Ethical dimensions of artificial intelligence in the South African banking sector

Working paper series

WP/2024/01

The Centre of Excellence in Financial Services is a non-profit organization focusing on collaborative research, bringing together local and international thought leaders, industry experts and academics to interrogate the role financial services can play in achieving national and international objectives. This working paper is available in electronic format on www.coefs.org.za.

Table of contents

Preface

The launch of ChatGPT in 2022 saw an unprecedented growth in users over just a few days, breaking all known records. By 2024, ChatGPT is not only a household name but has been woven into the fabric of our society. Artificial intelligence (**AI**) has largely evaded the attention of ordinary citizens, with depictions of androids in science fiction movies capturing their imagination instead. However, with ChatGPT, the power of its offering can no longer be ignored. With the many references to monetising data for profit and benefit, AI has become the conductor of this data orchestra, and essential questions are being asked about the impact of our colonial past, how our data is being used, and who really benefits from this AI revolution.

The banking industry has long been an early adopter of technology and in some circles, may be viewed not just as banks, but as some of the largest IT companies in the country. It is, therefore, a natural fit for banks to take a leadership role in adopting AI in South Africa, and they have done just that. We wanted to understand whether the current emphasis on ethical behaviour across the banks, was transferable to this emerging discussion of technological innovation. We also wanted to see if we could make actionable recommendations that could be considered by the banks, bank regulators, and financial policymakers alike.

In preparing this working paper, we partnered with three experts in AI from academia to join us on this journey, with the objective of separately publishing an academic paper that follows the rigour associated with academic research. This working paper combines the lessons from the academic research with our own observations, as we embrace the emerging topic of AI.

We canvassed the opinions of many stakeholders during this research. Given the importance of the topic, we made sure we engaged senior executives from the C suite, risk management, compliance, governance, and IT leadership, each well-versed in the subject and representing a broad range of banks. At the same time, we wanted to capture the contribution of a cohort of bankers, not those selected for interviews and

less familiar with the technical application of AI, but well-versed in the ethical considerations relevant to banks. This was done through an anonymous online survey. We also made time to engage bank regulators and policymakers who must take responsibility for regulating the concerns being raised about AI.

This working paper represents an important step when considering how advances in technological innovation must be embedded in ethical governance. It is intended to inform future discussions and policies, giving stakeholders the considerations required to navigate the complexities of AI in banking. By exploring these challenges, COEFS hopes to contribute meaningfully to developing an AI ecosystem that benefits all stakeholders while observing the highest ethical standards.

Mark Brits
Executive Director

Members of the working group

in alphabetical order

Mark Brits

Committee Chair
Executive Director
Centre of Excellence in Financial Services

Prof. Manoj Maharaj

Professor: Information Systems and Technology
University of KwaZulu-Natal

Asande Mahlaba

Research Administrator
Centre of Excellence in Financial Services

Dr Kiru Pillay

Honorary Senior Lecturer
University of KwaZulu-Natal

Prof. Bruce Watson

Chief Advisor
National Security Centre of Excellence (Canada)

Members of the work group participate in their personal capacities, without remuneration, and their views do not necessarily represent those of their employer.

Abbreviations

AI	Artificial intelligence
HCI	Human-computer interaction
ML	Machine learning
POPIA	Protection of Personal Information Act
EU AI Act	European Union Artificial Intelligence Act
FinTech	Financial technology

Executive summary

This working paper investigates the ethical dimensions of AI within South Africa's banking sector, focusing on its transformative potential and associated challenges. As AI technologies become further integrated into financial operations, they offer significant benefits, including enhanced efficiency and improved decision-making capabilities. However, these advances raise critical ethical concerns, such as data privacy, algorithmic bias, transparency, and accountability, which require immediate and sustained attention.

The working paper is based on an academic study facilitated by COEFS, that involved a range of banking professionals and regulators to identify sector-specific concerns and potential solutions. This working paper surfaces the prevailing sentiment, including a strong industry inclination towards embracing AI and a clear recognition of its ethical challenges. Key areas of concern include safeguarding data privacy, addressing biases in decision-making algorithms, and ensuring transparency in AI-driven processes.

To address these issues, the paper outlines a series of actionable recommendations. These include developing fair AI design frameworks, enhancing public and workforce education on AI, and adopting risk-based governance structures. It also highlights the need for adaptive regulatory frameworks that balance innovation with consumer protection, ensuring alignment with global standards while considering local socio-economic contexts. The working paper does not pretend to provide all the answers, so some questions are left open-ended for future exploration.

The insights presented in the working paper highlight the importance of collaboration among stakeholders, including banks, regulators, and technology developers, to create an ethical AI ecosystem that supports innovation and trust. In harnessing the powerful potential of AI technologies, banks have a responsibility to consider the associated ethical challenges. South African banks, as established leaders in many facets of banking innovation, have an opportunity to deploy AI's potential

responsibly, setting a benchmark for ethical AI implementation in South Africa and the broader African Continent. This paper aims to serve as a resource for decision-makers committed to navigating the ethical dimensions of AI.

Introduction

As far back as 1950, noted English mathematician and computer scientist Alan Turing first discussed the idea that computers could think. Since then, developments in AI have generally been low-key, with various growth spurts interspersed with stages of limited activity. More recently, even though not yet ingrained in the public consciousness, AI has been evident in various industries and aspects of daily life. In business, AI-powered tools have been widely used in customer service through chatbots. In healthcare, AI-supported diagnostic tools enable early detection of diseases like cancer through image recognition technology. Financial services rely on AI for fraud detection, credit scoring, and algorithmic trading. However, the arrival of ChatGPT and recent rapid technological developments have led to the widespread increase of AI's scope towards more complex and adaptive roles.

The ongoing development and integration of AI raises important ethical and societal questions, particularly concerning the responsible use and governance of AI technologies. The lack of transparency associated with AI models, whose inner workings are often called "black boxes," can make it difficult to understand and explain their decision-making processes. Additionally, the potential for AI to automate jobs raises concerns about the displacement of the human workers, highlighting the need for upskilling and reskilling. Various issues must be addressed as AI systems become more integrated into decision-making processes.

This working paper synthesises the findings from a comprehensive literature review on AI ethics, specifically focusing on applications relevant to the banking sector in South Africa. Additionally, the working paper explores the implications of a detailed survey, and a framework analysis carried out within the banking sector.

While AI is still an emerging topic, there is a growing body of commentary on the matter. Given the strategic nature of the research, we sought to engage with senior bankers who understood the issues. These bankers were drawn from a diverse range of banks. Separately, we canvased the views of the financial sector regulators and

policymakers who must take steps to address the concerns being raised. At the same time, we still wanted to capture the contribution of those bankers who were less familiar with the technical application of AI and algorithms but were up to speed in banking culture and ethics in general.

Decoding the ethics of AI

What is AI? The answer depends on who you ask. In the broadest sense, AI refers to computerized systems that perform tasks typically requiring human intelligence. Due to its ability to work more accurately than any human and without ever tiring, AI has become integral to virtually every industry.

Integrating AI is becoming critical in diverse fields such as finance, healthcare, small businesses, and beyond. AI has opened new avenues for innovation but has also introduced a complex web of ethical and operational challenges.

Stakeholders across these sectors are expressing a wide range of concerns. Privacy, security, fairness, accuracy, explainability, empathy, trust, and accountability are at the forefront of these discussions. The voices in this conversation are diverse, including customers, employees, AI developers, regulators, policymakers, experts in human-computer interaction (**HCI**) and machine learning (**ML**), philosophers, ethicists, and professional societies. Each stakeholder brings a unique perspective, highlighting the multifaceted nature of AI ethics and governance.

Defining AI

In the banking industry, where automation has been a reality for many years, arriving at a clear definition of AI can shape how ethics and regulations are applied in this space. Should AI be limited to decision-making algorithms and generative models that expand beyond traditional automation, or do we risk conflating it with futuristic systems, yet to become a reality, that show general human-like intelligence? This distinction matters. If regulations aim too far ahead, they could suffocate innovation in low-risk systems that are only designed to enhance task efficiency or customer service. On the other hand, failing to address emerging generative AI capabilities could lead to ethical blind spots. There is a need to create a definition that is clear but also flexible and that allows banks to innovate responsibly without taking their eye off ethical considerations.

Interventions aimed at addressing these concerns are as varied as the applications themselves. Diagnostic tools in healthcare, customer service chatbots, fraud detection systems, and credit scoring algorithms all demonstrate the broad utility of AI. AI frameworks are being developed within a governance framework to ensure

these systems are used responsibly. For example, human-centred AI systems focus on making interactions more intuitive and fairer, while explainable AI seeks to make the decision-making processes of AI more transparent. There is also a growing emphasis on empathetic and ethical AI, generative AI that respects user privacy, AI soft law that offers flexible regulatory guidance, and frameworks for assessing trustworthiness in AI systems.

The ethical challenges become even trickier when comparing traditional and AI-driven approaches. While still useful, traditional ethical frameworks often fall short when dealing with the complexities introduced by AI. Issues such as governance, human interaction, fairness, and transparency are far more nuanced in the context of AI. Modern approaches to AI ethics attempt to integrate deeper structural issues like racism and information asymmetries, areas that received little attention in traditional frameworks.

Evaluating the effectiveness of ethical practices in AI reveals the need for a more comprehensive approach. It is not enough to address individual biases. Effective practices must tackle structural issues that underpin these biases. There is also a delicate balance to be struck between accuracy and explainability, ensuring that AI systems are both precise and at the same time understandable. Integrating fairness and empathy into AI design is crucial, especially in a unique socio-economic landscape such as South Africa's. It is also important to address privacy and security concerns in a way that does not compromise the integrity or utility of the AI system.

Each sector faces its own unique set of challenges and opportunities. Addressing these requires a concerted effort from all stakeholders to ensure that AI is developed and deployed in a manner that is ethical, fair, and aligned with the broader values of society. In South African banking, AI integration has become increasingly central to operations, particularly in customer service and financial decision-making. The focus is on how AI can enhance banking processes, improve customer interactions, and streamline financial assessments. However, as AI becomes more entrenched in these operations, it brings a host of ethical and practical concerns that must be carefully navigated.

Stakeholders in this space, ranging from bank customers and employees to management and regulators, are deeply concerned about several key issues. Data privacy is at the forefront, as the vast amounts of personal information handled by AI systems must be protected against breaches and misuse. Algorithmic bias is another major concern, with stakeholders worried about the potential for AI to reinforce existing inequalities or introduce new forms of discrimination. Additionally, the impact of AI on employment is a significant issue, as automation in banking could lead to job displacement, raising questions about the future of work in this sector.

Balancing flexibility and oversight in AI regulation

Considering how fast AI is evolving, should regulators focus on detailed, rigid rules or embrace a principles-based approach? We would argue for the latter. High-level principles like fairness, transparency, accountability, and risk-based deployment can offer a strong guiding compass across industries without stifling innovation. A principles-based approach allows stakeholders to be adaptable as technology evolves. The complexity that comes from industry-specific challenges should rather be left to those industries to resolve. The right balance must be struck that ensures ethical oversight while protecting societal interests without dulling the transformative potential of AI in sectors like banking. If innovation is reliant on flexibility, can regulation (already a slow process) keep pace without becoming an obstacle?

Specific AI applications have been implemented within South African banks to address these concerns. Customer service chatbots are being used to provide efficient, 24/7 support, improving customer satisfaction while reducing operational costs. Fraud detection systems powered by AI are helping banks to identify and prevent fraudulent activities more effectively than traditional methods. Banks are also using credit scoring algorithms to assess customers' creditworthiness more quickly and accurately.

When comparing these AI-driven applications to traditional banking methods, several issues come to light. Transparency is a significant challenge. While AI systems can process vast amounts of data and make decisions rapidly, the reasoning behind these decisions is often opaque, leading to concerns about accountability. Fairness is another critical issue, as the data used to train AI systems may not always represent the full diversity of the population, potentially leading to biased outcomes.

Ethical evaluations of AI in banking emphasise the importance of maintaining fairness and preventing bias in all AI-driven processes. Ensuring privacy is paramount,

as customers need to trust that their personal information is secure. Finally, maintaining customer trust in AI implementation is crucial. Banks must be transparent about how AI is used and demonstrate a commitment to ethical practices to retain the confidence of their customers and the broader public.

As South African banks continue to adopt AI technologies, they must carefully consider the ethical implications of these tools. By focusing on fairness, privacy, and transparency and by addressing concerns around bias and employment, banks can leverage AI to enhance their operations while upholding the trust and confidence of all stakeholders involved.

A journey of discovery: The structure of the research

The journey to our findings began with a comprehensive literature review, which included both academic papers and broader desktop research, focusing on the ethical dimensions of AI as they could apply to the South African banking sector. This initial phase helped us identify key themes such as transparency, fairness, and data privacy.

Following this, we conducted a survey targeting diverse banking professionals to gauge their familiarity with AI and their perspectives on its ethical implications. The survey results highlighted a strong recognition of ethical challenges and a commitment to improving AI practices.

Building on these insights, we conducted a Delphi study comprising four discussion groups. The first two sessions involved separate cohorts of bankers, focusing on AI's role, transparency, bias, and data privacy. The third session synthesized these insights, bringing all previously interviewed bankers together to reach a consensus on key issues.

The final session engaged financial regulators and policymakers, presenting the bankers' concerns and discussing regulatory frameworks, public education, and the balance between innovation and regulation. This structured, iterative process allowed us to refine our understanding and develop actionable recommendations for ethical AI adoption in South Africa's banking sector.

What are the dominant themes in AI literature?

Our preparations for this working paper began with a systematic literature review based on the PRISMA protocol. This academic literature review included papers from the last four years, mostly set in the African and specifically South African context. Here is what many commentators agreed upon:

Is bias built into AI?

The reading in this area raises significant ethical challenges regarding AI, particularly its role in perpetuating societal biases and inequalities. Many discussions focus on bias mitigation and fairness but fail to address deeper systemic issues like structural racism and anti-blackness. AI systems designed and developed within historically unequal societal contexts risk amplifying these disparities.

A good example of this can be found in technologies like facial recognition, which have consistently struggled to identify black faces, reinforcing marginalisation and discrimination. The way in which data is collected and represented in AI further entrenches these biases, as shown by examples where specific search terms in early search engines linked blackness to dehumanizing terms.

Many authors argue that achieving fairness in AI calls for a broader approach than simply addressing technical biases. Ethical frameworks must consider the historical and structural realities that shape AI design and deployment. For instance, while increasing transparency or fairness in surveillance technologies might seem like a good idea, it doesn't address the systemic issues of over-policing or the surveillance of marginalized communities. The literature emphasizes the importance of designing AI systems that reflect the lived experiences of those most affected by bias, particularly black individuals, and calls for a shift from narrow technical fixes to addressing the foundational inequalities found in many AI systems.

How biased AI can affect societies

Some research dwelled on the broader societal impact of AI, particularly highlighting how systemic biases in AI systems affect social and economic structures. They argue that anti-blackness embedded in AI technologies can perpetuate inequities, impacting marginalized communities disproportionately. These biases can shape economic opportunities and social mobility. This reality clearly reveals the need for a more inclusive and equitable approach to AI design and deployment.

Transparency and accountability also emerge as critical factors in addressing these issues. Many commentators stress the importance of making AI systems more transparent to uncover and correct biases that disproportionately harm black communities. Additionally, some thinkers advocate for inclusive public participation, emphasizing the involvement of black communities and other marginalized groups in the development of AI systems. Engaging the public ensures that AI technologies better reflect diverse perspectives, reducing the risk of reinforcing harmful societal biases.

Specific ethical concerns

Some research explores how AI systems embed and perpetuate racial biases. Beyond the data itself, subjective and unconscious biases in AI reflect deeper societal inequalities, often overlooked in approaches that treat bias as a purely technical problem. These systemic flaws affect automated decision-making systems, even within the financial sector, which, without careful ethical consideration, can replicate and amplify structural racism.

Transparency is a recurring theme in these discussions. Many commentators call for greater auditing and accountability in AI systems to identify and address the biases embedded within them. Ethical data-sharing practices and more rigorous oversight are essential to ensure AI systems are fair and equitable. The papers argue that by exposing and mitigating these hidden biases, AI can be designed to challenge, rather than reinforce, existing societal injustices.

Key themes across our research

The list below captures the key themes identified across all our research.

1. **Fairness and non-discrimination:** AI algorithms should avoid biases and discrimination based on race, gender, or socioeconomic status.
2. **Transparency and explainability:** AI systems need transparency and explainable decision-making processes to ensure accountability and build trust, especially in high-stakes applications like finance and healthcare.
3. **Human oversight and control:** Maintaining human oversight and control over AI systems is vital to prevent unintended consequences and ensure ethical decision-making, particularly in autonomous systems.
4. **Privacy and data protection:** Ethical AI development should prioritise safeguarding personal information and complying with privacy regulations due to the vast amounts of data that AI applications rely on.
5. **Accountability and responsibility:** Developers, users, and organisations should be held accountable for AI systems' functioning and potential harm, addressing algorithmic bias and unintended consequences.
6. **Beneficence:** AI should be used for social good, improving human well-being and addressing global challenges like poverty, inequality, and climate change.
7. **Systemic risks:** AI systems in finance can lead to systemic risks like flash crashes and algorithmic collusion, necessitating ethical considerations beyond individual applications.
8. **Human-AI interaction:** The interaction between humans and AI raises ethical questions about trust, autonomy, and the potential for AI to manipulate or deceive, which is relevant in customer service, and financial advice.

9. **AI in developing countries:** Deploying AI in developing countries involves unique ethical considerations like exacerbating existing inequalities and the need for contextually relevant solutions.

10. **Environmental impact:** AI development and use significantly impact the environment, including carbon emissions and resource consumption, requiring sustainable, ethical AI development.

11. **The exploitation of labour:** AI development often relies on low-wage workers for data annotation and labelling tasks, necessitating fair labour practices and compensation.

12. **Job displacement and the future of work:** AI automation raises concerns about job displacement, requiring workforce reskilling and upskilling to ensure a just transition for workers.

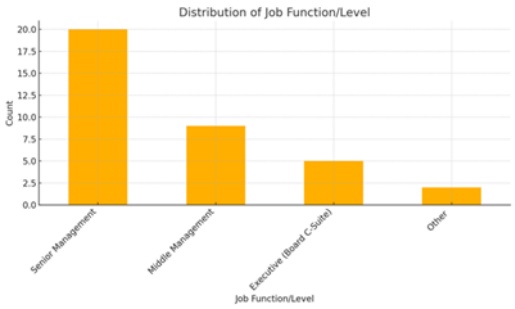
13. **Lack of diversity and inclusion:** AI development should prioritise diversity and inclusion to ensure representative and beneficial AI systems for all society members.

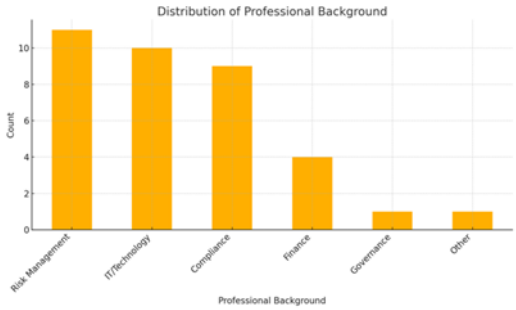
Using a survey to understand prevailing opinion

After completing the thematic analysis of the readings, we set about engaging far and wide to see how financial services workers in South Africa felt about the issues. We identified several respondents in the financial services sector for an online confidential survey.

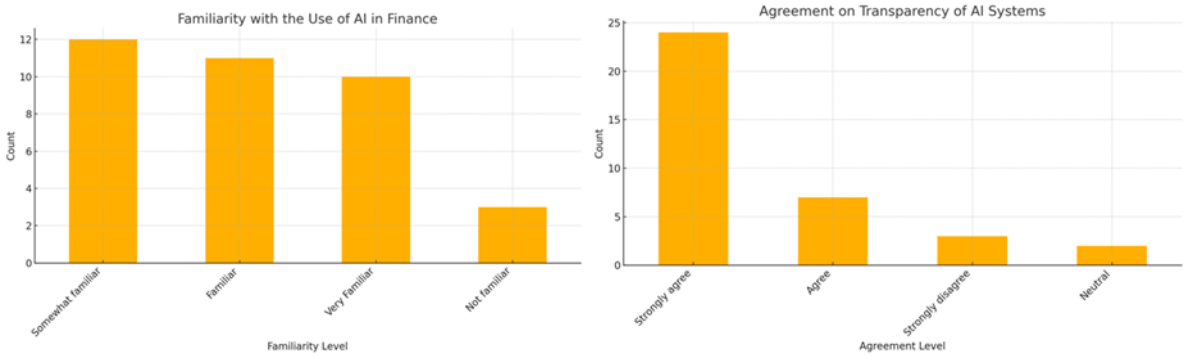
These bankers were more diverse in their skill sets and not immersed in the business of AI. They had other skills and areas of responsibility with the majority having more than15 years banking experience. Some of the roles they represented included HR, ethics, and digital transformation. Some were regional heads for African subsidiaries, and others were immersed in data privacy, data analytics, and operations. What connected them was that they all understood corporate ethics and were steeped in creating and taking part in diverse corporate cultures.

Seventy-five surveys were distributed. The two graphics below reveal the different job functions and professional backgrounds of the respondents.





The two graphics below reveal the respondents' familiarity with the use of AI in finance and how they mostly agreed that transparency of AI systems was important.



The 36 respondents were mostly familiar with the use of AI in finance and agreed that transparency was important.

In South Africa's banking sector, where issues of fairness, equity, and access are critical, AI adoption presents both opportunities and ethical challenges. The banking sector operates within a unique socio-economic landscape, where issues of fairness, equity, and access to financial services are never far from mind. As banks in South Africa increasingly adopt AI, they must do so with greater awareness of the ethical implications.

The challenge is to harness the power of AI in a way that balances innovation with responsibility. This involves adhering to international best practices and developing context-specific guidelines that reflect the values and priorities of the South African market.

In what follows, we provide a high-level summary of the survey results related to the ethics of AI in the banking sector. The findings serve as a foundation for the Delphi study to refine and build consensus on key ethical challenges and opportunities in AI adoption within the South African banking sector.

1. Positive attitudes toward AI:

Most respondents are enthusiastic about using AI in their organisations. With 90% of respondents indicated that they either "Agree" or "Strongly Agree", having

a positive attitude towards using AI in their organisation. This support suggests that banking professionals are ready to embrace AI technologies.

2. Recognition of ethical challenges:

There is a clear awareness of the ethical challenges associated with AI, reflecting a balanced and mature perspective on AI adoption. 57% of respondents acknowledged significant ethical challenges by agreeing or strongly agreeing with the relevant statement. This indicates that while there is optimism, there is also a strong recognition of the need to address ethical concerns.

Streamlining AI regulation for practical oversight

How do we make AI regulation both effective and efficient without drowning in redundancy? The answer lies in leveraging existing frameworks for data protection, cybersecurity, and operational risk. This approach can streamline oversight while addressing AI-specific concerns like bias in decision-making. However, this approach is not without its challenges. Regulators must balance addressing genuine societal threats without overregulating. We need to consider whether AI is always a threat that demands regulation or whether there are areas where the risks are exaggerated. Moreover, while essential in preventing discrimination, ethics may be a liability when dealing with fraud or criminal activity. There is a need to create AI policies that are principled yet practical. These policies must avoid obscure or theoretical ideals that get in the way of implementation.

3. Importance of industry standards and regulations:

Industry standards and regulatory frameworks are critical to guiding ethical AI practices, with 81% of respondents either agreed or strongly agreed that industry standards are important in shaping their AI practices. This consensus underscores the reliance on structured guidelines to ensure ethical AI implementation in the banking sector.

4. Commitment to ethical AI practices:

Organisations are strongly committed to continually enhancing ethical AI practices. 75% of respondents agreed or strongly agreed that their organisations are committed to continuously improving ethical AI practices.

5. Room for improvement in AI governance:

While there is confidence in current practices, some respondents see opportunities for improving AI governance. Only 27% of respondents agreed that their organisation needs to improve its AI governance and ethical practices, suggesting areas for further effort.

Four discussion groups

A Delphi study is a research method used for gathering and refining expert opinions on a specific topic or problem through a structured, iterative process. It is particularly useful for addressing complex issues such as ethics in AI where there is limited empirical data or a need for consensus among diverse experts.

After performing a research review of ethics in AI and supplementing with the survey we discussed in the last section, it was time to perform a Delphi study comprising four discussion groups. We separated bankers of the same brand between the two sessions to make sure we had open and frank discussions and, at the same time made sure smaller banks had a voice.

The first banking cohort

The session began with a discussion of the survey results highlighting concerns about AI ethics, particularly around fairness and bias.

Key topics and insights

1. **AI in banking:** AI's role in the banking sector is generally seen as evolutionary rather than revolutionary. Participants noted that banks had used AI in lower-risk applications and emphasised the need for a standardised governance framework. The implications of generative AI were explored, particularly its effect on junior employees' learning opportunities.

2. **Transparency and accountability:** Emphasis was placed on creating unified frameworks for AI use, to ensure transparency and allow clients to understand AI-driven decisions.

3. **Data privacy and security:** The complexity of managing data in AI environments, particularly with third-party vendors, was highlighted as a concern.

4. **Future of AI in banking:** The group expressed that, while AI integration will continue, the current hype might be overstated.

The second banking cohort

Building on the previous discussion, this session re-emphasised the four primary themes, focusing on transparency in AI applications and ensuring ethical AI usage.

Key topics and insights

1. **Transparency and accountability:** There was a consensus view on the necessity for public disclosures when AI is involved, especially in decision-making that impacts clients. The group discussed the importance of distinguishing between AI decision-making and bot interactions. Practical aspects, such as clear customer information on AI processes, were also debated.

2. **Bias and fairness:** Addressing bias in AI, especially in algorithmic decision-making, was underscored. The group highlighted the need for representative samples and frameworks to mitigate biased outcomes.

3. **Financial inclusion and job market impact:** This session explored AI's role in enhancing financial access and ethical concerns about job displacement. Participants noted the importance of AI literacy and financial education as tools to combat exclusion.

4. **Data privacy and security:** Harmonized vendor disclosure practices and robust data privacy protocols were recommended to ensure customer data protection, aligning with regulatory expectations like the Protection of Personal Information Act (**POPIA**).

The third working session

This session aimed to synthesise insights from prior discussions. Here, all the previously interviewed bankers were brought together to discuss their insights and reach a consensus.

Key topics and insights

1. **Transparency and accountability:** A key focus remained on transparency in AI applications, balancing customer understanding with proprietary limitations. Suggestions were made to clarify disclosures around AI's role and improve explainability.
2. **AI definition and governance:** Participants debated the appropriate governance structures for AI and the potential need for an industry-specific framework that does not overburden innovation. The European Union Artificial Intelligence Act (**EU AI Act**) was referenced as a model for South Africa to consider.
3. **Ethical AI deployment:** Delegates discussed the importance of the ethical deployment of AI, emphasising a risk-based approach and internal governance to ensure ethical standards are met while protecting business interests.

Consider risk-based implementation

A risk-based approach recognizes that not all AI applications pose equal threats. Systems influencing financial decisions, like credit scoring or fraud detection, clearly call for stricter oversight than tools assisting employees, such as call centres, and chatbots. However, the line between these applications blurs when considering the impact on different customer classes. Depositors, considered more vulnerable, need higher standards of care, as their exposure to AI can carry significant financial and ethical risks. On the other hand, investors operate in a landscape of informed risk-taking, perhaps requiring less intensive AI regulation. What is needed is to design a proportional framework that safeguards the most vulnerable while allowing flexibility for innovation in lower-risk applications.

4. **Future directions:** The study revealed the need for a dynamic document encapsulating these insights, with continued refinement based on evolving requirements.

The fourth session with financial policymakers and regulators

The fourth and final session was conducted with financial industry policymakers and regulators. We presented insights gleaned from the prior three sessions, which had surfaced the main concerns of bankers across a broad spectrum of subject matter expertise.

This session would allow us to put bankers' opinions and research thematics to a group of regulators with the freedom for them to express their views without being held to account. It included a representative from the National Treasury, the South African Reserve Bank, and the Financial Sector Conduct Authority. Each of these entities was deeply involved in AI discussions of their own.

Participants discussed the adequacy of current regulatory frameworks, the need for transparency, and the importance of balancing innovation with consumer protection. Here are some key highlights:

Key topics and insights

1. **Interconnected risks in AI:** Concerns were raised about biases in AI models due to data limitations, emphasising the need for transparency in AI-driven decisions. Additional risks discussed included cybersecurity vulnerabilities, especially adversarial attacks, and the operational risks arising from AI systems' complexity. Concentration risk due to reliance on major cloud providers was also noted.

2. **AI and cyber risk management:** Given their interconnectedness, it was agreed that AI risks should be viewed alongside cyber risks. However, it was noted that AI risks are multifaceted and may extend beyond current cyber risk frameworks.

3. **Regulatory evolution and ethical considerations:** The group discussed the need for regulatory adaptation to address AI's ethical challenges. A gap analysis was recommended to evaluate the adequacy of current regulations and propose new standards where needed, particularly to address AI's unique risks.

4. **Balancing innovation and regulation:** Participants emphasised the importance of balancing innovation and implementing regulatory safeguards. Other regions where minimal regulatory intervention has allowed FinTech innovations to thrive were noted. The need for a proactive versus reactive regulatory stance was debated, with recognition of AI's rapid development.

5. **Impact of AI on financial institutions:** Concerns were raised about AI's effect on ethical standards within the financial sector and the associated reputational risks, particularly around perceived biases in lending practices. Ongoing discussions at the international level to address financial stability risks associated with AI were highlighted.

6. **Explainability and ethical AI:** Participants emphasised the need for explainability in AI processes, with one attendee questioning how to make AI-driven decisions understandable to the public. Regular internal and external audits were recommended to support ethical AI adoption and compliance.

7. **Audience understanding and public education:** Several attendees highlighted the importance of public education and financial literacy, noting that clear communication tailored to different audiences is essential. Suggestions were made for a basic financial literacy initiative to empower citizens with knowledge for interacting with AI-driven banking services.

8. **Public engagement and consumer protection:** The group acknowledged the need for a consumer-focused approach to AI in banking. Concerns were raised about the

risks of one-size-fits-all educational strategies, and a call for targeted approaches to serve diverse demographics was considered better.

9. **Challenges of exclusivity and concentration risk:** Concerns were voiced regarding concentration risk in the market, particularly the over-reliance on specific providers, which could impact operational stability.

10. **Compliance with the EU AI Act:** Compliance with international AI regulations, such as the EU AI Act, was discussed, with concerns about its implications for local banks and the need for clarity in the roles of AI providers and deployers.

Align with best practices.

How can South Africa's AI frameworks ensure both local relevance and global competitiveness? Aligning with international standards like GDPR allows for integration into global markets. Yet, local nuances must not be overlooked. That's why South Africa requires tailored applications that address issues like financial inclusion and equity. Principles-based regulation offers the agility to bridge these needs, focusing on AI functionality rather than the broader, harder-to-regulate field of ethics.

Assume ethics is more aligned to culture, a discipline that guides behaviour rather than a codified set of rules. In that case, regulation should complement ethical considerations without attempting to govern them outright.

Again, the challenge is to foster a regulatory environment that allows ethical AI use without stifling innovation or creating rigid frameworks ill-suited to promoting a dynamic and global economy.

11. **Trust and innovation in banking:** Attendees discussed the challenge of trust in traditional banking versus FinTech, noting that differing regulatory expectations may create disparities in innovation capacities.

Building public trust in AI through regulatory oversight

The public needs to understand AI's capabilities to foster trust and adoption while addressing fears of bias and job loss. However, this may be outside the remit of an individual bank whose job is not to educate the public on AI. The best trust model may be to place responsibility on e.g. the banking regulator to set the oversight standards. In a highly regulated industry like banking, where oversight is already robust, a regulator-led model could set clear standards for transparency, bias auditing, and fairness, fostering trust without overburdening banks with public education initiatives. The public understanding that oversight is being provided by either the Prudential Authority or Financial Sector Conduct Authority should provide comfort and maintain the trust of the consumer.

12. **Human oversight in AI decisions:** The need for human oversight in AI-driven credit decisions was highlighted, with varying opinions on the extent of automation suitable for smaller-scale applications. Emphasis was placed on accountability and the importance of senior management's role in overseeing AI processes.

Did bankers and policymakers agree on the same issues?

The meeting underscored the complexities of integrating AI in the banking sector and highlighted the need for updated regulatory frameworks, effective public education, and ongoing collaboration. Ethical considerations, transparency, and trust-building were key themes, with agreement on the importance of balancing innovation with consumer protection, and a commitment to further discussions and comprehensive analyses to inform AI-related regulatory standards.

Although the focus of regulators was different in certain areas to that of the banks, there was a common thread. The willingness of regulators to continue engaging in dialogue with the industry will ensure that gaps and unintended consequences will be ironed out before legislation is promulgated.

Promoting change through regulation.

Regulation by its very nature is designed to change behaviour and the regulator monitors compliance with the regulation. If we were to develop facial recognition post a regulatory framework for AI, will this nascent technology be tolerated by regulators, with clear prejudice towards black people, or will the implementing bank be sanctioned for a discriminatory practice?

Final recommendations

Integrating AI into South Africa's banking sector represents both a significant opportunity and a profound challenge. As this working paper has shown, AI has the potential to revolutionize financial services, offering greater efficiency, enhanced decision-making, and new avenues for innovation. However, alongside these benefits lies a series of ethical, regulatory, and operational concerns. Issues such as data privacy, transparency, fairness, and the risk of algorithmic bias require immediate attention to ensure that AI is implemented to uphold trust and equity.

Our research journey with banking professionals, regulators, and diverse stakeholders has shown the importance of crafting a balanced approach to AI governance. By addressing these challenges thoughtfully, South African banks can position themselves as leaders in ethical AI adoption, setting standards that resonate locally and globally. The following recommendations aim to bridge the gap between innovation and ethical oversight in this complex landscape:

1. Strengthen transparency and explainability

Implement measures to enhance AI transparency, especially in high-stakes decision-making processes. This includes developing explainable AI systems that enable customers and regulators to understand the rationale behind automated decisions. Regular audits and customer disclosures on AI's role in banking interactions should be standard practice.

2. Enhance data privacy and security

Prioritise data privacy by aligning AI systems with the POPIA and other relevant regulations. Comprehensive data governance policies should be adopted,

focusing on secure data handling, transparency in data sharing, and ensuring customer information is protected from misuse and breaches.

3. Mitigate bias and ensure fairness

Develop frameworks to reduce bias in AI algorithms by ensuring training data is representative and inclusive of South Africa's diverse demographics. Regular bias audits and implementing fairness-centric AI design principles are recommended to mitigate potential discrimination in AI-driven decisions.

Positive bias.

When designing product solutions e.g. a cash economy, AI may need to use appropriate data sets not blended with data from a sophisticated digital economy.

4. Adopt a risk-based governance framework

Establish a risk-based approach to AI governance tailored to the banking sector's specific applications and ethical considerations. Risk management and internal governance structures should include ethics review panels, risk assessments, and compliance checks that support ethical AI deployment while safeguarding institutional and customer interests.

5. Promote workforce reskilling and public education

Address the potential impact of AI on employment by investing in workforce reskilling programs that equip employees with new skills relevant to AI-enhanced banking operations. Additionally, support public AI literacy initiatives to enable customers to confidently understand and engage with AI-driven services.

Invest in workforce reskilling and embed ethics in AI design.

As AI grows, workforce reskilling and education are already happening. But is reskilling enough to address the ethical challenges introduced by AI? There is an argument that the focus must extend beyond technical skills to include ethical programming and design. This is particularly important for developers creating systems that impact societal equity and fairness. It is simply too much to expect that workforce education can include the principles needed to navigate ethical dilemmas in AI deployment. If our approach is that ethical programming should be a cornerstone of AI development. In that case, this surely highlights the need for specialized expertise in every aspect of embedding ethics into every layer of design and implementation. How do you achieve this approach?

6. Foster public engagement and consumer protection

Strengthen consumer protection through targeted public education and financial literacy initiatives. Engage diverse demographic groups to ensure AI in banking is accessible and understood across customer segments. Emphasise ethical considerations in customer interactions to reinforce trust in AI-driven services.

7. Advocate for dynamic and adaptive regulation

Collaborate with regulators to support a flexible regulatory framework that evolves alongside AI advancements. This approach should balance innovation and regulatory compliance, ensuring ethical AI use without stifling growth. The framework can draw inspiration from global standards like the EU AI Act while contextually adapting to South African needs.

These recommendations aim to provide South African banks with a comprehensive framework for ethical AI adoption, positioning them to responsibly harness the potential of AI while upholding trust, accountability, and inclusivity in the financial sector.

Key findings

Based on the literature review, the survey, and the Delphi study, the following findings can be presented:

Common ethical issues associated with AI across different sectors

Ethical issues include transparency, accountability, fairness, data privacy, and bias. Stakeholders expressed concerns that unregulated AI may lead to discriminatory practices, lack of transparency in decision-making, and heightened data privacy risks. These issues are especially pertinent in fields that rely heavily on consumer trust, such as finance and healthcare.

Stakeholder perception of the ethical implications of AI

Stakeholders, including AI developers, regulators, and customers, view ethical AI as essential for maintaining trust and accountability. Many stakeholders advocate for explainable AI to demystify AI-driven decisions. There is a clear demand for AI systems that respect user privacy, ensure fairness, and remain transparent about data handling and decision-making processes.

Ethical challenges with using AI in South Africa's banking operations

Ethical challenges in South African banking include data privacy concerns, particularly compliance with POPIA, and the risk of bias in AI algorithms that could unfairly disadvantage certain demographics. There are also challenges related to transparency in AI-driven financial decisions and the need to mitigate AI's potential to exacerbate financial exclusion.

How AI in South African banks compares to transparency, accountability, and fairness methods

AI applications introduce a “black box” complexity that traditional methods lack, making it harder for customers and regulators to scrutinise decisions. Traditional methods allow more human oversight and more straightforward transparency, while AI systems require advanced measures, such as explainable AI, to maintain comparable levels of transparency and accountability. AI can potentially improve efficiency and personalisation but demands stricter frameworks to ensure these benefits do not come at the expense of fairness and inclusivity.

Conclusion

Integrating AI in the financial services industry is not a recent concept but a present reality. It will reshape how banks operate, innovate, and interact with stakeholders. This working paper shows AI's vast potential to enhance efficiency, mitigate risks, and drive customer satisfaction. However, it is just as clear that this power comes with profound ethical and operational challenges that demand immediate attention. Confronting issues of fairness, transparency, data privacy, and algorithmic bias are central to maintaining trust in the banking sector.

The future of AI in banking lies in its responsible adoption, guided by frameworks that balance innovation with ethical integrity. Industry professionals must collaborate to develop adaptive regulatory measures that address local contexts while aligning with global standards. This convergence of innovation and responsibility can position the banking sector as a leader in ethical AI implementation. With a commitment to transparency and fairness, the financial services sector has the opportunity to redefine its own practices in this brave new age.

Bibliography

Adedoyin Tolulope Oyewole, Bisola Beatrice Oguejiofor, Nkechi Emmanuella Eneh, Chidiogo Uzoamaka Akpuokwe, & Seun Solomon Bakare. (2024). DATA PRIVACY LAWS AND THEIR IMPACT ON FINANCIAL TECHNOLOGY COMPANIES: A REVIEW. *Computer Science & IT Research Journal*, *5*(3), 628–650. https://doi.org/10.51594/csitrj.v5i3.911

Ade-Ibijola, A., & Okonkwo, C. (2023). *Artificial Intelligence in Africa: Emerging Challenges*. https://doi.org/10.1007/978-3-031-08215-3_5

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Basel Committee on Banking Supervision. (n.d.). *Digitalisation of finance*. Retrieved July 23, 2024, from https://www.bis.org/bcbs/publ/d575.pdf

Bennett, M. T., & Maruyama, Y. (2022). Philosophical Specification of Empathetic Ethical Artificial Intelligence. *IEEE Transactions on Cognitive and Developmental Systems*, *14*(2). https://doi.org/10.1109/TCDS.2021.3099945

Big-four banks take lead in SA's GenAI, AI deployments | ITWeb. (n.d.). Retrieved September 3, 2024, from https://www.itweb.co.za/article/big-four-banks-take-lead-in-sas-genai-ai-deployments/G98YdqLGK9pMX2PD

Birhane, A., Ruane, E., Laurent, T., S. Brown, M., Flowers, J., Ventresque, A., & L. Dancy, C. (2022). The Forgotten Margins of AI Ethics. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3531146.3533157

Booth, A. (2006). Clear and present questions: formulating questions for evidence based practice. *Library Hi Tech*, *24*(3), 355–368. https://doi.org/10.1108/07378830610692127

Borokini, F., Wakunuma, K., & Akintoye, S. (2023). *The Use of Gendered Chatbots in Nigeria: Critical Perspectives*. https://doi.org/10.1007/978-3-031-08215-3_6

Buckley, R. P., Zetzsche, D. A., Arner, D. W., & Tang, B. W. (2021). Regulating Artificial Intelligence in Finance: Putting the Human in the Loop. *Sydney Law Review*, *43*(1).

Cavus, N., Mohammed, Y. B., & Yakubu, M. N. (2021). An artificial intelligence-based model for prediction of parameters affecting sustainable growth of mobile banking apps. *Sustainability (Switzerland)*, *13*(11). https://doi.org/10.3390/su13116206

Cihon, P., Kleinaltenkamp, M. J., Schuett, J., & Baum, S. D. (2021). AI Certification: Advancing Ethical Practice by Reducing Information Asymmetries. *IEEE Transactions on Technology and Society*, *2*(4), 200–209. https://doi.org/10.1109/tts.2021.3077595

Crockett, K., Colyer, E., Gerber, L., & Latham, A. (2023). Building Trustworthy AI Solutions: A Case for Practical Solutions for Small Businesses. *IEEE Transactions on Artificial Intelligence*, *4*(4). https://doi.org/10.1109/TAI.2021.3137091

Dancy, C. L., & Saucier, P. K. (2021). AI and Blackness: Toward Moving Beyond Bias and Representation. *IEEE Transactions on Technology and Society*, *3*(1). https://doi.org/10.1109/tts.2021.3125998

Department of Homeland Security. (2024). *Safety and Security Guidelines for Critical Infrastructure Owners and Operators*.

Dignum, V. (2023). *Responsible Artificial Intelligence: Recommendations and Lessons Learned*. https://doi.org/10.1007/978-3-031-08215-3_9

Doshi-Velez, F., & Kim, B. (2017). *Towards A Rigorous Science of Interpretable Machine Learning*. https://arxiv.org/abs/1702.08608v2

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2019). Re-examining the Unified Theory of Acceptance and Use of Technology (UTAUT): Towards a Revised Theoretical Model. *Information Systems Frontiers*, *21*(3), 719–734. https://doi.org/10.1007/S10796-017-9774-Y

Eke, D. O., Chintu, S. S., & Wakunuma, K. (2023). *Towards Shaping the Future of Responsible AI in Africa*. https://doi.org/10.1007/978-3-031-08215-3_8

Eke, D. O., Wakunuma, K., & Akintoye, S. (2023). *Introducing Responsible AI in Africa*. https://doi.org/10.1007/978-3-031-08215-3_1

Exploring artificial intelligence in the South African banking industry VP Darangwa. (n.d.).

Fares, O. H., Butt, I., & Lee, S. H. M. (2023). Utilization of artificial intelligence in the banking sector: a systematic literature review. *Journal of Financial Services Marketing*, *28*(4), 835–852. https://doi.org/10.1057/S41264-022-00176-7/TABLES/4

Gehl Sampath, P. (2021). Governing Artificial Intelligence in an Age of Inequality. *Global Policy*, *12*(S6). https://doi.org/10.1111/1758-5899.12940

Golda, A., Mekonen, K., Pandey, A., Singh, A., Hassija, V., Chamola, V., & Sikdar, B. (2024). Privacy and Security Concerns in Generative AI: A Comprehensive Survey. *IEEE Access*, *12*, 48126–48144. https://doi.org/10.1109/ACCESS.2024.3381611

Green, B. (2021). The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice. *Journal of Social Computing*, *2*(3). https://doi.org/10.23919/JSC.2021.0018

Green, B., Heider, D., Firth-Butterfield, K., & Lim, D. (2021). *Responsible Use of Technology: The IBM Case Study*. https://www3.weforum.org/docs/WEF_Responsible_Use_of_Technology_The_IBM_Case_Study_2021.pdf

Gwagwa, A., Kraemer-Mbula, E., Rizk, N., Rutenberg, I., & De Beer, J. (2020). Artificial Intelligence (AI) Deployments in Africa: Benefits, Challenges and Policy Dimensions. *The African Journal of Information and Communication*, *26*. https://doi.org/10.23962/10539/30361

Hagendorff, T. (2022). Blind spots in AI ethics. *AI and Ethics*, *2*(4). https://doi.org/10.1007/s43681-021-00122-8

Hassan, M., Aziz, L. A.-R., & Andriansyah, Y. (2023). The Role Artificial Intelligence in Modern Banking: An Exploration of AI-Driven Approaches for Enhanced Fraud Prevention, Risk Management, and Regulatory Compliance. *Reviews of Contemporary Business Analytics*, *6*(1).

Hassani, B. K. (2021). Societal bias reinforcement through machine learning: a credit scoring perspective. *AI and Ethics*, *1*(3). https://doi.org/10.1007/s43681-020-00026-z

He, H., Gray, J., Cangelosi, A., Meng, Q., McGinnity, T. M., & Mehnen, J. (2022). The Challenges and Opportunities of Human-Centered AI for Trustworthy Robots and Autonomous Systems. *IEEE Transactions on Cognitive and Developmental Systems*, *14*(4). https://doi.org/10.1109/TCDS.2021.3132282

Hentzen, J. K., Hoffmann, A., Dolan, R., & Pala, E. (2022). Artificial intelligence in customer-facing financial services: a systematic literature review and agenda for future research. *International Journal of Bank Marketing*, *40*(6). https://doi.org/10.1108/IJBM-09-2021-0417

How, M. L., Cheah, S. M., Khor, A. C., & Chan, Y. J. (2020). Artificial intelligence-enhanced predictive insights for advancing financial inclusion: A human-centric ai-thinking approach. *Big Data and Cognitive Computing*, *4*(2). https://doi.org/10.3390/bdcc4020008

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2023). An Overview of Artificial Intelligence Ethics. *IEEE Transactions on Artificial Intelligence*, *4*(4). https://doi.org/10.1109/TAI.2022.3194503

International Organisation of Securities Commissions. (2020). *The use of artificial intelligence and machine learning by market intermediaries and asset managers Consultation*. www.iosco.org

Jiang, L., Hwang, J. D., Bhagavatula, C., Bras, R. Le, Liang, J., Dodge, J., Sakaguchi, K., Forbes, M., Borchardt, J., Gabriel, S., Tsvetkov, Y., Etzioni, O., Sap, M., Rini, R., & Choi, Y. (2021). *Can Machines Learn Morality? The Delphi Experiment*. http://arxiv.org/abs/2110.07574

Keeping pace with global innovations by propelling SA's banking, insurance sectors | ITWeb. (n.d.). Retrieved September 3, 2024, from https://www.itweb.co.za/article/keeping-pace-with-global-innovations-by-propelling-sas-banking-insurance-sectors/KA3Wwqdzmxj7rydZ

Kelley, S. (2022). Employee Perceptions of the Effective Adoption of AI Principles. *Journal of Business Ethics*, *178*(4). https://doi.org/10.1007/s10551-022-05051-y

Kshetri, N. (2020). Artificial Intelligence in Developing Countries. *IT Professional*, *22*(4). https://doi.org/10.1109/MITP.2019.2951851

Kwanya, T. (2023). *Working with Robots as Colleagues: Kenyan Perspectives of Ethical Concerns on Possible Integration of Co-bots in Workplaces*. https://doi.org/10.1007/978-3-031-08215-3_4

Marchant, G. E. (2021). Professional Societies as Adopters and Enforcers of AI Soft Law. *IEEE Transactions on Technology and Society*, *2*(4). https://doi.org/10.1109/tts.2021.3116524

Martin, J. L. (2023). The Ethico-Political Universe of ChatGPT. *Journal of Social Computing*, *4*(1). https://doi.org/10.23919/JSC.2023.0003

Mashishi, A. (2023). *National AI Government Summit Discussion Document*. https://www.dcdt.gov.za/images/phocadownload/AI_Government_Summit/National_AI_Government_Summit_Discussion_Document.pdf

Matsepe, N. T., & Van der Lingen, E. (2022). Determinants of emerging technologies adoption in the South African financial sector. *South African Journal of Business Management*, *53*(1). https://doi.org/10.4102/sajbm.v53i1.2493

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, *27*(4), 12–12. https://doi.org/10.1609/AIMAG.V27I4.1904

Mhlanga, D. (2020). Industry 4.0 in finance: the impact of artificial intelligence (ai) on digital financial inclusion. *International Journal of Financial Studies*, *8*(3). https://doi.org/10.3390/ijfs8030045

Mhlanga, D. (2021). Financial inclusion in emerging economies: The application of machine learning and artificial intelligence in credit risk assessment. *International Journal of Financial Studies*, *9*(3). https://doi.org/10.3390/ijfs9030039

Minkkinen, M., & Mäntymäki, M. (2023). Discerning Between the "Easy" and "Hard" Problems of AI Governance. *IEEE Transactions on Technology and Society*, *4*(2). https://doi.org/10.1109/tts.2023.3267382

Mogaji, E., Farquhar, J. D., van Esch, P., Durodie, C., & Perez-Vega, R. (2022). AI in financial services marketing. *International Journal of Bank Marketing*, *40*(6), 1097–1101. https://doi.org/10.1108/IJBM-09-2022-617

Mogaji, E., & Nguyen, N. P. (2022). Managers' understanding of artificial intelligence in relation to marketing financial services: insights from a cross-country study. *International Journal of Bank Marketing*, *40*(6). https://doi.org/10.1108/IJBM-09-2021-0440

Monaro, M., Barakova, E., & Navarin, N. (2022). Editorial Special Issue Interaction With Artificial Intelligence Systems: New Human-Centered Perspectives and Challenges. In *IEEE Transactions on Human-Machine Systems* (Vol. 52, Issue 3). https://doi.org/10.1109/THMS.2022.3172516

Mullins, M., Holland, C. P., & Cunneen, M. (2021). Creating ethics guidelines for artificial intelligence and big data analytics customers: The case of the consumer European insurance market. *Patterns*, *2*(10). https://doi.org/10.1016/j.patter.2021.100362

Myers-West, S., Whittaker, M., & Crawford, K. (2019). Discriminating systems: Gender, Race, and Power in AI. In *AI Now Institute* (Issue April).

National Cyber Security Centre. (2024). *Deploying AI Systems Securely*. https://www.ic3.gov/Media/News/2024/240415.pdf

Ngobeni, D. A., Saurombe, M. D., & Joseph, R. M. (2022). The influence of the psychological contract on employee engagement in a South African bank. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.958127

Nilsson, N. J. (n.d.). *The Quest for Artificial Intelligence A History of Ideas and Achievements-The Quest for Artificial Intelligence: A History of Ideas and Achievements Nils J. Nilsson Frontmatter More information*. Retrieved September 1, 2024, from www.cambridge.org/9780521122931

NIST. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. https://doi.org/https://doi.org/10.6028/NIST.AI.100-1

NIST. (2024). *AI RMF PLAYBOOK*. https://airc.nist.gov/AI_RMF_Knowledge_Base/Playbook

Okolo, C. T., Aruleba, K., & Obaido, G. (2023). *Responsible AI in Africa—Challenges and Opportunities*. https://doi.org/10.1007/978-3-031-08215-3_3

Oluwatoyin Ajoke Farayola. (2024). REVOLUTIONIZING BANKING SECURITY: INTEGRATING ARTIFICIAL INTELLIGENCE, BLOCKCHAIN, AND BUSINESS INTELLIGENCE FOR ENHANCED CYBERSECURITY. *Finance & Accounting Research Journal*, *6*(4), 501–514. https://doi.org/10.51594/farj.v6i4.990

Oseremi Onesi-Ozigagun, Yinka James Ololade, Nsisong Louis Eyo-Udo, & Damilola Oluwaseun Ogundipe. (2024). AI-driven biometrics for secure fintech: Pioneering safety and trust. *International Journal of Engineering Research Updates*, *6*(2), 001–012. https://doi.org/10.53430/ijeru.2024.6.2.0023

Petkovic, D. (2023). It is Not "Accuracy vs. Explainability"—We Need Both for Trustworthy AI Systems. *IEEE Transactions on Technology and Society*, *4*(1). https://doi.org/10.1109/tts.2023.3239921

PRISMA statement. (n.d.). Retrieved August 3, 2024, from https://www.prisma-statement.org/

Privacy and responsible AI | IAPP. (n.d.). Retrieved September 3, 2024, from https://iapp.org/news/a/privacy-and-responsible-ai

Rizinski, M., Peshov, H., Mishev, K., Chitkushev, L. T., Vodenska, I., & Trajanov, D. (2022). Ethically Responsible Machine Learning in Fintech. *IEEE Access*, *10*. https://doi.org/10.1109/ACCESS.2022.3202889

Roche, C., Wall, P. J., & Lewis, D. (2023). Ethics and diversity in artificial intelligence policies, strategies and initiatives. *AI and Ethics*, *3*(4). https://doi.org/10.1007/s43681-022-00218-9

Ruttkamp-Bloem, E. (2023). *Epistemic Just and Dynamic AI Ethics in Africa*. https://doi.org/10.1007/978-3-031-08215-3_2

Ryan, S., Nadal, C., & Doherty, G. (2023). Integrating Fairness in the Software Design Process: An Interview Study With HCI and ML Experts. *IEEE Access*, *11*. https://doi.org/10.1109/ACCESS.2023.3260639

SA Banks and Generative AI | EY - South Africa. (n.d.). Retrieved September 3, 2024, from https://www.ey.com/en_za/industries/financial-services/sa-banks-and-generative-ai

Schiff, D., Borenstein, J., Biddle, J., & Laas, K. (2021). AI Ethics in the Public, Private, and NGO Sectors: A Review of a Global Document Collection. *IEEE Transactions on Technology and Society*, *2*(1). https://doi.org/10.1109/tts.2021.3052127

Sidley, S. B. (2024). *AI-what happens when there is nothing left to learn? l'D (Image: ideogram.ai)*. https://www.dailymaverick.eo.za/author/stevenboykeysidley/

Singh, A. (2023). *AI 42001:2023 AIMS*.

Stahl, B. C., Leach, T., Oyeniji, O., & Ogoh, G. (2023). *AI Policy as a Response to AI Ethics? Addressing Ethical Issues in the Development of AI Policies in North Africa*. https://doi.org/10.1007/978-3-031-08215-3_7

Stettinger, G., Weissensteiner, P., & Khastgir, S. (2024). Trustworthiness Assurance Assessment for High-Risk AI-Based Systems. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3364387

Svetlova, E. (2022). AI ethics and systemic risks in finance. *AI and Ethics*, *2*(4). https://doi.org/10.1007/s43681-021-00129-1

The Ethical Decision-Making Model. (n.d.).

The Ethics of AI In Business and Finance | StreetFins®. (n.d.). Retrieved September 3, 2024, from https://streetfins.com/the-ethics-of-ai-in-business-and-finance/

The processes of technological innovation : Tornatzky, Louis G : Free Download, Borrow, and Streaming : Internet Archive. (n.d.). Retrieved August 3, 2024, from https://archive.org/details/processesoftechn0000torn

The White House. (2022). Blueprint for an AI Bill of Rights. *White House*, *October*.

Tidjon, L. N., & Khomh, F. (2023). The Different Faces of AI Ethics Across the World: A Principle-to-Practice Gap Analysis. *IEEE Transactions on Artificial Intelligence*, *4*(4). https://doi.org/10.1109/TAI.2022.3225132

Tolulope Esther Edunjobi, & Opeyemi Abayomi Odejide. (2024). Theoretical frameworks in AI for credit risk assessment: Towards banking efficiency and accuracy. *International Journal of Scientific Research Updates*, *7*(1), 092–102. https://doi.org/10.53430/ijsru.2024.7.1.0030

Truby, J., Brown, R., & Dahdal, A. (2020). Banking on AI: mandating a proactive approach to AI regulation in the financial sector. *Law and Financial Markets Review*, *14*(2). https://doi.org/10.1080/17521440.2020.1760454

Venkatesh, V., & Davis, F. D. (2000). Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/MNSC.46.2.186.11926

World Economic Forum. (2019). *AI Governance A Holistic Approach to Implement Ethics into AI*. https://www.weforum.org/publications/ai-governance-a-holistic-approach-to-implement-ethics-into-ai/

Xivuri, K., & Twinomurinzi, H. (2021). A Systematic Review of Fairness in Artificial Intelligence Algorithms. *Lecture Notes in Computer Science (Including Subseries Lecture*

Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), *12896 LNCS*. https://doi.org/10.1007/978-3-030-85447-8_24

Yu, S., Carroll, F., & Bentley, B. L. (2024). Insights Into Privacy Protection Research in AI. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3378126

Zhang, J., Shu, Y., & Yu, H. (2023). Fairness in Design: A Framework for Facilitating Ethical Artificial Intelligence Designs. *International Journal of Crowd Science*, *7*(1). https://doi.org/10.26599/IJCS.2022.9100033

Zicari, R. V., Brodersen, J., Brusseau, J., Dudder, B., Eichhorn, T., Ivanov, T., Kararigas, G., Kringen, P., McCullough, M., Moslein, F., Mushtaq, N., Roig, G., Sturtz, N., Tolle, K., Tithi, J. J., van Halem, I., & Westerlund, M. (2021). Z-Inspection ®: A Process to Assess Trustworthy AI . *IEEE Transactions on Technology and Society*, *2*(2). https://doi.org/10.1109/tts.2021.3066209