

Title: Embedding AI in Investment Banking

Author: Sarvesh Aditya Murali Desan
Intern, Group Business Ethics & Integrity
Kenanga Investment Bank Berhad

Executive Summary

Artificial Intelligence (AI) is revolutionising financial services. It enhances customer insights, automates decisions, and improves compliance monitoring. But with these gains come serious ethical risks—bias, data misuse, lack of explainability—that must be addressed. At Kenanga, where AI is being piloted in credit scoring and financial modelling, we must embed ethical governance at the core of innovation.

This whitepaper explores how ethical frameworks can guide responsible AI development and proposes the CLEAR Framework—Consent, Limitations, Equity, Accountability, and Regulation—to align AI with trust, transparency, and compliance. It integrates key insights from the Group Ethics Blueprint, the 3-Year Ethics Implementation Plan, and international benchmarks (e.g. DBS, HSBC, IBM). The paper outlines GBEI's role in AI ethics oversight and provides practical tools to cascade responsible AI across all functions.

Kenanga is well-positioned to lead ASEAN's financial sector in responsible AI, leveraging our strong ethical foundation: the AFBC policy, Vendor Code of Conduct, Whistleblowing Framework, and ethics training infrastructure. Aligning AI with the bank's ethics roadmap is not only good governance—it is essential for resilience, inclusion, and long-term success.

1. Introduction: Why Ethical AI Matters

AI is no longer an experiment—it is part of the financial system. Kenanga is using AI in credit scoring, SME lending, customer analytics and fraud detection. While AI improves performance, it introduces new risks: algorithmic bias, opaque logic, and overreliance on machines. Regulators like BNM, the EU, and MAS are increasing scrutiny.

Ethical AI means designing, deploying and monitoring AI in a way that is fair, accountable, and transparent. For financial institutions, this protects customer rights, avoids discrimination, and builds sustainable trust. For Kenanga, it aligns with our Ethics Blueprint and reinforces our long-standing commitment to integrity and governance.

2. The Role of Ethical Frameworks in Guiding AI

Ethical frameworks offer structure when laws lag behind. At Kenanga, they help:

- Meet Section 17A MACC Act obligations by defining clear accountability.
- Embed trust and integrity into digital systems.

- Use grievance and whistleblowing channels to flag AI misuse.
- Audit high-risk models and ensure human oversight.

The CLEAR Framework (next section) adapts best practices from benchmarks and global case studies to KIBB's current approach and 3-year implementation roadmap. It serves as a practical AI ethics guide tailored to Kenanga's needs.

3. The CLEAR Framework: A Governance Model for Ethical AI

Principle	Definition	Benchmark Institution	Application at Kenanga
Consent	Stakeholders are aware when AI is used	DBS (discloses AI use in customer interactions, part of Human-Centric AI strategy)	Client disclosure for robo-advisory, chatbot transparency, vendor policy alignment
Limitations	Boundaries of AI use are clear	Maybank (via RMiT controls) and Standard Chartered (AI model risk frameworks)	HITL (human-in-the-loop) procedures for credit scoring, disclaimers on AI tools. Helps ensure there is also a bit of human intervention to prevent full AI override
Equity	Models are tested and corrected for bias	Hong Leong (public fairness commitment) and DBS (model review guidelines)	Fairness audits in lending & recruitment (e.g. SHAP* tools—see note), inclusive training data
Accountability	Clear ownership and escalation channels	HSBC (AI Governance Board assigns responsibility post-deployment)	Model registries, AFBC review process, GBEI oversight

Regulation	Compliance with law and ethical principles	Maybank (RMiT integration), Standard Chartered (Ethics risk taxonomy), EU AI Act	BNM RMiT, MACC Act, EU AI Act alignment, quarterly ethics reviews
-------------------	--	--	---

*SHAP (SHapley Additive exPlanations) is a machine learning explainability tool. It identifies how much each input contributed to an AI model's decision—key for bias audits and fairness reviews.

The CLEAR framework connects seamlessly with Kenanga's existing ethics pillars—Organisational, Operational, Technical, and Reputational—and provides tools to operationalise AI ethics (see Appendix A).

4. GBEI's Role in Responsible AI

As custodian of ethics at Kenanga, GBEI can lead responsible AI by:

- Including AI risks in AFBC reviews
- Expanding conflict-of-interest reviews to include algorithmic bias
- Using grievance and whistleblowing frameworks for AI harm escalation
- Integrating AI ethics in onboarding, vendor management, and staff training
- Tracking KPIs: % explainable models, flagged grievances, model fairness scores

These actions are aligned with the 3-Year Ethics Implementation Plan:

- **Year 1:** Identify AI-related ethical risks and build foundations
- **Year 2:** Strengthen controls and introduce AI audits
- **Year 3:** Evaluate impact and refine training for emerging risks

5. Cascading AI Ethics Across Kenanga

Ethical AI must be a shared responsibility. CLEAR can be cascaded as follows:

Function	Role
Business Units	Apply CLEAR to AI products; escalate concerns using grievance process

IT & Data Science	Document models, conduct bias testing, ensure explainability
Compliance & Risk	Incorporate AI ethics into RCSA reviews, quarterly reports
HR	Review AI used in hiring or performance tools
Vendor Management	Enforce vendor AI compliance using the Vendor Code of Conduct

These efforts should be centralised under an **AI Ethics Registry**, coordinated by GBEI and updated quarterly.

6. What Global Leaders Are Doing

Kenanga can benchmark itself against:

- **DBS**: AI Ethics Council, fairness reviews in lending
- **Maybank**: Ethics integrated into Risk Management in Technology (RMIT) and credit scoring
- **Hong Leong Bank**: Annual ethics statements and audits
- **HSBC**: Global AI Governance Board for explainability and accountability
- **Standard Chartered**: AI ethics risk assessments and internal policies
- **IBM & Novartis**: Dynamic AI ethics protocols updated with technology

Kenanga can lead by combining these best practices with its ethics infrastructure.

7. Challenges & Strategic Next Steps

Challenges:

- Black-box models reduce transparency (Burrell, 2023)
- Training data can encode social bias (Bender et al., 2023)
- Legal frameworks lag behind AI growth
- Pressure to innovate may outpace governance

Action Plan:

- Launch a dedicated AI Ethics Council within GBEI
 - Start a cross-functional AI Landscape Assessment
 - Identify and document key AI use cases (e.g. SME credit scoring)
 - Conduct staff interviews to assess awareness
 - Develop 2 internal case studies on AI implementation and risks
 - Design a Responsible AI Framework tailored to Kenanga
 - Pilot test the framework in one unit and collect feedback
-

8. Conclusion: Ethics as a Strategic Advantage

Responsible AI is not just about compliance—it's a leadership imperative. Kenanga's strong ethics governance and 3-Year Implementation Plan give it a unique platform to lead on AI ethics in ASEAN.

Embedding AI governance into our ethical DNA will:

- Build trust with clients and regulators
- Minimise risk from bias and algorithmic harm
- Support innovation aligned with our values

With the CLEAR Framework and GBEI leadership, Kenanga can ensure AI becomes not just a tool of automation—but a force for accountability, fairness, and sustainable success.

Let us shape AI before it shapes us.

Appendix A: CLEAR Implementation Checklist

Principle	Key Questions	Team	Status
Consent	Have users been informed?	Compliance, Legal	Yes/No
Limitations	Are model limits disclosed?	IT, Data Science	Yes/No
Equity	Has fairness been tested	Data Science, HR	Yes/No
Accountability	Is ownership clear?	GBEI, Risk	Yes/No
Regulation	Are we compliant?	Legal, Audit	Yes/No

Appendix B: Key Ethics Policies & Alignment

Policy	How it Supports Ethical AI
AFBC	Prevents bribery & enables ethical audit trails
Conflict Management	Mitigates bias from relationships or personal gain
Whistleblowing	Enables AI harm reporting
Vendor Code of Conduct	Holds third-party AI vendors accountable
GEH Policy	Protects against influence in AI procurement
Group Code of Ethics	Sets standards for employee conduct in AI-related decisions
Sponsorship Policy	Prevents reputational or ethical misuse of AI-linked funding
Reporting Procedures	Tracks fraud or bias linked to AI deployment

Appendix C: Citations & Research

- Arendt, H. (1958). *The Human Condition*. University of Chicago Press.
- Andreessen, M. (2023). *The Techno-Optimist Manifesto*. Andreessen Horowitz.
- Bender, E. et al. (2023). “AI and Linguistic Bias.” *Journal of Machine Learning*.
- Burrell, J. (2023). “Algorithmic Opacity.” *Nature Digital Society*.
- Dastin, J. (2018). *Reuters* – AI bias in hiring.
- Harari, Y. N. (2015). *Homo Deus*. Harvill Secker.
- Suleyman, M. (2023). *The Coming Wave*. Crown Publishing.
- Romero, V. (2023). *LA Times* – Algorithmic misclassification.
- Stanford AI Lab (2023). Chatbot bias study.
- Cho, A. (2023). *Quanta Magazine* – AI failure case study.
- IBM & Novartis – Industry case studies (referenced, internal notes).

Appendix D: Summary of Benchmarked Institutions

Institution	Practice
DBS	Human-centric AI, Ethics Council
Maybank	Integrated AI reviews into RMiT
Hong Leong	Digital ethics statements & fairness testing
HSBC	AI Governance Board for explainability
Standard Chartered	Working group on AI fairness risk
IBM/Novartis	Live AI ethics dashboards and adaptive frameworks

Appendix E: Research Paper - How AI will shape our future (June 2025)

TITLE: AI: How intelligence forces us to ask who we are

"The real test isn't AI's power — it's our humanity."

For generations, science fiction imagined tomorrow. Today, artificial intelligence isn't fiction — it's fact, accelerating into our lives faster than any technology before it. But AI isn't just an invention — it's a turning point in how we define what it means to be human. While the printing press and internet changed how we lived, AI is transforming how we think, feel, and choose." As Mustafa Suleyman puts it, AI is no longer a tool — it's infrastructure. As it enters our courts, classrooms, clinics — it raises a deeper question: not "what can it do?" but "what will we become in response?"

AI's ability to learn, adapt, and make decisions gives it a hint of sentience. From chatbots to self-driving cars, AI is already an attempt to replicate or even surpass human intelligence using machines. But as machines grow more capable undertaking things like diagnosing illness, writing music and generating art, they confront us with uncomfortable questions. If a machine can think, where does that leave us? What makes us human? What defines a machine?

This is why the very act of creating AI must be interrogated from the outset. As AI reshapes everything from schools, hospitals, hiring practices, court and military systems, we need to be clear on how we manage what we create and manage ourselves once it begins to shape our world. Mary Shelley asked these same questions in her Gothic novel Frankenstein, which warned of the dangers of unchecked creation (Shelley, 1818). Is AI our new "Franken-Machines" built not from flesh, but from code? Powerful systems can learn, decide and act but remain morally unaccounted for (Matthias, A. 2004). In Harari's view, AI is actually imposing a new order on our lives (Harari, 2024). They already predict our moods and guide our choices. This predictive power may reduce freedom, in fact, render it an illusion. When Spotify plays the perfect song to match one's day, this notion of choice, is it with the person or the algorithm? So when a chatbot answers faster than a friend, it is becoming inevitable we turn to the option that is most accessible (Matthias, A., 2004).

The risks are not hypothetical and the warnings are apparent. As Hannah Arendt warned, "progress and catastrophe are two sides of the same coin," (Arendt, 1965). A 2023 report from the Stanford AI Lab found that chatbot responses often varied based on the perceived identity of the user (Stanford AI Lab., 2023). Empathy levels differed, especially in terms of bias in race and gender leaked into dialogue (Bender et al., 2023). In California, an automated welfare system misclassified many people as fit for work, wrongly cutting off their support (Romero, 2023). Similarly, Dastin J. (2018) highlighted how hiring tools filtered out underrepresented candidates using flawed data. These aren't rare bugs as they are reflections of societal bias amplified by code (Arendt, H., 1958).

By contrast, Marc Andreessen sees AI as a great enabler and even a cure-all. His manifesto argues that innovation will resolve more issues than it causes (Andreessen, 2023). There's truth in that hope. In 2023, Goldman Sachs estimated that AI could add \$7 trillion in productivity gains over the next decade (Silverman, 2023). DeepMind's AlphaMissense, developed by Jumper et al. (2023), may redefine how we diagnose genetic conditions. At the same time, science still has unexpected directions. As Adrian Cho (2023) reports in Quanta Magazine, even optimistic technologies like room-temperature superconductors are not immune to surprise reversals. Similarly, Nature Editorial Board (2023) was forced to withdraw a much-cited paper reminding us that even peer-reviewed progress is answerable. But seeing technology as saviour masks the need for caution. Progress must be balanced with preparedness. A rush to deploy often outpaces the laws, protections and values we need to govern it.

Ethical oversight often trails far behind capability. Machines influence decisions in policing, banking and healthcare, yet many operate as black boxes. Burrell (2023) describes this opacity problem. This is especially so when systems make decisions we cannot interpret and accountability disappears. In Estonia, digital courts use automated tools to guide judgements (Tammet, 2023). While the process gains efficiency, questions remain about justice without human nuance. When people are reduced to patterns, decisions lose empathy.

Cultural shifts often reveal more than statistics. In South Korea, Zen chatbots have been developed to challenge capitalist logic. These programs aren't just quirky experiments — they're reflections of what we value. Likewise, galleries now display AI-generated artworks that stir debate on authorship, soul, and intention (Chen, L. 2024). Can a system with no childhood, no grief, no memory of failure or joy, truly create? Art divorced from the messiness of life may impress, but can it mean? This author himself has been able to generate a deeper, data enriched understanding of the topic because he has benefited from research gleaned from AI.

There are limits that technology cannot cross. Therapy bots in Japan, for instance, can mimic attentive listening but cannot feel sorrow or joy. Machines don't experience grief. They cannot forgive. They cannot fall in love. They offer patterns but not presence. Human experience is not programmable because the elements that matter most in relationships, in healing, in moral judgement are completely intangible and remain out of reach.

Ultimately, power, not just progress, shapes how these tools are made and used. In the right hands, AI becomes a powerful tool, not a sword. This is because algorithms are not neutral, they carry assumptions about what matters and who counts. If their training reflects only dominant voices, the systems they create will exclude others. Linguistic bias, as highlighted by Bender et al. (2023), silences minority dialects and marginalised worldviews. This is narrowing of what counts as knowledge. AI is like an infinite onion. Each layer peeled away reveals new possibilities for progress. Yet at its core lies the architecture of what already exists, with all its inherited biases and power structures.

The environmental and political costs are just as real. Training large systems consumes water and electricity at vast scales, often in regions facing scarcity. Data extraction from users in the Global South, reveals how benefits and burdens are unequally shared (Roy, A. 2014). When just a handful of corporations control development, knowledge becomes a commodity, not a public good.

So where do we go from here? Transparency is only the beginning. Systems must be open to challenge. Decisions must be revisable. Brynjolfsson (2023) warns of assigning rights to machines before we've secured rights for those most impacted by them. Inclusion and restraint, not just innovation, should guide our next steps. Some

problems demand human care, not computation. And the most important decisions must remain in human hands.

AI is not a crystal ball. It is a mirror reflecting our choices and contradictions. The focus must be based on empathy, nuance and moral responsibility. We must ask not, "How intelligent can machines become?" but, "how courageous, wise and humane can we remain?" The answer lies with our own intentions. This technology will not shape the future if we don't shape it first. We must anchor AI in values that resist automation. We face a choice. Either forge tools that amplify humanity or surrender control to code.

APPENDIX A:

"In perfecting artificial intelligence, we risk perfecting our own intellectual surrender"
"An algorithm that can't explain itself is the modern equivalent of 'because I said so'."
"The danger isn't that AI will rebel, but that it will do exactly what we ask."

APPENDIX B:

- Andreessen, M. (2023). The Techno-Optimist Manifesto. Andreessen Horowitz.
- Arendt, H. (1958). *The Human Condition*. University of Chicago Press.
- Bender, E. M., et al. (2023). AI and Linguistic Bias. *Journal of Machine Learning*, 24(1), 511–530.
- Brynjolfsson, E. (2023). Machine Rights Debate. *Boston Review*, 48(2), 65–72.
- Burrell, J. (2023). Algorithmic Opacity. *Nature Digital Society*, 2(4), 417–432.
- Chen, L. (2024). AI Art and Value. *ArtNet News*.
- Cho, A. (2023). Superconductor Surprise. *Quanta Magazine*.
- Dastin, J. (2018). AI Hiring Tools and Discrimination. *Reuters*.
- Harari, Y. N. (2015). *Homo Deus: A Brief History of Tomorrow*. Harvill Secker.
- Jumper, J., et al. (2023). AlphaMissense and Mutation Tracking. *Nature*.
- Matthias, A. (2004). *The responsibility gap: Ascribing responsibility for the actions of learning automata*. *Ethics and Information Technology*, 6(3), 175–183.
- Nature Editorial Board. (2023). Retraction of Graphene Paper. *Nature*, 620(7972), 10.
- Romero, V. (2023). California Algorithm Error. *LA Times*.
- Roy, A. (2014). *Capitalism: A Ghost Story*. Verso Books.
- Shelley, M. (1818). *Frankenstein; or, The Modern Prometheus*. Lackington, Hughes, Harding, Mavor & Jones.
- Silverman, S. (2023). Goldman Sachs and AI Reports. *Wall Street Journal*.
- Stanford AI Lab. (2023). Chatbot Empathy Bias Study.
- Suleyman, M. (2023). *The Coming Wave: Technology, Power, and the Twenty-First Century's Greatest Dilemma*. Crown Publishing.
- Tammet, M. (2023). *Digital Courts in Estonia*. Tallinn Univ. Press.