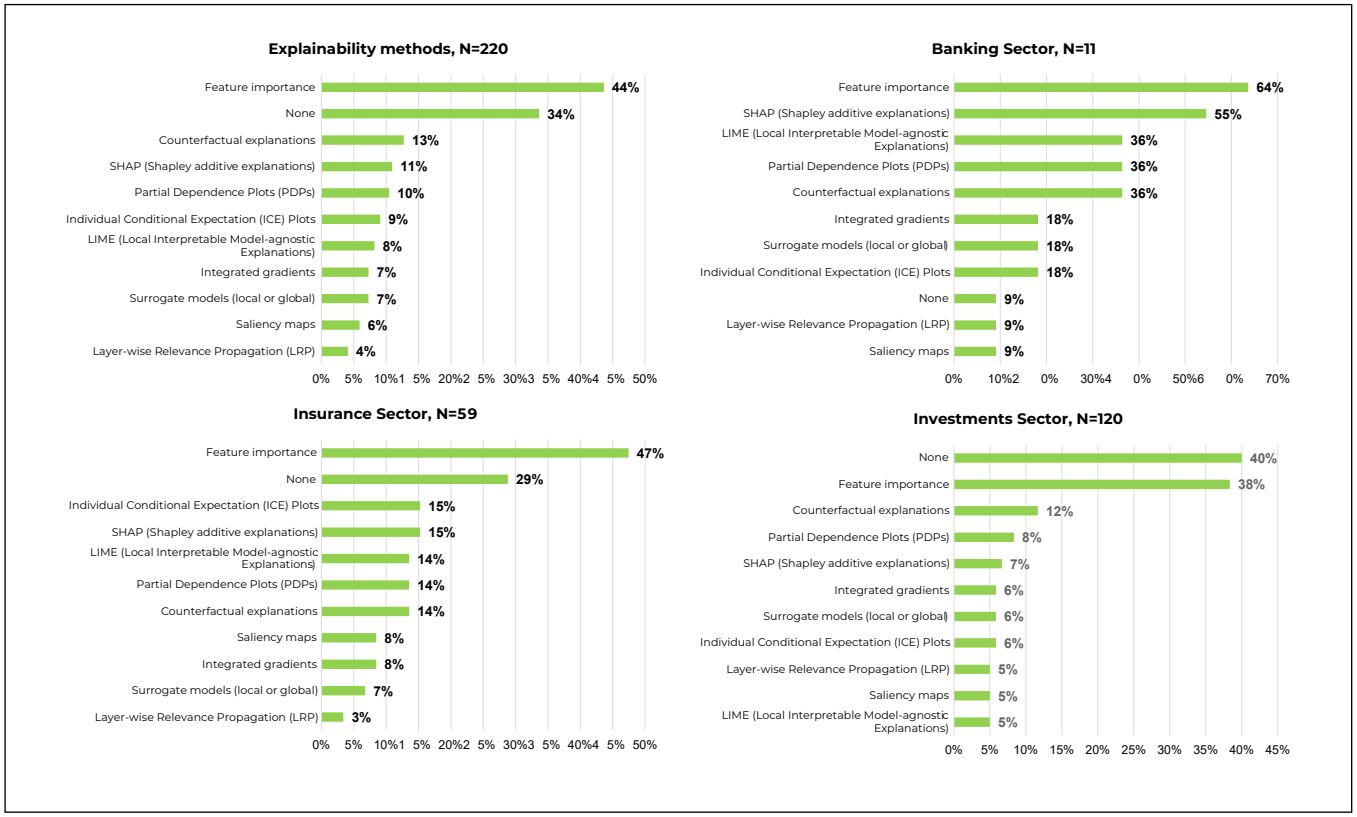
Figure 14: Regulatory Challenges




## 6.6 EXPLAINABILITY METHODS AND ETHICAL CONSIDERATIONS

Feature importance[41] is the leading explainability method employed as illustrated in Figure 15, thereby enhancing transparency in AI models. However, 21% of financial institutions do not use any form of explainability method, which raises concern around the governance of the use of traditional AI technologies in finance. The banking sector has more explainable models in the usage of traditional AI as compared to the insurance and investment sectors. Explainability is key for financial disclosure for consumers.

Figure 15: Employed exaplainability methods




[41] *Feature importance is a technique used in machine learning to identify which features (variables) have the most significant impact on predicting the target variable. It helps in understanding the model's decision-making process, improving model performance, and gaining insights into the data. Methods to determine feature importance include model-based approaches, permutation importance, SHapley Additive exPlanations (SHAP) values, and Local Interpretable Model-agnostic Explanations (LIME).*

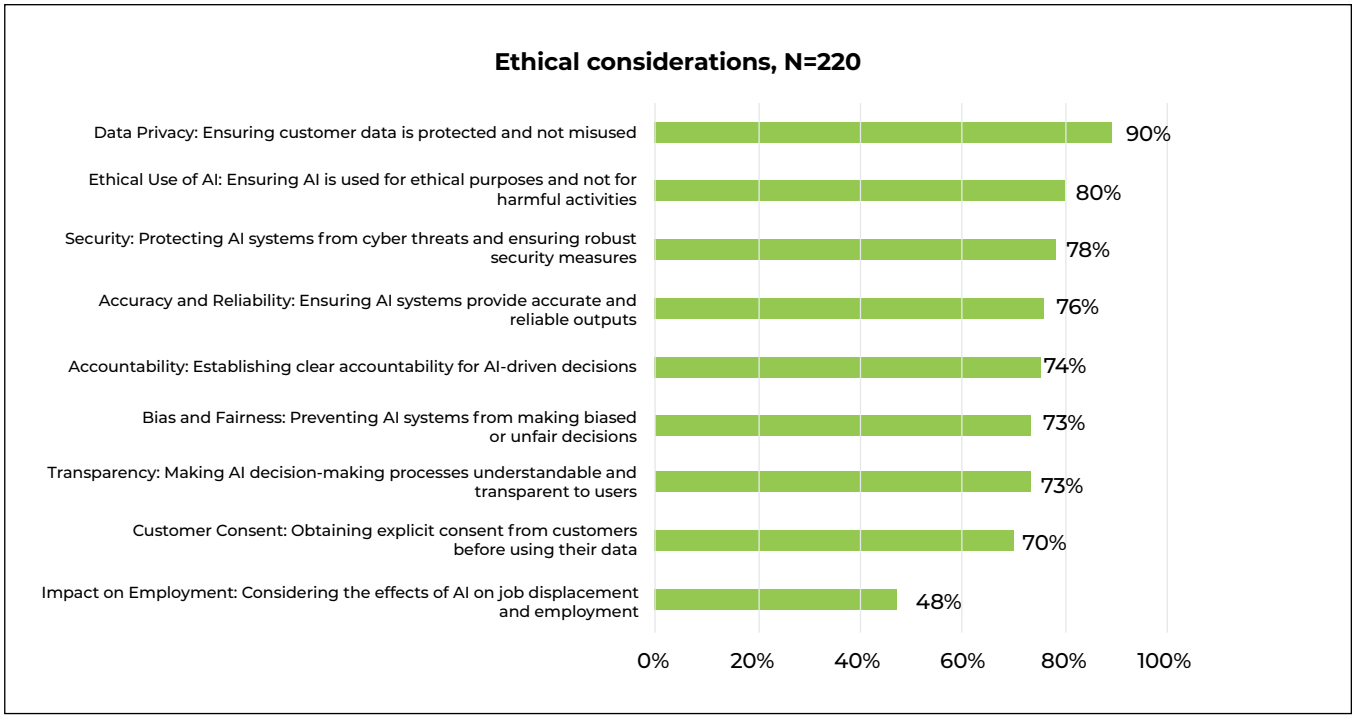
*SHAP (SHapley Additive exPlanations) values are a method used to explain the output of machine learning models by attributing the contribution of each feature to the prediction. They are based on cooperative game theory and provide a unified measure of feature importance.*

*LIME (Local Interpretable Model-agnostic Explanations) is a technique used to explain the predictions of any machine learning model by approximating it locally with an interpretable model. It works by perturbing the input data and observing the changes in the predictions to understand the model's behaviour in the vicinity of a specific instance.*

Ethical considerations are paramount in the deployment of AI in the financial sector (see Figure 16). The leading ethical considerations include data privacy, bias and fairness, transparency, accountability, and security. Ensuring customer data is protected and not misused is crucial. Financial institutions may consider implementing robust data protection measures, obtain user consent, and maintain transparency about data usage. Preventing AI systems from making biased or unfair decisions is essential. Institutions should train AI models on diverse datasets, regularly audit these models, and implement fairness metrics to ensure ethical decision-making.

Transparency in AI decision-making processes is vital. When financial institutions provide clear explanations of AI-driven decisions this enables customers to understand and trust these systems. Establishing clear accountability for AI-driven decisions is crucial. Institutions should assign responsibility for the development, deployment, and monitoring of AI systems, ensuring compliance with ethical standards and regulatory requirements. Protecting AI systems from cyber threats and ensuring robust security measures are essential. Institutions need to implement advanced encryption techniques, secure data storage, and establish clear data governance policies. By focusing on these ethical considerations, the finance sector can harness the power of AI responsibly and ethically, ensuring that it benefits all stakeholders while maintaining trust and integrity.

Figure 16: AI ethical considerations



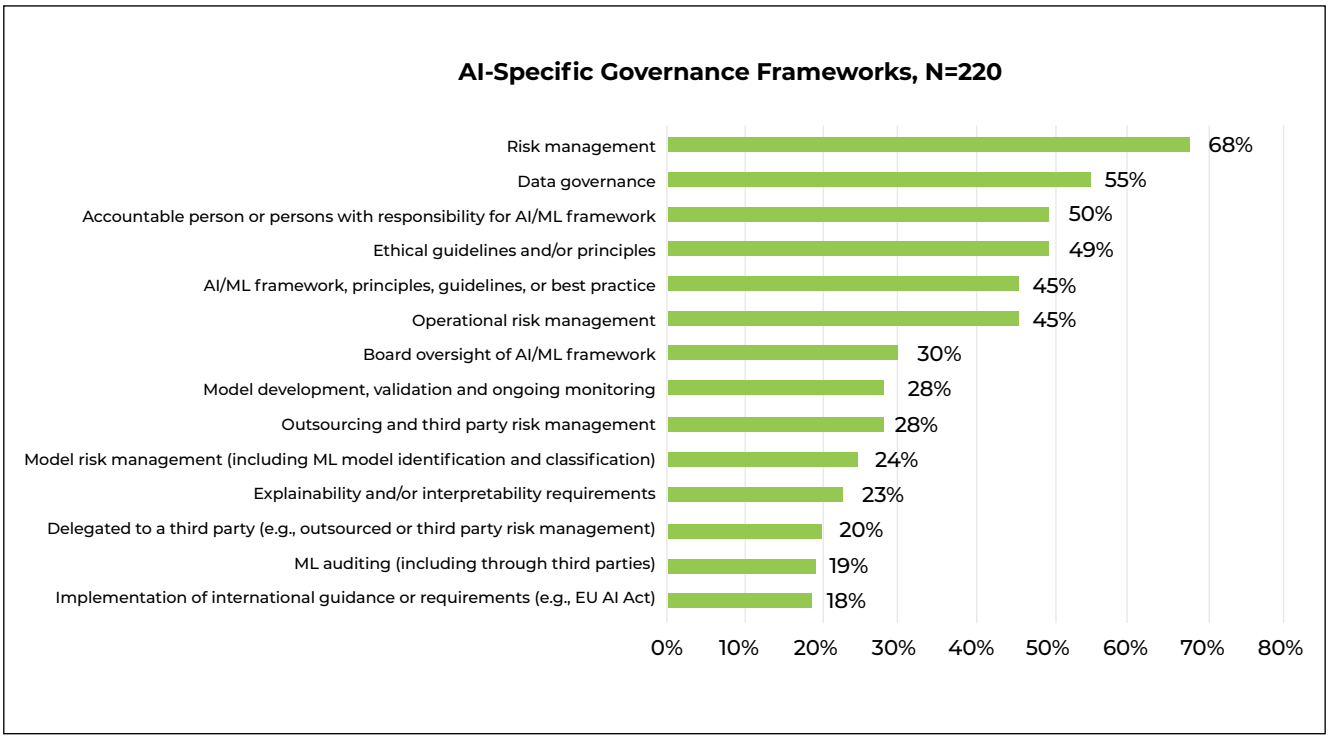
# 6.7 GOVERNANCE

## 6.7.1 GOVERNANCE FRAMEWORKS

The survey data indicates that South African financial institutions have implemented various governance frameworks to manage the deployment and operation of AI technologies (see Figure 17). These frameworks are essential for ensuring that AI systems are used responsibly, transparently, and ethically. Figure 17 highlights the key governance considerations from the survey. Governance frameworks tend to focus on data governance, model development, validation, ongoing monitoring, and risk management and these are illustrated in Figure 18.

Figure 17: Key governance frameworks applicable to AI

Figure 18: Governance frameworks applied to AI




From a conduct perspective: Governance frameworks ensure that AI systems operate transparently and ethically, fostering consumer trust and protecting their interests. Effective data governance and ethical guidelines help in preventing biases and ensuring fair treatment of customers. Regular monitoring and validation of AI models enhance the reliability of AI-driven decisions, improving customer satisfaction and loyalty.

From a prudential perspective: Robust governance frameworks are essential for managing the risks associated with AI systems. Effective risk management and cybersecurity measures protect financial institutions from potential threats and ensure the stability of the financial system. Board oversight and accountability frameworks ensure that AI initiatives align with regulatory requirements and strategic goals, promoting financial stability and resilience.

## 6.7.2 TRANSPARENCY AND ACCOUNTABILITY

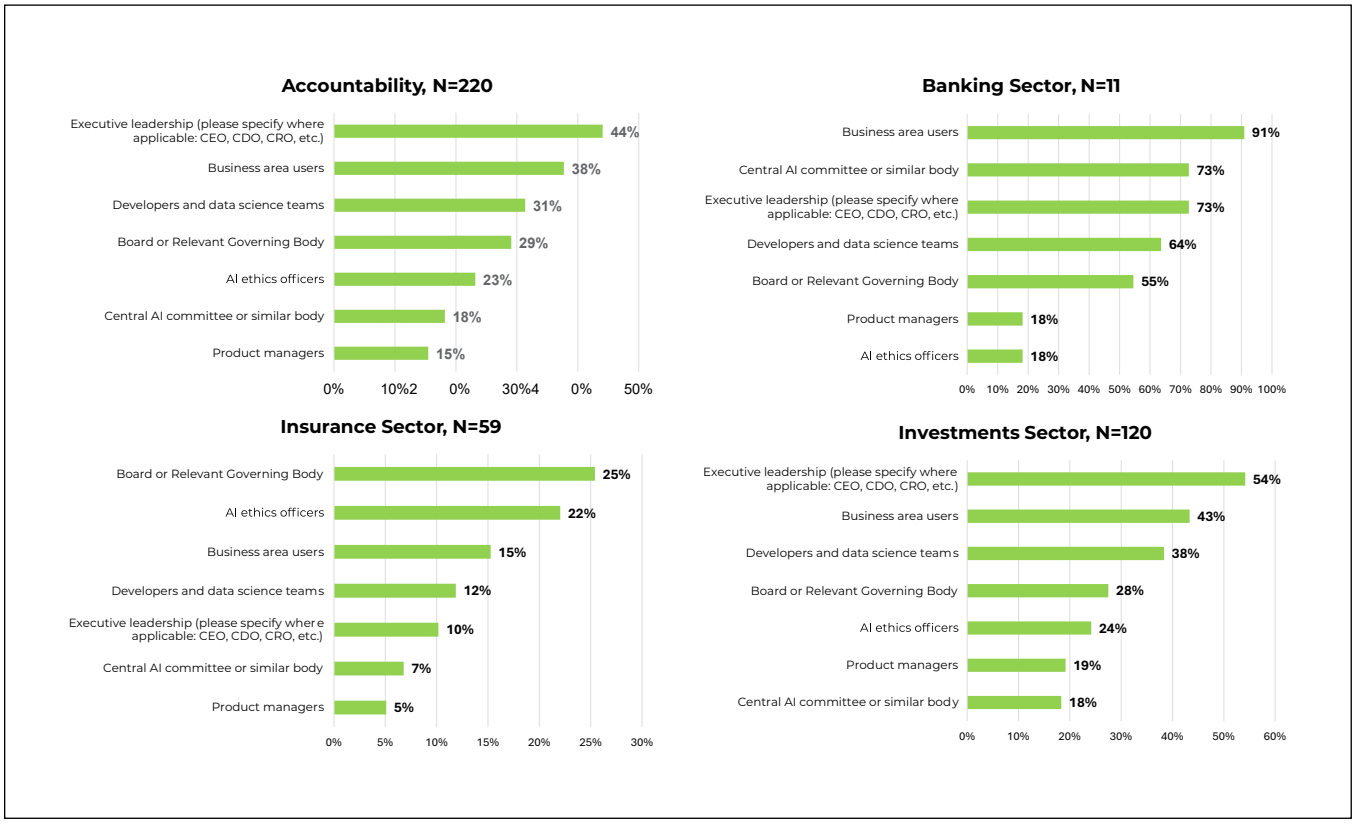
Figure 19 indicates various roles and structures responsible for the accountability of AI, ML, and GenAI within South African financial institutions. These roles include executive leadership, AI ethics officers, product managers, central AI committees, developers and data science teams, business area users, and boards or relevant governing bodies.

From a conduct perspective, having clear accountability structures ensures that AI systems operate ethically and transparently. Roles such as AI ethics officers and central AI committees help prevent biases and ensure fair treatment of customers. Executive leadership and boards provide strategic oversight, fostering consumer trust and protecting their interests.

From a prudential perspective, robust accountability frameworks are essential for managing the risks associated with AI systems. Executive leadership and central AI committees ensure that AI initiatives align with regulatory requirements and strategic goals. Developers and data science teams address technical issues, ensuring the stability and reliability of AI systems. Effective governance structures help mitigate risks and maintain financial stability.

By establishing clear accountability for traditional AI/GenAI, South African financial institutions can leverage these technologies responsibly and ethically, ensuring that they benefit all stakeholders while maintaining trust and integrity in the financial sector.

Figure 19: Accountability for Traditional AI/GenAI decisions




# 6.8 OTHER SURVEY INSIGHTS

The survey further explored some qualitative questions to industry, and these are discussed below:

## 6.8.1 PRINCIPLE-BASED VS RULES-BASED REGULATION

**Question: How can regulation be simplified, strengthened and/or extended to better encompass AI and address potential risks and harms?**

From the survey, principle-based regulation should be preferred over rules-based regulation, building on existing regulations and governance principles. Human oversight is imperative, and education and training methods are essential to onboard AI effectively. Implementing a risk-based approach where AI systems are categorised based on their potential impact is crucial, with stricter regulations for high-risk applications. Transparency in AI operations and decision-making processes may be mandated, requiring companies to disclose how their AI systems work, the data they use, and potential biases.

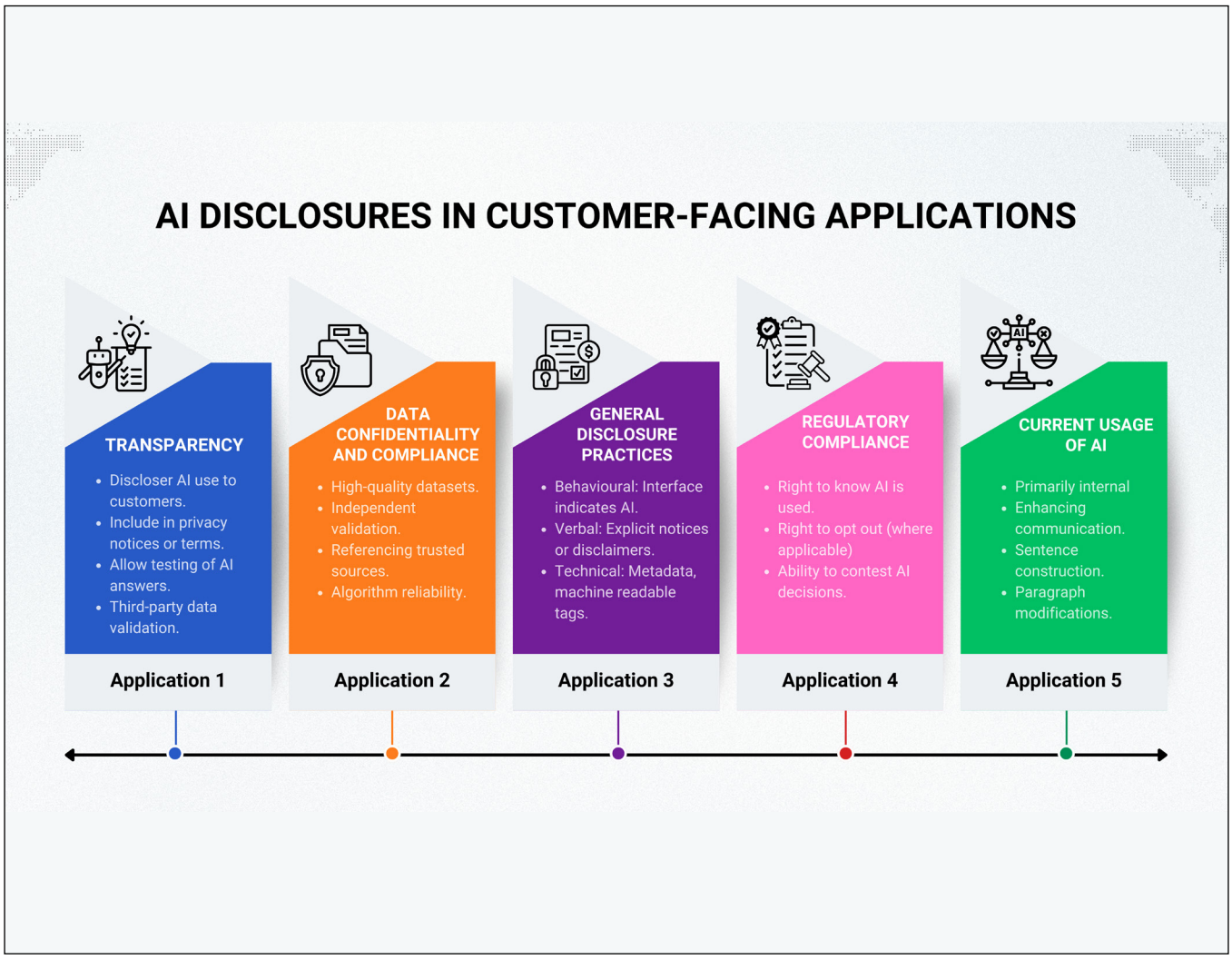
Establishing clear ethical guidelines and best practices for AI development and deployment is necessary, focusing on fairness, accountability, and non-discrimination. Ongoing collaboration with industry players is vital. Standardising global frameworks for accountability and aligning with international regulators is essential for effective AI governance.

### 6.8.2 TRANSPARENCY AND DISCLOSURE

**Question: What types of disclosure are there for customer-facing AI applications?**

There are different types of disclosures necessary for customer-facing AI applications, emphasising transparency, data quality, confidentiality, and compliance. The survey responses to this question are illustrated in Figure 20.

Figure 20: Disclosure for Consumer-facing applications

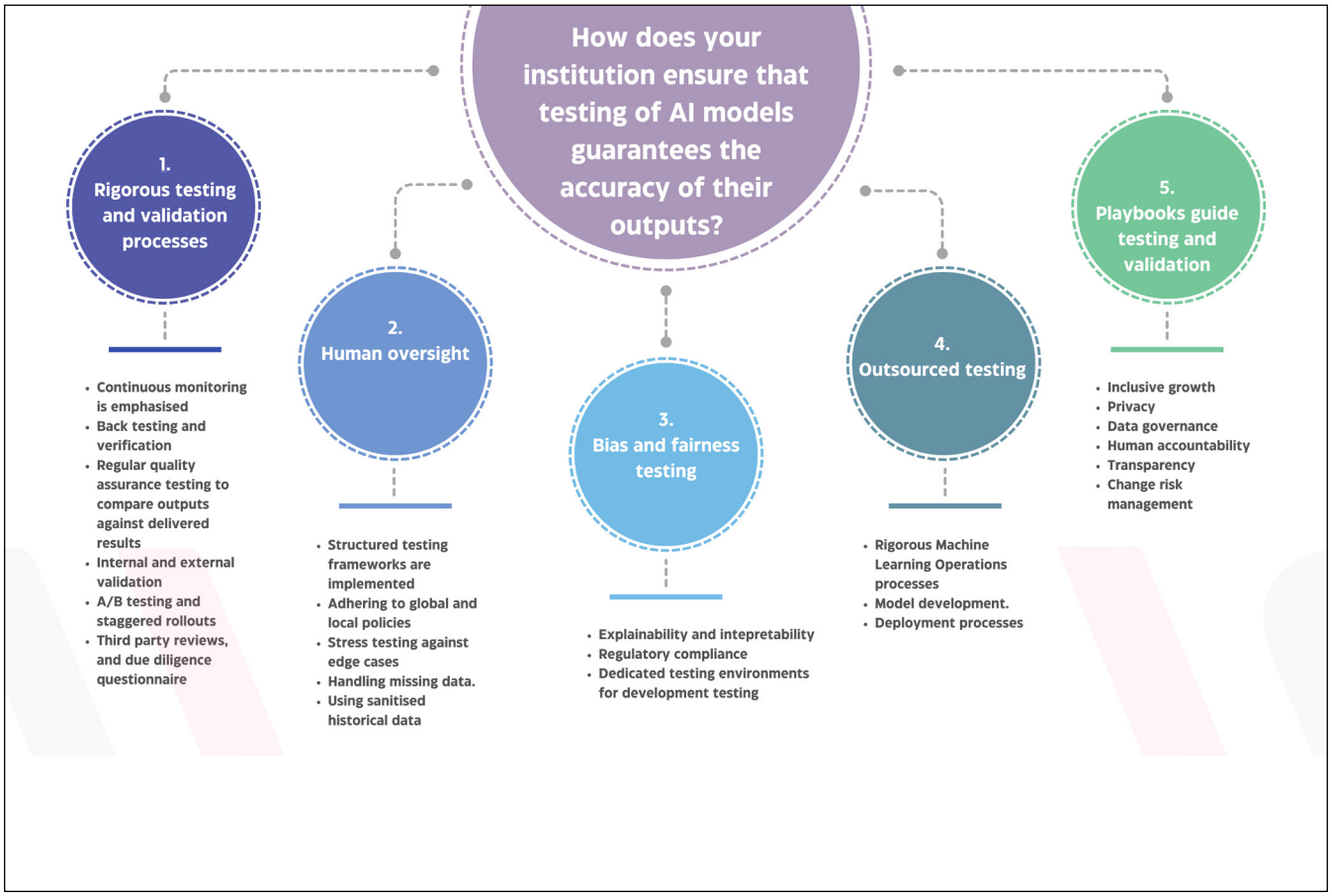


### 6.8.3 TESTING AND VALIDATION

**Question: How does your institution ensure that testing and validation of AI models is performed to guarantee the accuracy and reliability of their outputs?**

There are comprehensive strategies for ensuring the accuracy and reliability of AI model outputs through rigorous testing and validation processes. Continuous monitoring is emphasised, including back testing, verification, and regular quality assurance testing to compare outputs against delivered results. Internal and external validation involves thorough internal testing, such as A/B testing and staggered rollouts, along with independent model validation, third-party reviews, and due diligence questionnaires. Figure 21 illustrates the responses received to this question.

Figure 21: Strategies for ensuring the accuracy and reliability of AI model outputs

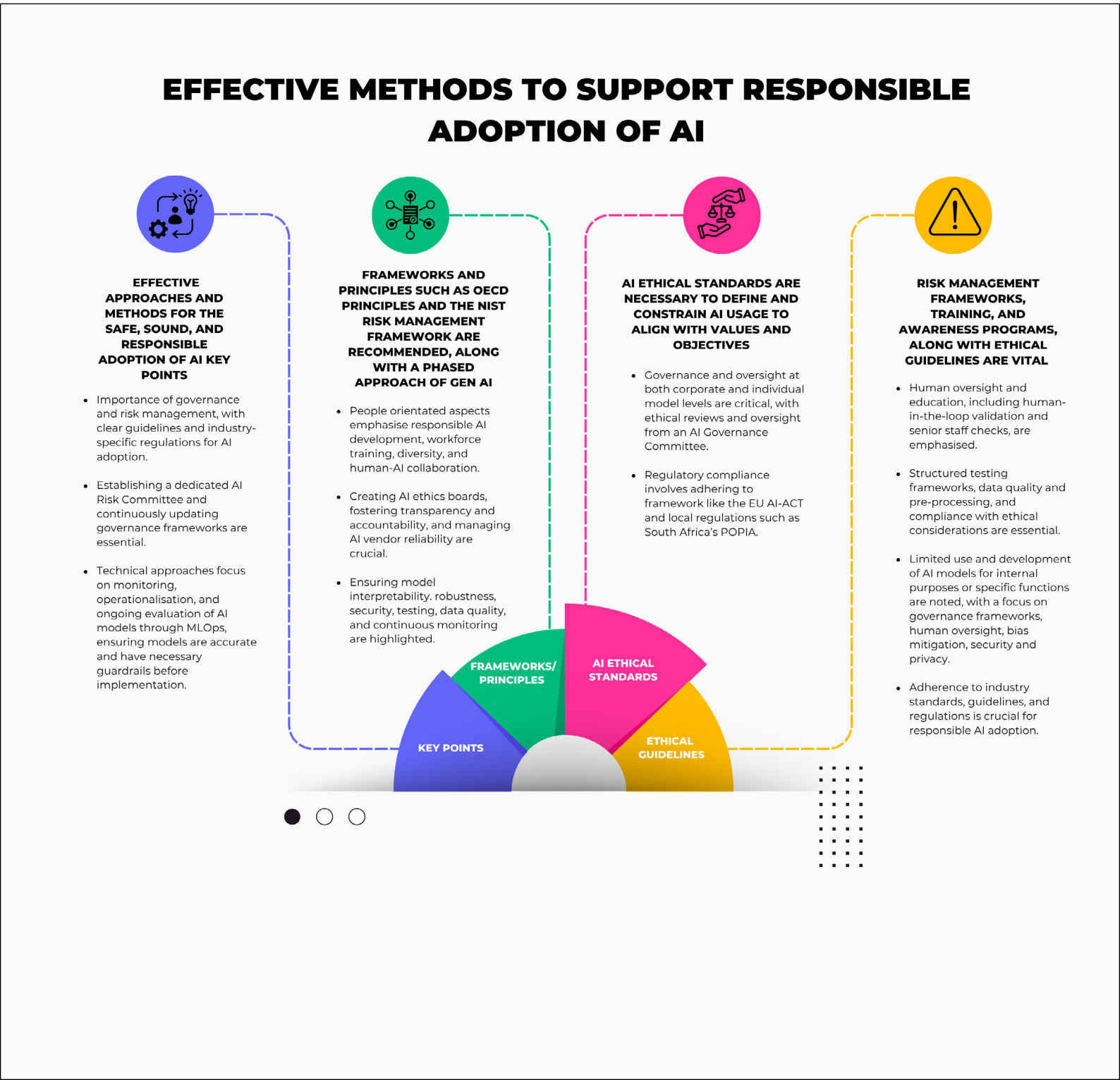


## 6.8.4 ARE THERE EFFECTIVE APPROACHES/METHODS TO SUPPORT THE SAFE, SOUND, AND RESPONSIBLE ADOPTION OF AI?

**Question: What types of disclosure are there for customer-facing AI applications?**

This section summarises the survey responses to effective approaches and methods for the safe, sound, and responsible adoption of AI, particularly in the financial sector, which are illustrated in Figure 22.

Figure 22: Effective approaches/methods to support the safe, sound, and responsible adoption of AI



# PART 7: LESSONS AND INSIGHTS

As South Africa's financial sector continues to integrate AI technologies, it becomes increasingly important to look beyond current applications and consider the foundational elements required for sustainable, ethical, and effective AI adoption.

The path forward requires a shift in mindset: from viewing AI as a purely technical innovation to understanding it as a socio-technical system that intersects with governance, ethics, consumer protection, and financial stability. As such, the future of AI in finance will depend not only on technological capability, but also on the strength of institutional frameworks, the clarity of regulatory expectations, and the inclusiveness of public dialogue. Below, we summarise the overall lessons learned and insights gained. The intention is to ignite the necessary dialogue and inspire support for a coordinated approach to AI that is both ambitious and grounded in the realities of South Africa's financial landscape.

**Encourage the usage of defined explainability methods**

Explainability methods, such as Feature Importance, SHAP values and LIME, are crucial for understanding how AI models make decisions. In the South African financial sector, these methods can help to ensure transparency and build trust among consumers and regulators. By adopting explainability techniques, financial institutions can provide clear insights into AI-driven decisions, making it easier to identify and mitigate biases, improve model accuracy, and comply with regulatory requirements.

**Encourage institutions to establish governance frameworks**

Robust governance frameworks, including data governance, model risk and risk management, are essential for the responsible deployment of AI. These frameworks help institutions manage data quality, ensure compliance with regulations, and mitigate risks associated with AI systems. In South Africa, establishing these frameworks would support financial institutions in maintaining ethical standards, protecting consumer data, and enhancing the overall stability of the financial sector.

**Coordination with the Information Regulator on compliance with POPIA**

Given the particular focus on data privacy, coordination with South Africa's Information Regulator is vital to ensure compliance with POPIA. The Information Regulator is mandated to promote and protect personal information through education, complaint investigation, and enforcing compliance with POPIA. Collaboration with the regulator will therefore help financial institutions navigate the complexities of data protection laws, implement robust privacy measures, and foster consumer trust. Ensuring compliance with POPIA is crucial for safeguarding sensitive financial information and maintaining the integrity of AI systems.

**Prioritisation of certain use cases**

High-impact use cases should be ideally prioritised to ensure responsible and effective application in areas that significantly affect the financial system and consumer protection. These high-impact use cases include fraud detection and prevention, risk management, regulatory compliance, credit scoring and lending, and consumer protection. By focusing on these areas, financial institutions and regulators can maximise the benefits of AI while mitigating risks and ensuring a fair and transparent financial system.

**Standards for ethical, fair, and responsible AI in finance**

Developing standards for ethical, fair, and responsible AI is crucial for ensuring that AI technologies are used in a manner that benefits all stakeholders. These standards should ideally emphasise accountability, transparency, human-in-the-loop, and fairness, guiding financial institutions in the ethical deployment of AI. In South Africa, establishing such standards will help mitigate risks, prevent discriminatory practices, and promote trust in AI-driven financial services.

**Robust oversight to mitigate risks of bias and inaccuracies**

Robust oversight mechanisms to monitor AI systems and mitigate risks of bias and inaccuracies should be prioritised. Some mechanisms could include regular audits, bias detection tools, and continuous monitoring to help identify and address issues in AI models. Implementing these oversight measures will ensure that AI systems operate fairly and accurately, protecting consumers and maintaining the credibility of financial institutions.

**Appropriate and effective disclosure particularly for AI**

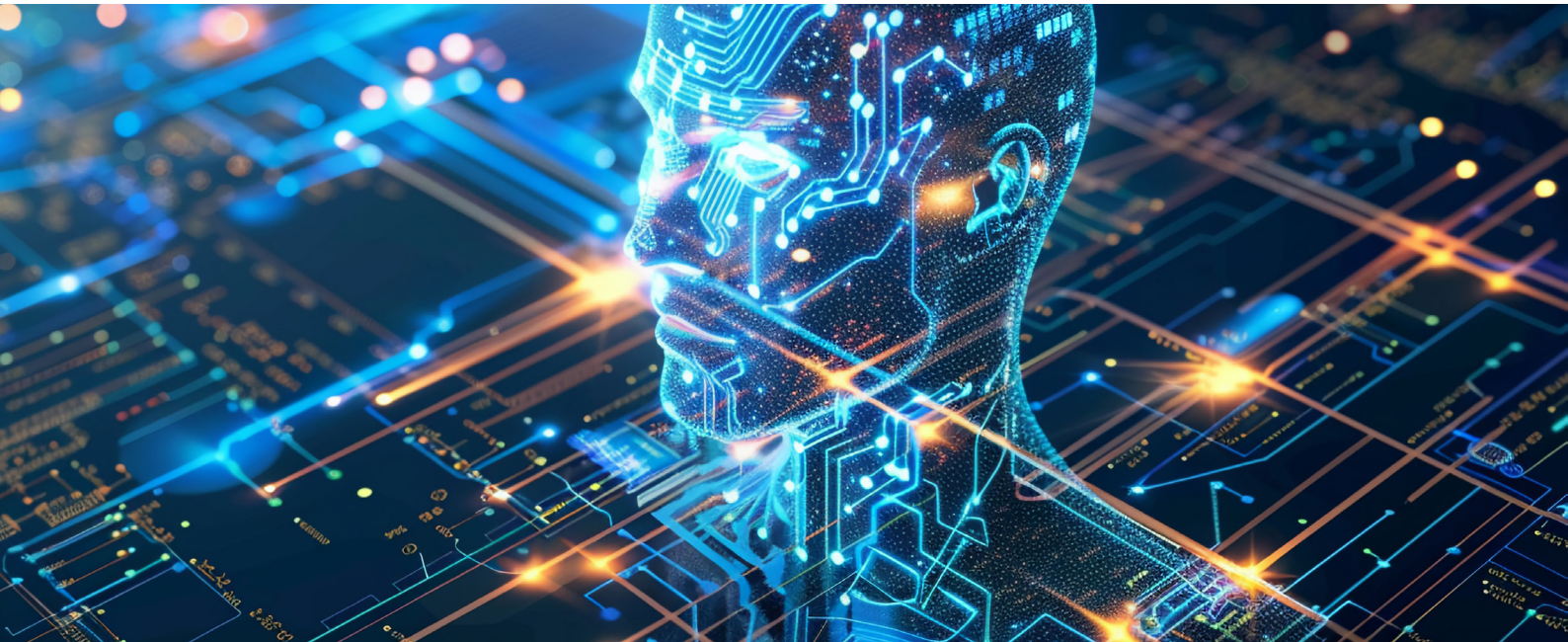
Appropriate and effective disclosure is essential for maintaining transparency in AI-driven decision-making processes. Customers should know when AI is used, for example in evaluating creditworthiness, determining insurance premiums, or making other significant decisions. Clear communication about AI usage will help consumers understand how their data is being used and will help build public trust in AI systems.

**Promote digital literacy, AI literacy, and financial education**

An often overlooked but essential component of ethical and inclusive AI deployment in the financial sector is the promotion of digital and AI literacy among consumers. For transparency and accountability mechanisms to be meaningful, consumers must have a foundational understanding of how AI works, how it may affect their financial decisions, and what rights and recourse options are available to them.

- Digital and AI literacy: Consumers need to be equipped to recognise when they are interacting with an AI system and to understand the implications of algorithmic decision-making in areas such as credit, insurance, and investment advice. Financial institutions and regulators should provide accessible educational materials and awareness campaigns that demystify AI technologies.
- Financial education in the AI age: As AI reshapes financial products and services, financial literacy programmes must evolve accordingly. Consumers need support in navigating automated platforms, understanding AI-generated advice, and exercising judgement in digital financial ecosystems. Tailored financial education initiatives can bridge the gap for vulnerable or digitally excluded populations.
- National collaboration: Regulators and financial institutions should collaborate on the development of national AI literacy strategies, including:
    - Plain-language disclosures for AI-driven decisions
    - Community-based digital literacy initiatives
    - Inclusion of AI awareness modules in school curricula or consumer outreach
    - Platforms for consumer feedback and recourse in cases of AI misuse.

Without adequate public understanding, even well-intentioned transparency and fairness mechanisms may fail to protect consumers. Embedding AI and financial literacy in the broader governance agenda is essential to ensure inclusive and informed participation in AI-powered financial services.



# PART 8: CONCLUSION AND WAY FORWARD

The integration of AI in South Africa's financial services sector offers transformative potential to enhance efficiency, improve decision-making, and deliver personalised customer experiences. However, the adoption of AI also brings forth challenges such as data privacy, ethical considerations, and regulatory compliance, which must be addressed to ensure responsible and sustainable AI deployment.

The insights gathered from the industry survey underscore the cautious yet progressive approach of South African financial institutions towards AI investment and implementation. While sectors like banking and fintech are leading in AI adoption, others like insurance and lending are more reserved, reflecting a need for greater confidence and clarity in AI's potential and regulatory landscape.

To fully harness the benefits of AI while mitigating its risks, it is imperative for regulators to establish robust governance frameworks, invest in upskilling their workforce, and foster a culture of continuous learning and innovation. Collaboration with regulatory bodies and adherence to ethical standards will be crucial in navigating the complexities of AI integration.

Importantly, this is not just a financial sector initiative – it is part of a broader countrywide endeavour to build a sovereign, inclusive, and ethical AI ecosystem. The National Advisory Council on Innovation (NACI) has emphasised the need for a human-centred, African-oriented, and globally competitive approach, grounded in principles of trust, fairness, transparency, accountability, and collaboration. Aligning financial sector efforts with national strategies will be key to ensuring that AI serves the public good, supports sustainable development, and reflects South Africa's constitutional values.

The Financial Sector Conduct Authority and the Prudential Authority are mindful of legal developments that may take place in other parts of government and will endeavour to align to these as appropriate, including whether a countrywide approach is developed that encompasses all sectors and not just finance.

As a next step, the Financial Sector Conduct Authority and the Prudential Authority will develop a discussion paper that will build on the findings of this joint FSCA–PA AI research report and will engage stakeholders on key regulatory and supervisory questions. The discussion will also support alignment with broader national AI strategies and promote coordination across financial sector regulators.

# PART 9: LIST OF ACRONYMS

| Acronym | Full Term |
|---|---|
| AI | Artificial Intelligence |
| AML | Anti-Money Laundering |
| BIS | Bank for International Settlements |
| DL | Deep Learning |
| DLM | Data Lifecycle Management |
| FCA | Financial Conduct Authority |
| FSCA | Financial Sector Conduct Authority |
| GAN | Generative Adversarial Network |
| GenAI | Generative Artificial Intelligence |
| LLM | Large Language Model |
| ML | Machine Learning |
| MLOps | Machine Learning Operations |
| NLP | Natural Language Processing |
| NN | Neural Network |
| OECD | Organisation for Economic Co-operation and Development |
| PA | Prudential Authority |
| RL | Reinforcement Learning |
| SARB | South African Reserve Bank |
| SEC | Securities and Exchange Commission |
| SHAP | SHapley Additive exPlanations |
| TCF | Treat Customers Fairly |
| XAI | Explainable Artificial Intelligence |

# PART 10: GLOSSARY OF TERMS

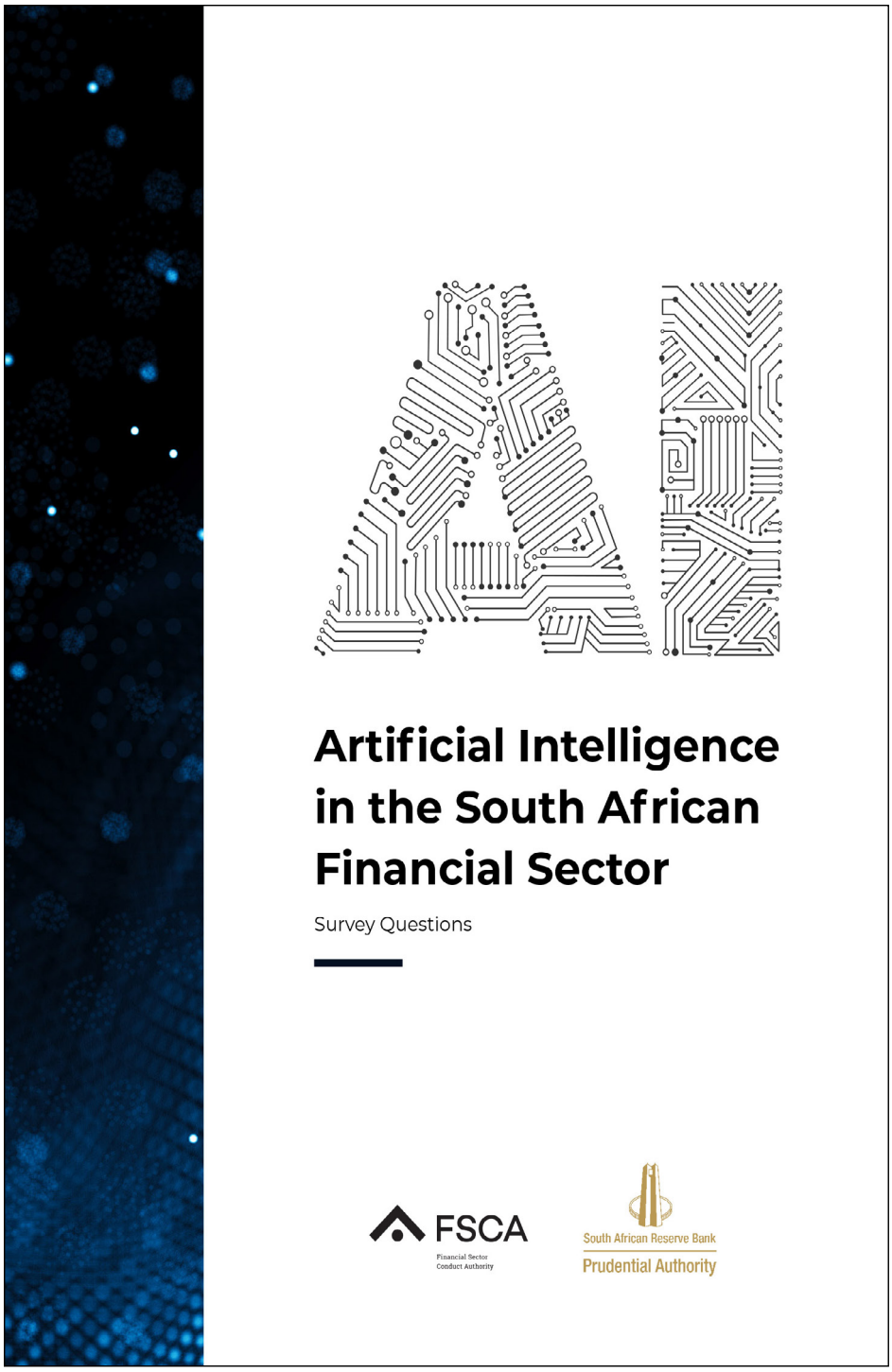
| Term | Definition |
|---|---|
| **Artificial Intelligence System** | An AI system is a machine-based system that, for explicit or implicit objectives, infers from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment. |
| **AI System Autonomy** | The degree to which a system can learn or act without human involvement following the delegation of autonomy and process automation by humans. |
| **Big Data** | Data that are too large or complex to be handled by conventional data processing tools and techniques. |
| **Consumer Protection** | Consumer protection is the safeguarding of consumers' rights and interests in the financial sector. This involves ensuring that financial institutions operate fairly, transparently, and responsibly, providing consumers with accurate information and protecting them from unfair practices and financial harm. |
| **Cloud Computing** | Cloud systems and applications are digital storage and computing resources remotely available on-demand via the Internet. |
| **Data Lifecycle Management** | Data Lifecycle Management (DLM) refers to the process of managing data throughout its entire lifecycle, from creation to deletion. |
| **Data Privacy** | Data privacy is the protection of personal information, ensuring that individuals have control over how their personal data is collected, used, and shared. This is primarily governed by the Protection of Personal Information Act (POPIA), which aims to safeguard personal information by enforcing principles such as accountability, transparency, and security |

| | |
|---|---|
| **Deep Learning** | Deep learning is a subset of machine learning that uses multilayered neural networks, known as deep neural networks, to simulate the complex decision-making power of the human brain. These networks consist of multiple layers of interconnected nodes, each building on the previous layer to refine and optimise predictions or categorisations.<br><br>Deep learning models can handle large volumes of unstructured data and are used in various applications, such as digital assistants, self-driving cars, and fraud detection. |
| **Financial Stability** | Financial stability refers to a financial system that inspires confidence through its resilience to systemic risks and its ability to efficiently intermediate funds. Protecting and enhancing financial stability can be achieved through monitoring the environment and mitigating systemic risks that might disrupt the financial system. A stable and well-functioning financial system significantly contributes to balanced and sustainable economic growth. |
| **Foundation Model** | A foundation model, also known as large X model (LxM), is a machine learning or deep learning model that is trained on vast datasets so it can be applied across a wide range of use cases. Generative AI applications like Large Language Models are often examples of foundation models. |
| **Generative Artificial Intelligence** | Generative AI (GenAI) is a type of artificial intelligence that can create original content—such as text, images, video, audio, or software code—in response to a user's prompt or request. GenAI relies on sophisticated machine learning models, particularly deep learning models, which simulate the learning and decision-making processes of the human brain.<br><br>These models identify and encode patterns in large amounts of data, enabling them to generate new, relevant content based on user inputs. |
| **Neural Network** | A neural network is a type of machine learning (ML) process in artificial intelligence (AI) that teaches computers to process data like the human brain. It uses a layered structure of interconnected nodes, called neurons, to process data. |
| **Knowledge Representation** | Knowledge representation is a field within artificial intelligence (AI) that focuses on representing information about the world in a form that a computer system can use to solve complex tasks. This involves creating models that capture the relationships between different entities and concepts, enabling machines to reason and make decisions based on this structured information. |

| | |
|---|---|
| **Large Language Model** | Large Language Models (LLMs) are a category of foundation models trained on immense amounts of data, enabling them to understand and generate natural language and other types of content. These models are designed to perform a wide range of tasks, such as text generation, translation, summarisation and answering questions.<br><br>LLMs leverage deep learning techniques and vast datasets to capture intricate patterns in language, making them capable of producing coherent and contextually relevant responses. |
| **Machine Learning** | A method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns, and make decisions with minimal human intervention. |
| **Natural Language Processing** | Natural Language Processing (NLP) is a subfield of artificial intelligence (AI) and computer science that uses machine learning to enable computers to understand, interpret, and generate human language. NLP combines computational linguistics with statistical, machine learning, and deep learning models to process and analyse large amounts of natural language data. This technology powers various applications, including search engines, chatbots, voice-operated systems, and digital assistants. |
| **Neural Network** | A neural network is a machine learning model that mimics the way biological neurons work together to identify phenomena, weigh options, and arrive at conclusions. Neural networks consist of layers of nodes, or artificial neurons, including an input layer, one or more hidden layers, and an output layer. Each node connects to others and has its own associated weight and threshold1. If the output of any node exceeds a specified threshold, it activates and sends data to the next layer. |

# PART 11: SURVEY QUESTIONS

Click to access the survey questions



**Financial Sector Conduct Authority**
Riverwalk Office Park, Block B
41 Matroosberg Road
Ashlea Gardens
Pretoria
0002

012 428 8000
www.fsca.co.za

**Prudential Authority**
370 Helen Joseph Street
Pretoria
0001

012 313 3911
www.resbank.co.za