CHARTERED GOVERNANCE
Hong Kong

Responsible AI Policy Development:

A GOVERNANCE PLAYBOOK

September 2025

PRELUDE

AI is no longer a distant prospect; it is increasingly shaping industries, business models, and workforce expectations. Ignoring its impacts, therefore, is no longer an option. For organisations, the challenge is no longer whether to adopt AI but how to adopt it responsibly. As AI becomes embedded in core business processes and decision-making, the need for a clear, well-governed AI policy has become pressing, which is the subject of this governance playbook.

Organisations must first clarify ownership of their AI policy. Oversight should sit at a sufficiently senior level, typically the board or a delegated risk or technology committee, to ensure strategic alignment and accountability for outcomes. The board's role is to approve the AI policy, set risk appetite, and monitor management's progress in implementation.

Senior management should be responsible for drafting and maintaining the policy, drawing input from key stakeholders, including IT, cybersecurity, data governance, legal, compliance, risk management, HR (to address workforce impact), and business leaders adopting AI solutions. Premature or poorly governed adoption can create legal, ethical, and reputational harms. Implementation should be operationalised by cross-functional teams so that controls, ethical standards, and regulatory requirements are embedded into daily processes. Periodic review and regular board-level reporting on AI risks, benefits, and incidents should ensure that the policy remains current as technology, regulations, and business priorities evolve.

But governance cannot focus solely on risk containment. There are also risks in delaying or avoiding engagement with the potential upsides of AI technologies. Organisations that do not explore these tools may find themselves at a disadvantage if competitors achieve gains in efficiency, innovation, or service delivery. The talent dimension is equally critical: next-generation professionals increasingly expect to work in AI-enabled environments, and organisations that cannot offer this may face recruitment and retention challenges.

This is why AI governance must strike a careful balance between risk mitigation and enablement. The objective is to create the conditions for responsible experimentation, measured adoption, and continuous learning. Boards and governance professionals can support this by enabling safe "sandbox" pilots, setting clear oversight parameters, and scaling successful use cases with appropriate controls. Overly rigid frameworks may protect the organisation from present risks, while limiting its capacity to respond and adapt to tomorrow's opportunities and challenges.

Governance professionals are uniquely positioned to make this balance work in practice. They bridge the board, management, and operational teams, translating regulatory expectations into practical policies and embedding accountability, transparency, and ethical standards across the organisation. By ensuring that governance is both a guardrail and a catalyst, they help organisations turn AI governance from a brake on progress into a driver of sustainable growth.

FOREWORD

Responsible AI Policy Development: A Governance Playbook

At the Institute, we take pride in our thought leadership in governance. A risk that is often overlooked is the risk of not adopting AI. Companies that are too cautious may fall behind in cost competitiveness, product innovation, and their ability to attract next-generation talent. Governance should face this challenge head-on, enabling responsible AI adoption rather than stifling innovation.

In this report, in the context of artificial intelligence (AI) as a driver of innovation and efficiency, we focus on how to adopt responsible AI policies to capitalise on related opportunities and manage new and evolving risks. The aim is for the ethical, transparent, and accountable use of AI.

There is no one-size-fits-all approach to AI governance. The diverse range of AI applications across industries means that organisations must tailor their policies to address their specific challenges, regulatory obligations, and values. While it's understandable that many businesses are adopting versions of AI governance policies modelled on others, we have observed that this approach often overlooks the unique risk profiles and circumstances of individual organisations. This is especially true as the AI landscape is still emerging, and we are only beginning to fully grasp the complexities of AI's risks and opportunities.

For organisations that are sophisticated users of AI technologies, a more bespoke approach is essential. Responsible AI governance requires careful consideration of what constitutes responsible use in the context of their specific operations, objectives, and risk tolerance. As such, the matters discussed in this report are not only timely but necessary for organisations looking to build robust, risk-appropriate AI frameworks.

This report offers a comprehensive playbook to responsible AI governance, structured across five areas:

- **AI Governance Matters** – An exploration of why effective AI governance is crucial in today's business landscape, and how it can safeguard both organisations and society from the risks of AI deployment.
- **Operationalising AI Governance** – A deep dive into real-world examples of AI applications and the complex governance challenges that arise as AI systems become more advanced and integrated into business processes.
- **Dynamic AI Governance: Building Policies That Evolve** – Insights on how to craft adaptable policies, enabling organisations to keep pace with the fast-changing nature of AI technologies and regulatory landscapes.
- **Responsible AI Policy Framework and Example** – A practical framework for developing responsible AI policies, alongside an example to guide organisations in their policy process.
- **Director Briefing Template** – A customisable template designed for directors to quickly understand the key considerations of AI governance, enabling them to make informed decisions on AI adoption and policy development.

We trust that this report will serve as a valuable resource, equipping organisations with the knowledge and tools needed to develop AI governance frameworks that not only promote innovation but also mitigate risks, ensuring the responsible use of AI in a rapidly evolving environment.

I want to thank the authors for their contributions to the Institute's thought leadership.

Mr David Simmonds FCG HKFCG
President, The Hong Kong Chartered Governance Institute
Chief Strategy, Sustainability & Governance Officer, CLP Holdings Limited

ACKNOWLEDGEMENTS

Authors

Roshan Bharwaney, Ed.D

Roshan is a global technology, operations, talent, and organisational development professional with over 20 years of experience in individual, team, and business transformation. He is a strategic and hands-on trusted adviser, collaborator, and thought partner to senior stakeholders across global, matrixed, multicultural organisations. His previous roles include Strategy & Innovation Principal at Meta and Associate Director at WPP. With a doctorate in Adult Learning & Leadership and a Master's in Organisational Psychology from Columbia University, he brings a research-driven approach combining insight, design, and execution to help organisations adapt, learn, and grow across cultures and markets.

Mohan Datwani FCG HKFCG(PE) Deputy Chief Executive, The Hong Kong Chartered Governance Institute

Mohan is a seasoned solicitor and governance professional with nearly 35 years of experience spanning law, corporate leadership, and public service. Formerly an owner of a prominent US international law firm and General Counsel of a listed multinational, he is now Deputy Chief Executive of The Hong Kong Chartered Governance Institute (HKCGI). He spearheads policy, research, and advocacy across Hong Kong and mainland China, serving a community of about 10,000 professionals. Notably, he submitted HKCGI's first formal proposal to the Financial Services and the Treasury Bureau, advocating for a Hong Kong re-domiciliation regime, and authored a pivotal report for the Equal Opportunities Commission, which was presented to LegCo, driving institutional reform and the adoption of a victim-centric approach. His work was recognised with the Directors of the Year Award by the Hong Kong Institute of Directors in 2018.

Roshan P. Melwani, MPP

Roshan has a decade of experience spanning litigation, public policy, and applied AI research. He has advised organisations on the responsible use of AI in high-stakes decision-making contexts, most recently at Climate Policy Radar and the Helen Bamber Foundation. His work centres on the intersection of human rights and emerging risks, with previous roles at UNHCR, The Migration Observatory, and the private sector. Grounded in frontline experience with vulnerable communities, Roshan brings regulatory insight and a nuanced understanding of the ethical challenges posed by technology. He holds a Master of Public Policy from the University of Oxford and a law degree from the London School of Economics.

Dylan Williams FCG HKFCG

Dylan is EVP, General Counsel and Company Secretary of Sands China Ltd. (HKEX: 1928), where he has spent 19 years helping guide one of Macao's major integrated resort operators through significant growth and transformation. A New York-licensed attorney with nearly three decades of experience in Hong Kong, Macao, and Greater China, Dylan leads a team of over 40 legal professionals and serves on the company's Executive Committee. His work has spanned complex transactions, including multi-billion-dollar project financings, Hong Kong IPOs, gaming concession negotiations, and navigating regulatory compliance across multiple jurisdictions.

With a practical approach to bridging legal and business priorities, Dylan has focused on modernising legal operations through technology and AI, developing custom platforms that streamline workflows across a broad range of legal disciplines. Before joining Sands China, he held leadership roles in technology and media companies in Hong Kong and served as a Governor of the Chinese International School for 16 years. Dylan is a Fellow of the Hong Kong Institute of Chartered Secretaries.

Contributors

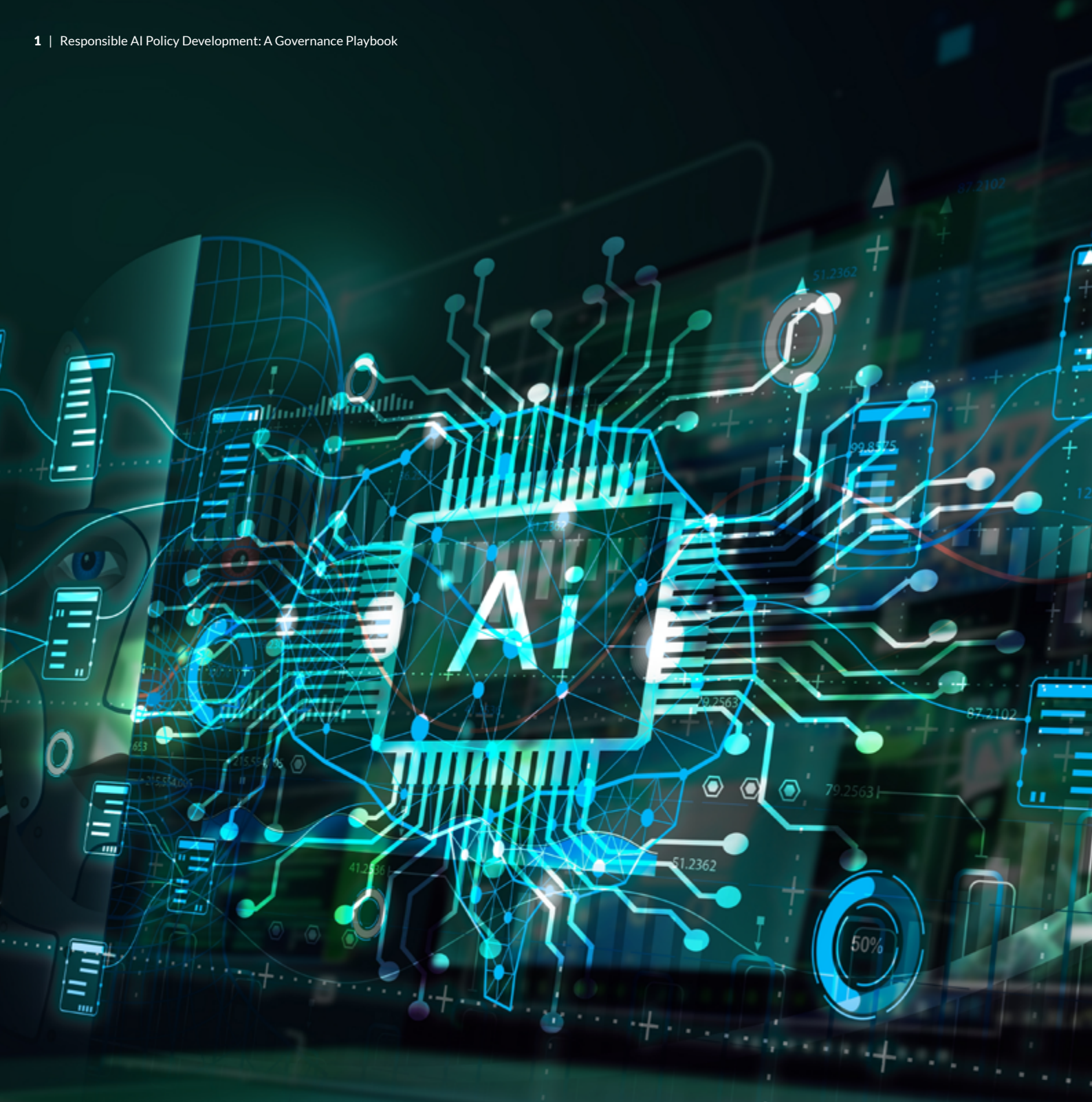
Michael Ling FCG HKFCG

Chair, Institute Technical Consultation Panel (TCP) as well as TCP Members.

Ellie Pang FCG HKFCG(PE)

Chief Executive, The Hong Kong Chartered Governance Institute.

CONTENTS



EXECUTIVE SUMMARY

Responsible AI Policy Development: A Governance Playbook

This playbook provides governance professionals, directors, and senior management with practical tools to develop tailored AI governance frameworks that strike a balance between innovation, ethical responsibility, and risk management.

1. Core To AI Development: Six Responsible AI Principles

2. Five Imperatives for AI Governance

- **Build a Use Case Inventory** - Maintain a central registry of all AI systems, capturing purpose, data sources, risk level, and ownership. You can't govern what you can't see.
- **Translate Principles into Practice** - Link fairness to bias audits, transparency to explainability standards, and accountability to escalation protocols. Make values operational.
- **Embed AI Governance into Existing Structures** - Incorporate AI risk into board agendas, enterprise risk frameworks, procurement reviews, and internal audit plans.
- **Treat Policy Development as a Living Process** - Establish regular review cycles to keep frameworks current as regulations evolve and risks emerge. A static policy is a risk.
- **Equip Boards and Staff with the Right Questions** - Provide tailored guidance through practical tools like risk checklists and oversight questions to build a culture of accountability.

3. What This Playbook Provides

Chapter	Key Deliverable
1: AI Governance Matters	Six-principle framework and risk mapping
2: Operationalising AI Governance	Use case inventories, risk assessments, lifecycle oversight tools
3: Dynamic AI Governance	Continuous review mechanisms and institutional learning approaches
4: Policy Framework	Ready-to-adapt AI policy template with operational controls
5: Director Briefing	Board-ready questions and oversight guidance

4. Key Takeaway

AI governance requires tailored approaches—not one-size-fits-all solutions. This playbook helps organizations build frameworks aligned with their unique risk profiles, regulatory obligations, and values while maintaining flexibility to evolve with technology and regulation.

CHAPTER 1

AI GOVERNANCE MATTERS

1.1 Introduction

Governance must look both ways: mitigating the harms of irresponsible AI use and avoiding the strategic risk of falling behind. Failing to adopt AI can result in loss of market share, inefficiency, and weakened competitiveness. Organisations should treat AI as a core capability, actively encouraging responsible experimentation and scaling successful pilots. Governance professionals should advocate for clear frameworks that enable innovation within safe boundaries, allowing risk management and competitiveness to advance in tandem.

Artificial intelligence (AI) technologies have the potential to enhance how organisations make decisions, deliver services, manage risks, optimise resources, and analyse data. From process automation to data-informed innovation, AI-based tools and systems can offer strategic opportunities for companies. However, what is required to govern AI technologies safely will depend on how they are deployed within an organisation, their integration with human expertise, and the careful management of their limitations.[1]

As such, AI technologies also introduce distinct categories of risk that organisations must carefully manage. Technical risks include algorithmic bias, model errors, and safety failures that can lead to problematic downstream consequences.[2] Organisational and legal risks encompass issues like legal non-compliance, intellectual property infringement, data misuse,

privacy violations, and security breaches.[3] These risks can, in turn, cause operational disruptions, damage stakeholder trust, and jeopardise stakeholder relationships with partners, regulators, and the public.[4]

Effective application requires a clear understanding of what AI systems can and cannot do, as well as ongoing oversight to ensure they align with organisational goals, ethical standards, and regulatory requirements.[5] Governance professionals play a vital leadership role in bringing these governance issues to the fore, ensuring that AI is not adopted as a purely technical or commercial tool but as a capability that demands proper oversight, alignment with purpose, and accountability across the organisation.

While governance professionals are often the conveners and framers of policy conversations, the development of a credible AI Policy requires active board oversight and organisation-wide involvement, particularly at the senior management level.[6] There is no one-size-fits-all approach — each organisation must tailor its policy to fit its unique business model, AI applications, deployment scenarios, risk profile, and stakeholder expectations.[7]

This report provides practical guidance for governance professionals to initiate, structure, and lead AI governance processes, ensuring that organisational policies and frameworks are not only aspirational but also credible, enforceable, and adaptable.

1.2 Benefits of AI Policy Development

A well-crafted AI policy should do more than codify good intentions. It should act as a practical mechanism to ensure the deployment of AI systems reflects meaningful responsibility, mitigates harms, and is subject to appropriate oversight.[8] Developed rigorously, an AI policy helps organisations to:

- **Anchor organisational values and ethical principles** in enforceable standards to guide safe deployment and restrict misuse.[9]
- **Strengthen accountability**, enabling checks and balances, clear oversight mechanisms, and transparent decision-making processes.[10]
- **Stay responsive to public scrutiny**, treating regulatory compliance as a floor while engaging meaningfully with civil society.[11]
- **Foster trust and open communication** with internal and external stakeholders, promoting inclusive and socially grounded applications of AI.[12]
- **Mitigate downstream and systemic risks**, particularly where these disproportionately impact vulnerable groups or the environment.[13]
- **Facilitate responsible innovation**, enabling teams to deploy and scale AI solutions confidently where appropriate - knowing that clear guidelines, redress pathways and risk controls are in place.[14]

1.3 Overview of the Regulatory Landscape for AI Technologies

This section offers a generalised picture of the regulatory landscape in Hong Kong, and should be read alongside emerging regulatory developments in jurisdictions of operation:

Hong Kong has adopted a context-specific, sector-led approach to AI governance,[15] rather than introducing a single, overarching law like the EU AI Act. Instead, the government has relied on existing laws and sectoral guidelines, supplemented by voluntary frameworks, to manage AI-related risks.[16] Across industries, two Hong Kong government bodies have taken the lead in promulgating AI standards:

Office of the Privacy Commissioner for Personal Data (PCPD)

August 2021:
Published Guidance on the Ethical Use and Development of AI, outlining high-level principles for responsible development and use of AI.[17]

June 2024:
Developed the Model Personal Data Protection Framework, providing practical measures for organisations to establish robust AI governance strategies, conduct comprehensive risk assessments, manage AI models securely, and engage transparently with stakeholders.[18]

Digital Policy Office (DPO)

July 2024:
Issued the Ethical AI Framework, an internal reference for government departments that is also recommended to external organisations. It lays out ethical principles, governance models and assessment templates.[19]

April 2025:
Published the Generative AI Technical & Application Guidelines, a best-practice guide for developers, platform providers and users of generative AI systems.[20]

While compliance with these frameworks is voluntary, the underlying Personal Data and Privacy Ordinance (PDPO) obligations are not. As of February 2025, the PCPD have launched a new round of AI security compliance checks for organisations across various sectors, including telecommunications, banking and finance, insurance, beauty services, retail, transportation, education, medical services, public utilities, social services and government departments.[21]

Moreover, a range of sector-specific circulars have been published by industry bodies - most notably in banking and finance,[22] healthcare,[23] and insurance.[24] So while the regulatory landscape remains fragmented, AI governance in Hong Kong is **already enforceable** through existing laws and sectoral guidance in the absence of specific legislation. Consequently, the above standards increasingly reflect what regulators expect to see during investigations, reviews, or licensing. Governance professionals should therefore treat voluntary frameworks as practical compliance imperatives, as these quickly become de facto expectations in boardrooms and compliance reviews.

To stay ahead of the curve, governance professionals should also monitor jurisdiction-specific developments, alongside the convergence of leading international AI frameworks. Arranged from broad principles through to specific operational guidance, useful touchstones include:

1. OECD AI Principles;[25]
2. Singapore Model AI Governance Framework.[26]
3. EU AI Act (and General-Purpose AI Code of Practice).[27, 28]
4. The NIST AI Risk Management Framework.[29]
5. ISO/IEC 42001,[30] and ISO/IEC 23894.[31]
6. China's Interim Measures for the Management of Generative AI Services.[32]

1.4 Six Core Principles: Learning from Microsoft's Responsible AI Standard (RAIS)

Microsoft's Responsible AI Standard (RAIS) offers a useful reference for maintaining effective board-level oversight in a rapidly evolving landscape.[33] First published in 2022 and updated regularly, including major revisions in 2024 to address generative AI risks, it is built around six core principles:

- Fairness.
- Reliability & Safety.
- Privacy & Security.
- Inclusiveness.
- Transparency.
- Accountability.

Each principle is backed by structured requirements and measurable safeguards, including red-teaming (i.e., simulating adversarial attacks or challenges to test and improve the effectiveness, security, or resilience of systems), safety testing, documentation standards, and defined human oversight.[34]

Governance professionals can draw on this model, but must resist the temptation to replicate it without critical thought. This is particularly crucial given that "alignment" between what humans want and what AI systems can do remains an open research problem.[35] Organisations must therefore carefully reflect on their values, risk appetite, regulatory context, and technical realities, working closely with senior leadership to embed tailored policy into strategic priorities and day-to-day operations.

1.5 Linking Principles to Risks: How Governance Professionals Add Value

The chart below distils the six core principles and their implications, illustrating how each principle might show up in governance. As a starting framework, the issues outlined are not a complete catalogue. But for each principle, we flag representative risk categories to show the path from value statement to operational exposure. Governance professionals should therefore expand or revise this mapping to reflect their own sector, jurisdictions, impact profile and emerging external standards, then design controls in proportion to the level of risk identified:

1. Fairness

Fairness asks whether an AI system treats individuals without discrimination and avoids unjustified disparate impact across protected or vulnerable groups. In practice, this covers **training data representativeness, model design choices, and outcome monitoring.**

Example governance risks:

- ✔ **Legal exposure:** Algorithmic discrimination can lead to liability issues under Hong Kong's anti-discrimination ordinances and infringe on the right to equality protected under Article 25 of the Basic Law or the HK Bill of Rights.
- ✔ **Reputational licence:** Perceived injustice can trigger media backlashes, activist litigation and regulator intervention.[36]
- ✔ **Capital allocation:** Using AI models can potentially distort credit, hiring or pricing decisions, embedding systemic bias in business outcomes.[37]

2. Reliability & Safety

Reliability and safety entail whether an AI system performs as intended under expected and unexpected conditions, while avoiding harm and unintended side effects. This covers issues of technical robustness, fault tolerance, risk containment, and safe deployment across diverse environments and user contexts.

Example governance risks:

- ✔ **Operational disruption:** Unreliable models can hallucinate, drift, malfunction, or fail under stress, leading to system downtime, service degradation, or cascading failures.[38]
- ✔ **Legal exposure:** Unsafe AI systems may breach product liability or duty-of-care obligations or consumer protection statutes.[39]
- ✔ **Trust erosion:** Safety failures, especially in high-stakes use cases, can undermine user trust, investor confidence, and long-term adoption.[40]

3. Privacy & Security

Privacy and security entail whether an AI system protects individuals' personal data and prevents unauthorised access, manipulation, or misuse. This spans data collection, storage, and usage practices, as well as cybersecurity safeguards across the AI lifecycle. Organisations must ensure their usage of AI tools or systems complies with data protection laws.

Example governance risks:

- ✔ **Regulatory penalties**: Mishandling personal data can trigger formal investigations, audits and substantial fines, under both Hong Kong law and in other jurisdictions.[41]
- ✔ **System compromise**: Insecure AI models and infrastructure are targets for adversarial attacks, hacks, data breaches, and model inversion techniques.[42]
- ✔ **Reputational loss**: Perceived misuse, leakage, or unauthorised use of user data can erode public trust and damage stakeholder relationships.

4. Inclusiveness

Inclusiveness asks whether AI systems are accessible, usable and beneficial across demographic, linguistic, cultural and disability dimensions, and whether they avoid creating new digital divides.

Example governance risks:

- ✔ **Market share loss:** If products don't work in minority languages, fail to cater to disability needs, exclude certain demographics by design, or underperform in certain markets.
- ✔ **Monolingual bias:** Over-reliance on English training data can degrade performance for Cantonese and Putonghua users.
- ✔ **Regulatory non-compliance:** May contravene ESG-related standards, such as those found in HKEX's Corporate Governance Code.[43]
- ✔ **Innovation blind spots:** May overlook use cases, risks, or opportunities relevant to broader populations.

5. Transparency

Transparency covers both **model explainability** (stakeholders can understand how outputs were produced), and **organisational disclosure** (being open about AI usage, limitations and governance). This includes internal traceability, user-facing explanations, documentation of design choices, and openness about limitations and risks.[44]

Example governance risks:

- ✔ **Litigation exposure** if decisions cannot be explained to regulators or the Court. Many jurisdictions increasingly require data transparency and model explainability in high-risk domains; non-compliance can result in sanctions, service rollbacks, or product bans.[45]
- ✔ **Consumer confidence:** Knowing whether and how AI technologies are used will enable consumers to make informed decisions, ensure trust, and prevent backlash. Consumers may reject products and services if they cannot understand their outputs or challenge their decisions.
- ✔ **Accountability gaps:** The intrinsic opaqueness of AI models means that they can be hard to debug, especially if key technical staff leave. This can make it difficult to identify root causes of errors or harms, hindering redress and oversight.

6. Accountability

Accountability ensures that clear ownership, oversight, and redress mechanisms are in place for the design, deployment, and impact of AI systems. In other words, that identifiable humans, and ultimately the board, remain answerable for AI-generated outcomes, with clear mechanisms to trace responsibilities, remedy harms, and learn from incidents. It includes assigning responsibility across functions, tracking decisions, and ensuring consequences for misuse or failure.

Example governance risks:

- ✔ **Incident under-reporting and ethical drift:** If escalation routes are unclear, this can amplify harm and delay remediation. Lack of accountability can lead to unmonitored deployment, scope creep, or misalignment with organisational values.[46]
- ✔ **Vendor risk transfer:** If relying on external models without contractual recourse, residual liabilities may land on the organisation for unanticipated failures.[47]
- ✔ **Organisational blind spots:** When no one owns AI outcomes end-to-end, risks can fall between the cracks and go unaddressed.

1.6 Structuring Your AI Policy: Securing Buy-In

The AI Policy must reflect whole-of-organisation participation, not only from governance, legal, and IT, but also from business heads, operations, marketing, HR, and internal audit. Senior management ownership or buy-in is critical for sustained adoption and effectiveness.

- **Initiate Leadership Engagement**. Secure board support and appoint a cross-functional lead group reporting to senior management.
- **Facilitate a Cross-Functional Workshop.** Use the six principles to identify practical applications and trade-offs. Ensure input from all relevant business units.
- **Draft a Policy Charter and Gap Assessment**. Map existing policies (e.g., cybersecurity, procurement, data protection) to identify overlaps and blind spots.
- **Define Oversight and Review Mechanisms**. Integrate AI governance into existing board risk or ethics committees, with regular reporting and oversight to ensure effective management.

CHAPTER 2

OPERATIONALISING AI GOVERNANCE

2.1 Introduction: Governance as AI Use Matures

As organisations deploy more sophisticated AI systems, such as large language models, predictive analytics, and automated decision-making tools, additional governance challenges emerge. Legal exposure, stakeholder scrutiny, and operational complexity tend to rise sharply.

At this stage, the governance professional should play a critical role in translating principles into practice. Acting as a facilitator, the governance professional connects technical, legal, and business functions, ensuring that responsible AI practices are integrated into daily operations, risk frameworks, and compliance structures. They also connect and align with external stakeholders, such as regulatory bodies, and assume the roles of horizon scanning and liaison.

2.2 Translating Risks into Governance: Promoting End-to-End Accountability

For chartered governance professionals, the imperative is clear: establish robust internal governance structures that proactively manage AI risks across its lifecycle.

This means moving beyond reactive measures to embed ethical considerations and accountability across the organisation. Governance professionals should take the lead in:

- Interpreting each principle from an internal governance lens.
- Translating abstract values into policy elements.
- Embedding policy elements into an end-to-end accountability framework.
- Ensuring the board and senior management understand their oversight responsibilities.
- Engaging operational teams to design and implement meaningful processes.

The governance professional should coordinate risk-mapping workshops across functions, adapting oversight structures to reflect the nature and purpose of AI applications. The table below provides a non-exhaustive list of governance actions that teams can consider:

Principle	Non-Exhaustive Governance Actions
Fairness	✔ Require fairness audits before and after deployment to detect harms and address bias over time. ✔ Ensure human-in-the-loop oversight in critical decisions, particularly to catch edge cases not handled well by automation. ✔ Use diverse, representative and up-to-date datasets, documenting provenance, gaps and limitations. ✔ Define and approve fairness metrics, drawing from broad stakeholder input. ✔ Secure board approval of chosen fairness trade-offs through a metric-justification memo. ✔ Establish clear escalation and remediation processes if unfair outcomes are detected.
Reliability & Safety	✔ Introduce red-teaming protocols, stress testing and scenario planning, simulating technical and/or organisational failures. ✔ Maintain fallback procedures and post-launch safety checks. ✔ Conduct robust testing and validation, redundancy, and fail-safes. ✔ Establish model performance benchmarks to validate outputs under expected, edge-case and adversarial conditions. ✔ Monitor for hallucination and model drift. ✔ Utilise intrusion detection, data encryption, and secure channels, and perform regular data audits.
Privacy & Security	✔ Map data flows and perform thorough legal/privacy compliance reviews of all AI systems. ✔ Employ adversarial testing: Test for inversion attacks, membership inference attacks, prompt injection and data leakage. ✔ Ensure board-level visibility over AI incident response readiness. ✔ Extend privacy and security requirements to third-party vendors through contracts, audits, and monitoring.
Inclusiveness	✔ Conduct inclusive user testing and solicit feedback from diverse communities during design and post-deployment. ✔ Use diverse and representative datasets, inclusive design principles, and social impact assessments to mitigate and monitor exclusionary outcomes. ✔ Integrate inclusiveness KPIs into internal risk reporting.
Transparency	✔ Introduce model documentation templates. ✔ Lead policy development on AI explainability, including tiered requirements proportionate to risk level and regulatory expectations. ✔ Commission periodic external audits or assurance reviews of transparency claims and documentation. ✔ Train frontline staff to discuss AI outputs, risks, and limitations effectively.
Accountability	✔ Establish a governance structure that identifies responsible individuals and outlines clear escalation paths for addressing issues. ✔ Require vendor accountability clauses covering risk disclosures, remediation obligations, and audit rights. ✔ Develop clear internal guidelines for employees on using AI responsibly. ✔ Support board committees in reviewing AI risk reporting and approving deployments. ✔ Foster a culture of ownership, where AI accountability is not outsourced to vendors or technical teams.

2.3 Building the AI Governance Toolkit

To embed AI oversight into everyday processes, the governance professional should help develop and promote internal governance instruments tailored to the organisation's AI maturity and risk profile.

1. Use Case Inventory

Establish a central inventory of all AI systems in use. This registry enables internal visibility and facilitates board-level oversight. Suggested inputs are:

- ✔ Business owner and system purpose
- ✔ Model type and data sources
- ✔ Risk classification (e.g. customer impact, regulatory sensitivity)
- ✔ Explainability level and lifecycle stage

Additional documents:

- ✔ Product requirements document (PRD)
- ✔ Stakeholder mapping (clarifying who is affected and who should be consulted)
- ✔ System mapping (flow of inputs into model to outputs to downstream impacts)

2. AI Risk Assessment Addendum

Integrate AI-specific questions into existing enterprise risk assessments or product development checklists to enhance their effectiveness and accuracy.

These should include:

- ✔ Does the system influence decisions with legal or ethical consequences?
- ✔ Are personal or sensitive data used?
- ✔ Can outputs be explained and challenged?
- ✔ How do these risks map onto regulatory standards?

A **failure modes and effects analysis** (FMEA) and **"what-if" anticipatory scenario worksheet** can facilitate the early identification of downstream impacts, promoting proactive risk mitigation.[48]

3. Model Artefacts & Traceability Standards

The governance professional should require all teams deploying AI to complete standardised documentation, addressing:

- ✔ Intended use and limitations
- ✔ Data used for training and validation
- ✔ Bias mitigation techniques
- ✔ Explainability, performance, and safety metrics

Such documents can include:

- ✔ **Model cards:** This document should clarify the intended applications of AI models, accompanied by details of their performance characteristics, assumptions made, harms anticipated and mitigation actions taken;[49]
- ✔ **Datasheets for datasets:** This document accompanies datasets used for a model, outlining the reasons for the data, its composition, the collection process and recommended uses;[50]
- ✔ **Algorithmic Design History File** - A running log of design changes, decisions and test results;[51]
- ✔ **Checklists** to confirm all model artefacts are present before each gate of the AI lifecycle

4. Ethics or Exception Review Process

For high-risk or novel use cases, the governance professional should convene an oversight forum—either an existing committee or a new review board—drawing on legal, compliance, technical, and risk functions.

The forum should conduct and review **social impact assessments** to make "go / no-go / revise" decisions.[52] It should be empowered to delay or reject deployments until mitigations, risk thresholds and critical stakeholder concerns are addressed.

5. Post-Deployment Monitoring & Assurance

As AI systems become integrated and scaled in workflows over time, the governance professional must continuously oversee their impacts. Oversight tools include:

- ✔ Real-time dashboards for **performance drift, bias drift, hallucination rates** and security anomalies;
- ✔ Using **audit checklists** to ensure document completeness;
- ✔ Scheduled **re-audits** and red-team tests after material changes or system updates;
- ✔ Regular reviews of remediation and risk mitigation plans

2.4 From Assurance to Enablement

As AI becomes increasingly central to strategic operations, the governance professional plays a crucial role in transforming principles into actionable governance mechanisms. By coordinating policy implementation, enabling alignment across teams, and embedding continuous oversight, the governance professional supports not only compliance but also sustainable innovation.

Done effectively, this work:

- Reduces ethical, legal, and reputational risks.
- Strengthens clarity and accountability within the organisation.
- Builds the conditions for the adoption of trusted, socially grounded, and responsible AI.

CHAPTER 3

DYNAMIC AI GOVERNANCE: BUILDING POLICIES THAT EVOLVE

3.1 Introduction: Governance Beyond Implementation

End-to-end accountability requires that AI oversight extends beyond the initial development and deployment stages. In this chapter, we focus on what comes next: embedding, evolving, and institutionalising AI governance. Governance does not end with a policy or a set of procedures—it is a continuing process of reflection, adaptation, and improvement.

Governance professionals play a crucial role in ensuring that AI governance frameworks remain dynamic, integrated, and fit for purpose as technologies evolve and regulatory landscapes mature. In this phase, the goal is to move from compliance to confidence, enabling organisations to govern AI responsibly while supporting innovation and agility. This includes building structures that can adapt to regulatory changes and emerging risks, ensuring long-term resilience and trustworthiness.

The governance professional should ensure that AI systems are managed throughout their entire lifecycle, with clearly defined responsibilities and regular reviews to ensure effective oversight.

3.2 Governance in Motion: The Need for Continuous Review

AI governance must adapt to rapid changes in use cases, stakeholder expectations, and regulatory requirements. A well-structured review process ensures that AI policies remain relevant and effective.

Recommended actions:

- **Establish a regular review cycle** (e.g. every 6–12 months), embedded into board and committee agendas (with a focus on fairness, safety, reliability and risk).
- **Trigger ad hoc reviews** in response to, for example, material AI-related incidents or near misses; high-impact or experimental deployments; and significant regulatory changes (e.g. EU AI Act, China's AI regulations, updates in data protection laws).
- **Ensure inclusive review participation**, involving legal, risk, IT, operations, compliance, and frontline units.
- **Document updates through defined governance pathways**, including sign-off by the board and senior management.

This approach ensures that AI governance remains a living framework, not a static rulebook.

3.3 Institutional Learning: Reflecting on Practice to Inform Policy

AI governance cannot succeed solely through policies and procedures: promoting a reflective, "lessons learnt" culture is essential. This requires a consistent feedback loop and horizon scanning. The governance professional should collaborate with senior management and other groups across the organisation to foster an internal culture that promotes psychological safety, transparency, trust, responsible AI use, and equips staff with the necessary skills and expertise to effectively utilise it.

Recommended actions:

- **Facilitate structured pre-mortems and post-deployment reviews**.[53] Asking, for example, were ethical trade-offs documented and discussed? Did the oversight mechanisms function as intended? Were the affected stakeholders properly considered?
- **Construct "user stories"** to understand an AI's functionality from a particular user's view.[54]
- **Delivering targeted training** for business functions on the AI policy's practical implications.
- **Record and integrate lessons learned** into future risk assessments, policy updates, training, and design standards. Ensure cross-functional input (legal, compliance, IT, user teams) to generate comprehensive insights.
- **Actively track** AI-related developments (regulatory, technological, best practices), and how peers are operationalising their AI policies.

The governance professional's role is to ensure this reflection is formal and part of the governance culture.

3.4 Cross-Functional Stewardship: Building Capacity and Coordination

As AI governance matures, strong internal coordination and robust tools are essential. Beyond deployment, organisations must learn from their own and others' real-world use to improve future policies, controls, and decision-making processes. Governance professionals should lead the development of practices to ensure that these initiatives are well-coordinated and effective:

Recommended actions:

- **Establish cross-departmental forums**: working groups that include risk, compliance, legal, technology, business, and data teams. These groups should meet regularly to share insights, challenges, and evolving practices.

- **Develop short, role-specific guides** (e.g. "AI Risk Checklist for Marketing Teams").

- **Maintain a central AI governance resource hub** (e.g. policies, FAQs, templates, real-world case examples). Review data from tools that collect feedback on AI system performance, incidents, audits, and user experience, to refine policies and governance processes.

- **Maintain decision logs** and capture rationale for high-impact governance decisions. Track exceptions and flag deviations from policy and provide remediation steps.

- **Standardise internal assessments**, approvals, and reviews using consistent templates and criteria. Track risks, edge cases, policy adoption and usage metrics as part of internal audits or KPI.

- **Benchmark and gather external intelligence**, regularly scanning the environment to see how other organisations and sectors are governing AI and integrating lessons learned and best practices.

- **Prepare emergency fallback mechanisms** pre-deployment.

- **Ensure retired models are decommissioned** with no unintended residual influence.

These help reduce blind spots, support transparency, and enable effective institutional learning.

3.5 Strategic Enablement: Supporting AI Innovation with Confidence

Governance professionals must help organisations strike a balance between control and enablement, ensuring that responsible governance does not stifle innovation but builds the foundation for sustainable adoption. A restrictive framework can inadvertently slow innovation and erode competitiveness. Strategic enablement involves assessing both risk and opportunity, ensuring that high-potential use cases receive appropriate support and guardrails rather than blanket rejection. Some organisations deploy 'sandbox' environments, allowing safe experimentation under controlled conditions, followed by scaled deployment once risks are addressed.

Recommended actions:

- **Reinforce key ethical principles** through leadership communications and team discussions.

- **Track and report AI developments** through regular dashboards and updates to the board, including deployment trends, exceptions, incident trends, emerging risk indicators, and regulatory developments and their strategic implications.

- **Support safe whistle-blowing** and meaningful escalation mechanisms, ensuring staff have clear, trusted channels to raise issues before downstream impacts emerge.

- **Conduct scenario planning and reputational risk assessments** for novel or high-profile AI initiatives.

By reinforcing escalation pathways and promoting open dialogue, governance professionals create a culture of early issue resolution, strategic foresight, and responsible innovation.

3.6 Building Trust and Delivering Value

In today's fast-moving AI environment, a static policy is a risk. AI governance must be treated as an ongoing, organisation-wide effort—one that evolves in step with operational realities, stakeholder expectations, and regulatory shifts.

Governance professionals play a crucial role in this journey. By leading review cycles, promoting learning, fostering coordination, and supporting innovation, they help transform AI governance from a compliance obligation into a source of strategic advantage.
With the right structure and mindset, governance professionals can ensure their organisations govern AI with confidence, responsibly, resiliently, and in the public interest.

CHAPTER 4

RESPONSIBLE AI POLICY FRAMEWORK AND EXAMPLE

This chapter offers a sample AI Policy template for organisations seeking to formally codify their AI governance practices. This template is designed to help governance professionals facilitate the development of a policy that is:

- Aligned with international best practices and evolving regulatory frameworks;
- Rooted in organisational values and risk priorities;
- Practical and enforceable across operational settings;
- Adaptive to technological and legal developments over time.

Governance professionals are not expected to act as AI developers or technologists. Their role is to coordinate across functions, ensure appropriate oversight structures, and help tailor policies that support responsible, risk-informed innovation.

This sample policy is non-exhaustive and illustrative, and should be adapted and expanded in consultation with legal, risk, compliance, and technical teams to reflect the organisation's specific risk profile, business model, internal structure, decision-making culture, applicable laws and regulations, and the types of AI technologies used, as these technologies bring their own nuances:

- **Tier 1: Minimum Viable AI Policy** - Covers essential elements such as core principles, key requirements, basic governance structure, and prohibited uses.
- **Tier 2: Comprehensive Policy Additional Elements** – For mature or technology-forward organisations to consider adopting in addition to Tier 1 elements, with the caveat that there is no one-size-fits-all.

Sample AI Policy Template

1. Policy Objective and Scope (please align to your numbering and formatting)	
This AI Policy outlines the principles, governance roles, and operational controls that guide the development, acquisition, deployment, and use of AI systems within [Organisation Name]. The purpose of this Policy is to ensure that AI is used **safely, ethically, lawfully**, and in alignment with our organisational values, including **privacy protection** and **responsible innovation**.	
Tier 1: ✔ All departments and business units using or procuring AI tools. ✔ All AI systems supplied by third parties. ✔ All use cases where AI supports, influences, or replaces human decision-making.	**Tier 2 additional elements:** ✔ All AI systems that are developed internally.

2. Definitions

Term	Definition
Artificial Intelligence (AI)	A system that simulates human intelligence processes and performs tasks normally requiring human intelligence, such as learning, reasoning, problem-solving, and language understanding.
High-Risk AI	An AI system that poses serious risks to health, safety or the fundamental rights of protected groups
Generative AI	AI that creates content (e.g. text, images, audio, video), including LLMs, by learning patterns from existing data and generating original outputs.
Responsible AI Principles	Fairness, reliability, safety, privacy, security, inclusiveness, transparency, and accountability.

3. Policy Principles

All AI systems used by [Organisation Name] must adhere to six Responsible AI principles:

- ✔ **Fairness** – Prevent discriminatory or biased outcomes. Conduct fairness audits where appropriate.
- ✔ **Reliability & Safety** – Ensure systems are stress-tested and robust, with fallback mechanisms.
- ✔ **Privacy & Security** – Comply with applicable data protection laws. Conduct privacy impact assessments (PIAs), ensure lawful processing, and respect data subject rights.
- ✔ **Inclusiveness** – Design for accessibility and consider diverse user needs and impacts.
- ✔ **Transparency** – Inform users when AI is used. Ensure outputs are explainable where feasible.
- ✔ **Accountability** – Assign responsibility for AI decisions and outcomes. Maintain human oversight.

These principles reflect both legal and ethical imperatives, supporting long-term trust with customers, staff, and regulators.

4. Governance and Oversight

Tier 1:

- ✔ Board & Senior Management: Strategic oversight and policy approval.
- ✔ AI Governance Lead: Coordinates implementation, monitoring, and training. *[Note: A governance professional often facilitates this role.)*
- ✔ Legal & Risk: Draft and review policies. Ensure regulatory alignment. Remain aware of new regulation and risks.
- ✔ Business Units: Apply the Policy in operational settings and coordinate with Risk functions.
- ✔ Internal Audit: Periodically review policy compliance.

Tier 2 additional elements:

- ✔ Legal & Risk: Review and approve high-risk use cases.
- ✔ Internal Audit: Audit for adverse impacts of AI use.

5. Operational Controls

Tier 1:

- ✔ **AI Use Case Inventory:** Maintain a registry capturing:
  - Purpose, data used; and business owner.
- ✔ **Risk Assessment:** Assess AI systems for:
  - Ethical risks, bias, and explainability.
  - Data privacy.
- ✔ **Transparency and Notification:** Notify users when AI is involved in decisions.
- ✔ **Monitoring and Incident Reporting:**
  - Monitor models for fairness, accuracy, and data misuse
  - Report AI-related incidents to the AI Governance Lead and DPO within [x] days.
    [Note: Hong Kong is expected to adopt mandatory reporting in due course.]
- ✔ **Procurement and Third-Party AI: Vendors must disclose:**
  - Data usage and protection mechanisms.
  - Use of synthetic or identifiable data.
  - Contracts must include privacy terms, audit rights, and breach reporting clauses.
- ✔ **Human Oversight**
  - Staff must retain responsibility and intervene as needed.
  - AI must not make unreviewed decisions in critical contexts.

Tier 2 additional elements:

- ✔ **AI Use Case Inventory:** Maintain a registry capturing:
  - Privacy impact, cross-border data flow, and regulatory classification.
  - Lifecycle stage and responsible function.
- ✔ **Risk Assessment:** Assess AI systems for:
  - Data protection impact assessment (DPIA) outcomes.
  - Cross-functional review for high-risk use cases.
- ✔ **Transparency and Notification:** Notify users when AI is involved in decisions. Disclose:
  - The logic and potential impact of AI tools.
  - Rights to explanation, appeal, and human review.
  - Public performance metrics.
- ✔ **Monitoring and Incident Reporting**
  - Conduct safety evaluations on models for fairness, accuracy, and data misuse.

6. Training and Awareness

Practical training, whether licensed or internally developed, must include:

- ✔ Appropriate use of public AI tools (e.g. no confidential input).
- ✔ Verification of outputs before use.
- ✔ Awareness of limitations and biases.

Tier 1:

- ✔ Mandatory onboarding and refresher training (annually).
- ✔ Role-specific quick guides for frontline teams.

Tier 2 additional elements:

- ✔ Co-developed modules with HR, compliance, and IT functions.

7. Acceptable Use of AI by Staff and Contractors

- ✔ **Acceptable Use**
  - AI may enhance productivity, research, and customer service.
  - Avoid inputting sensitive data into public AI platforms.
  - AI-generated content must be human-validated before external use.
  - Disclose AI involvement in communications and decisions.

- ✔ **Prohibited Use**
  - Fabricated, defamatory, discriminatory, or misleading content.
  - Unauthorised impersonation.
  - Circumventing compliance or security controls.
  - Unreviewed automated decisions affecting rights.

- ✔ **Ownership and Oversight**
  - Staff remain responsible for AI-assisted work.
  - Professional responsibility cannot be delegated to AI tools.
  - Critical decisions must remain subject to human oversight.

8. Policy Review and Updates

Tier 1:

- ✔ Annual review cycle or upon:
  - Regulatory changes.
  - Significant incidents.
  - High-impact deployments.
- ✔ All updates require approval from the board or committee.
- ✔ Maintain version control and internal communications.

Tier 2 additional elements:

- ✔ Detailed policy review on AI applications and impact.

9. Policy Exceptions and Escalation

Tier 1:

- ✔ High-risk cases escalated to the board or oversight committee.

Tier 2 additional elements:

- ✔ Policy exceptions must be documented with justification.
- ✔ Legal and risk must review all exceptions.

10. Jurisdictional Compliance and Governance Standards (Expanded)	
Tier 1: ✔ **Jurisdiction / Key Law / Notes** • Hong Kong / PDPO / Data Protection Principles (DPPs), notice and consent requirements. • EU - GDPR / AI Act - DPIAs, data subject rights, and compliance with high-risk systems. • PRC / PIPL / Consent, data localisation, and regulatory filing obligations.	**Tier 2 additional elements:** ✔ **Jurisdiction / Key Law / Notes** • US (State) / CCPA, CPRA, VCDPA, etc. / Transparency, opt-out, and anti-discrimination clauses. • International Standards: ISO/IEC 42001, 23894, NIST AI Risk Management Framework

✔ **Reporting and Queries**

- Staff members must report any concerns, misuse, or policy breaches.
- Reports are sent to the AI Governance Lead, DPO, or the compliance function.
- A confidential channel should be available.

Serious violations may result in disciplinary action.

Conclusion

This Policy reflects [Organisation Name] 's commitment to using AI in an ethical, transparent, privacy-respecting, and aligned manner with our organisational purpose.

AI governance is not a destination; it is a continuous journey. This Policy is a tool to help navigate it with confidence, integrity, and foresight.

Note: Please follow the appointment of the responsible persons for policy implementation and other provisions of your organisation for consistency. The above are some sample suggestions, and there is no one-size-fits-all approach for any policy.

CHAPTER 5

DIRECTOR BRIEFING TEMPLATE

Developing credible AI policies is not just a compliance exercise: it is a strategic act of governance. The governance professional should anticipate the risks, frame the issues, and guide the process, but outcomes must be endorsed by the board and co-owned by senior management. We now turn to how the governance professional can facilitate directors in asking the right questions about AI implementation, interrogate key risks, and assess organisational readiness.

What Every Director Needs to Know About AI: A Practical Governance Briefing
Prepared by: [Company Secretary/Governance Professional]

AI Oversight Is About Readiness and Trust

A reasonable director is not expected to understand how an algorithm works in code, but they are expected to ensure the organisation is equipped to use AI responsibly and in accordance with the law. That includes:

- ✔ Understanding purpose.
- ✔ Clarifying risk.
- ✔ Setting expectations.
- ✔ Monitoring accountability.
- ✔ Supporting transparency.

By asking the right questions and relying on governance professionals to facilitate sound oversight, the board can ensure that AI becomes an asset—not a liability—to the organisation's future.

1. Start With the Right Question: What Is the AI Being Deployed For?

As a director, you **need to know where AI is being used in your organisation** and **what it's being used for**. This is the foundation for effective board oversight.

Ask:

- ✔ What decisions or processes are being influenced, supported, or made by AI?
- ✔ Who owns each AI system or use case?
- ✔ What business problem is it trying to solve — and why use AI to solve it?

These questions help determine whether the deployment is routine (e.g., email sorting), sensitive (e.g., recruitment filtering), or high-risk (e.g., credit assessments or predictive policing).

2. Understand the Real Risks — Legal, Reputational, and Operational

AI is not inherently safe or neutral. Risks arise depending on how AI is trained, applied, and governed. Directors should understand and probe the following (note: this is not an exhaustive list of risks):

- ✔ Bias and unfairness – Does the AI treat certain individuals or groups unfairly?
- ✔ Lack of transparency – Can we explain how the AI arrives at its decisions?
- ✔ Data misuse – Is personal or sensitive data being used lawfully and ethically?
- ✔ System failure – What happens if the AI fails or produces harmful outputs?
- ✔ Lack of accountability – Who is ultimately responsible for decisions made using AI?

These risks can lead to public backlash, regulatory fines, loss of stakeholder trust, or strategic damage, depending on the location of the business operations and the applicable laws and regulations.

3. AI Governance Is About Principles, Not Just Technology

Your organisation should have a formally adopted AI Policy. It should not be just an IT policy, but a governance framework shaped around the following six principles:

- ✔ Fairness – Avoid discrimination or bias.
- ✔ Reliability and Safety – Ensure the system performs as intended, even under pressure.
- ✔ Privacy and Security – Comply with laws, protect personal data.
- ✔ Inclusiveness – Serve all user groups appropriately.
- ✔ Transparency – Make systems explainable to users and regulators.
- ✔ Accountability – Ensure human responsibility is never outsourced to a machine.

Directors should be familiar with these principles and ensure that they are reflected in their policies, risk management practices, and organisational culture.

4. Expect to See an Inventory of AI Use Cases

A reasonable director should ask: "Can you show me where AI is currently used in our business?"

A responsible organisation should be able to produce a use case inventory showing:

- ✔ Business purpose of the AI system.
- ✔ System owner and data used.
- ✔ Risk level (e.g. high-impact, regulatory sensitive).
- ✔ Whether fairness, privacy, and explainability controls are in place.

This is the modern equivalent of knowing your organisation's financial systems or major contracts — it is about visibility and control.

5. Know Who Is Accountable Internally

There should be named individuals responsible for:

- ✔ Oversight of AI risk and compliance (often a governance lead or committee).
- ✔ Reviewing high-risk AI deployments.
- ✔ Reporting incidents or policy exceptions.
- ✔ Coordinating across legal, IT, business units, and data protection.

As a board member, you should not assume technical teams are "handling it". Ask:

- ✔ Is there clear ownership for each AI system?
- ✔ Who reports to the board on AI risk and performance?
- ✔ What happens when things go wrong?
- ✔ Are there examples of AI issues/risks that have emerged so far and if so, how have they been handled?

6. Board's Role: Ask the Right Questions and Expect the Right Information

Some practical questions to ask at board or committee meetings:

- ✔ Governance & Accountability:
  - What are our top 5 AI use cases, and how do they support our strategy?
  - When was our AI Policy last reviewed, and who owns it?
  - Who is responsible for each AI system across its lifecycle?
  - What are the board-level triggers for escalation in the event of an AI-related incident or risk?
  - How are governance bodies (e.g. ethics boards, risk committees) empowered to intervene in high-risk AI decisions?
  - What training or capacity-building measures are in place to ensure responsible AI ownership across the organisation?
  - How is leadership and the board regularly informed about AI performance, risks, and oversight outcomes?
- ✔ Risk & Trade-Offs:
  - What standards and internal thresholds guide decisions around fairness, privacy, safety, and inclusion in AI systems?
  - How are trade-offs between model performance, explainability, and broader societal impact evaluated and documented?
  - How are risks to vulnerable or excluded groups assessed throughout the AI lifecycle?
  - What indicators or thresholds trigger escalation, rollback, or redesign of a system?
  - How is privacy risk assessed during model design, training, and deployment?
  - Have we assessed whether any systems are considered high-risk under new laws (e.g. EU AI Act, Chinese mainland measures)?
  - Are we deploying any AI in ways that affect individual rights, regulatory compliance, or reputation?
- ✔ Assurance & Audit:
  - What independent audits, red-teaming exercises, or stress tests are conducted to validate the robustness, fairness, and safety of AI systems?
  - How are fairness claims and performance metrics independently verified?
  - Are our third-party models, data sources, and tools vetted for alignment with our internal standards on fairness, privacy, security, and inclusion?
  - What security controls are in place to protect models, bias, failure, data pipelines, and inference processes from adversarial attacks?
  - What documentation and controls support continuous monitoring of compliance with regulatory and ethical expectations?
  - Are staff trained on the safe and responsible use of AI?
- ✔ Transparency & Explainability:
  - How are explainability requirements tiered based on model impact or risk (e.g., high-stakes vs. low-stakes)?
  - Can we clearly and meaningfully explain AI-driven decisions to both internal stakeholders and external parties, including affected individuals?
  - What documentation is maintained throughout the AI system's lifecycle to support transparency, accountability, and future audits?
  - Are trade-offs between performance and interpretability explicitly recorded and justified?
  - Do we disclose the use, limitations, and risks of AI systems to users, partners, and regulators?
  - What mechanisms exist for impacted individuals to question, appeal, or seek redress for AI-driven outcomes?

- ✔ Inclusion & User Impact:[55]
  - Are AI systems tested across diverse user groups before deployment, particularly those at risk of exclusion?
  - Are model performance metrics (e.g. precision, recall) disaggregated by protected characteristics such as race, gender, language, and disability?
  - What benchmarks or frameworks guide our approach to inclusive AI design and deployment?
  - Are accessibility and cultural considerations integrated into product development and model evaluation?
  - How are inclusive practices embedded into hiring, team composition, and vendor selection for AI projects?
  - Are exclusionary impacts monitored post-deployment, and how are findings acted upon?
- ✔ Incident Management:
  - Are incident response and recovery plans in place for AI system failures, including clear roles and cross-functional coordination?
  - How are findings from model audits, complaints, near-misses, or observed harms integrated into system redesign, oversight, and board reporting?
  - When exclusionary or harmful outcomes are detected, what remediation pathways and escalation processes are followed?
  - How are breach detection, response, and notification protocols adapted for AI-specific risks?

7. How the Company Secretary/Governance Professional Helps the Board

Company secretaries, general counsels and governance professionals are not AI developers — but they are the facilitators of responsible governance. They support the board by:

- ✔ Developing and maintaining the AI Policy.
- ✔ Coordinating the cross-functional AI use case inventory.
- ✔ Tracking regulatory developments and industry benchmarks.
- ✔ Facilitating training and awareness across functions.
- ✔ Ensuring board visibility of incidents, exceptions, and lessons learned.

You can expect your governance team to help translate complex technical risks into governance language — and frame the right issues for board review.

8. Next Steps for the Board

- ✔ Confirm that the organisation has adopted a fit-for-purpose AI Policy.
- ✔ Request a summary of current AI use cases and associated risk controls.
- ✔ Ensure AI governance is reviewed annually and monitored through appropriate board committees.
- ✔ Ask for regular briefings on new regulations, incidents, and major deployments.
- ✔ Encourage management to benchmark AI governance practices against industry norms.

CHAPTER 6
CONCLUSION

Thank you for taking the time to engage with **HKCGI's Responsible AI Governance Playbook**. We encourage you to apply the tools, frameworks, and engage in the activities described to embed responsible AI practices into your governance processes.

While governance may not be newsworthy, it's often what prevents negative headlines. By translating ethical principles into clear policies, documenting decisions transparently, and ensuring accountability as AI technologies and business circumstances evolve, your work strengthens organisational resilience, public trust, and long-term stakeholder value.

AI governance is not a one-time initiative. Continue to review, refine, and strengthen your practices. Incremental improvements made consistently will do more than a broad overhaul done once. We welcome your feedback and insights to shape future HKCGI resources. Together, we can ensure AI technologies serve not only organisational objectives, but society at large.

REFERENCES AND FURTHER READING

[1] See Collina, Luca, et al. "Critical Issues about A.I. Accountability Answered." *California Management Review Insights*, 6 Nov. 2023, cmr.berkeley.edu/2023/11/critical-issues-about-a-i-accountability-answered/.

[2] See Bender, Emily, et al. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *FAcc'21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 1 Mar. 2021, pp. 610–623, faculty.washington.edu/ebender/papers/Stochastic_Parrots.pdf, https://doi.org/10.1145/3442188.3445922.

[3] See Schneider, Johannes, et al. "Governance of Generative Artificial Intelligence for Companies." *ArXiv.org*, 5 Feb. 2024, arxiv.org/abs/2403.08802.

[4] See Bird & Bird. "AI Governance: Essential Insights for Organisations: Part I – Understanding Meaning, Challenges, Trends, and Best Practices in AI Governance." *Twobirds.com*, 2025, www.twobirds.com/en/insights/2025/ai-governance-essential-insights-for-organisations-part-i--understanding-meaning-challenges-trends-a.

[5] See Hickman, Eleanore, and Martin Petrin. "Trustworthy AI and Corporate Governance: The EU's Ethics Guidelines for Trustworthy Artificial Intelligence from a Company Law Perspective." *European Business Organization Law Review*, vol. 22, no. 4, 6 Oct. 2021.

[6] See Mökander, Jakob, et al. "Challenges and Best Practices in Corporate AI Governance: Lessons from the Biopharmaceutical Industry." *Frontiers in Computer Science*, vol. 4, 10 Nov. 2022, https://doi.org/10.3389/fcomp.2022.1068361.

[7] See [3]

[8] See [3]

[9] See Olteanu, Alexandra, et al. "Rigor in AI: Doing Rigorous AI Work Requires a Broader, Responsible AI-Informed Conception of Rigor." *ArXiv.org*, 2025, arxiv.org/abs/2506.14652.

[10] See [4]

[11] See Camilleri, Mark Anthony. "Artificial Intelligence Governance: Ethical Considerations and Implications for Social Responsibility." *Expert Systems*, vol. 41, no. 7, 18 July 2023, https://doi.org/10.1111/exsy.13406

[12] Ibid.

[13] See [2]

[14] See Raji, Inioluwa Deborah, et al. "Closing the AI Accountability Gap: Defining an End-To-End Framework for Internal Algorithmic Auditing." *ArXiv:2001.00973 [Cs]*, 3 Jan. 2020, arxiv.org/abs/2001.00973.

[15] See White & Case. "AI Watch: Global Regulatory Tracker - Hong Kong | White & Case LLP." *Whitecase.com*, 6 June 2025, www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-hong-kong

[16] Ibid.

[17] Privacy Commissioner's Office (Hong Kong). "Publishes Guidance on Ethical Development and Use of AI and Inspection Report on Customers' Personal Data Systems of Two Public Utility Companies." *Pcpd.org.hk*, 18 Aug. 2021, www.pcpd.org.hk/english/news_events/media_statements/press_20210818.html.

[18] Privacy Commissioner's Office (Hong Kong). "Privacy Commissioner's Office Publishes "Artificial Intelligence: Model Personal Data Protection Framework."" *Pcpd.org.hk*, 11 June 2024, www.pcpd.org.hk/english/news_events/media_statements/press_20240611.html.

[19] Digital Policy Office. "Ethical Artificial Intelligence Framework | Digital Policy Office." *Digitalpolicy.gov.hk*, 25 July 2024, www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/ethical_ai_framework/.

[20] Han, Sirui, et al. "Hong Kong Generative Artificial Intelligence Technical and Application Guideline." *SSRN Electronic Journal*, 2025, www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/ethical_ai_framework/doc/HK_Generative_AI_Technical_and_Application_Guideline_en.pdf, https://doi.org/10.2139/ssrn.5285630.

[21] Privacy Commissioner's Office (Hong Kong). "PCPD Has Completed Compliance Checks on 60 Organisations to Ensure AI Security." *Pcpd.org.hk*, 8 May 2025, www.pcpd.org.hk/english/news_events/media_statements/press_20250508.html.

[22] HK Monetary Authority. "HKMA Banking Regulatory Document Repository." *Hkma.gov.hk*, 2024, brdr.hkma.gov.hk/eng/doc-ldg/docId/20241122-3-EN.

[23] HK Department of Health. *Medical Device Administrative Control System (MDACS) Artificial Intelligence Medical Devices (AI-MD)*. 3 Jan. 2024.

[24] HK Insurance Authority. *Conduct in Focus*. 7 May 2023, www.ia.org.hk/en/legislative_framework/files/Eng_Conduct_in_Focus_7_May_23.pdf

[25] *OECD*. "AI Principles." OECD, 2024, www.oecd.org/en/topics/sub-issues/ai-principles.html.

[26] *Infocomm Media Development Authority*. "Singapore Model AI Governance Framework." Infocomm Media Development Authority, 16 Jan. 2024, www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2024/public-consult-model-ai-governance-framework-genai

[27] *European Parliament*. "EU AI Act: First Regulation on Artificial Intelligence." European Parliament, 8 June 2023, www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

[28] European Parliament. "The General-Purpose AI Code of Practice." *Shaping Europe's Digital Future*, 2025, digital-strategy.ec.europa.eu/en/policies/contents-code-gpai.

[29] National Institute of Standards and Technology. "AI Risk Management Framework." *NIST*, 12 July 2023, www.nist.gov/itl/ai-risk-management-framework.

[30] International Organisation for Standardisation. "ISO/IEC 42001:2023." *ISO*, 2023, www.iso.org/standard/42001.

[31] International Organisation for Standardisation. "ISO/IEC 23894:2023." *ISO*, Feb. 2023, www.iso.org/standard/77304.html.

[32] *Ashurst*. "New Generative AI Measures in China." Ashurst, 26 Sept. 2023, www.ashurst.com/en/insights/new-generative-ai-measures-in-china/.

[33] Microsoft. "Responsible AI Principles and Approach | Microsoft AI." *Www.microsoft.com*, 2024, www.microsoft.com/en-us/ai/principles-and-approach.

[34] Ibid.

[35] See [3].

[36] See [11].

[37] See [5].

[38] See Huang, Lei, et al. "A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions." *ArXiv (Cornell University)*, 9 Nov. 2023, https://doi.org/10.48550/arxiv.2311.05232.

[39] See [4]

[40] See [5]

[41] PCPD (2012). The Personal Data (Privacy) Ordinance. [online] Pcpd.org.hk. Available at: https://www.pcpd.org.hk/english/data_privacy_law/ordinance_at_a_Glance/ordinance.html.

[42] See Centre for Emerging Technology and Security. (2023). Adversarial AI: Coming of age or overhyped? [online] Available at: https://cetas.turing.ac.uk/publications/adversarial-ai-coming-age-or-overhyped.

[43] HK Stock Exchange. "Exchange Publishes Consultation Paper on Corporate Governance Code Enhancements." *Hkex.com.hk*, 2024, www.hkex.com.hk/News/Regulatory-Announcements/2024/240614news?sc_lang=en.

[44] See [14].

[45] See [27].

[46] See Weidinger, Laura, et al. "Sociotechnical Safety Evaluation of Generative AI Systems." *ArXiv (Cornell University)*, 18 Oct. 2023, https://doi.org/10.48550/arxiv.2310.11986.

[47] See Gesser, Avi. "Good AI Vendor Risk Management Is Hard, but Doable – Debevoise Data Blog." Debevoisedatablog.com, 26 Sept. 2024, www.debevoisedatablog.com/2024/09/26/good-ai-vendor-risk-management-is-hard-but-doable/. Accessed 1 Aug. 2025.

[48] See [14]

[49] Mitchell, Margaret, et al. "Model Cards for Model Reporting." *Proceedings of the Conference on Fairness, Accountability, and Transparency* - FAT* '19, 2019, pp. 220–229, https://doi.org/10.1145/3287560.3287596.

[50] See [14].

[51] See [14].

[52] See [14].

[53] Burns, Mary, et al. "Imagining Failure to Attain Success: The Art and Science of Pre-Mortems." *Brookings*, 6 Feb. 2025, www.brookings.edu/articles/the-art-and-science-of-pre-mortems/.

[54] Halme, Erika, et al. "How to Write Ethical User Stories? Impacts of the ECCOLA Method." *Lecture Notes in Business Information Processing*, 2021, pp. 36–52, https://doi.org/10.1007/978-3-030-78098-2_3. Accessed 12 Aug. 2021.

[55] Mitchell, Margaret, et al. "Diversity and Inclusion Metrics in Subset Selection." *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 4 Feb. 2020, https://doi.org/10.1145/3375627.3375832.

The Hong Kong Chartered Governance Institute 香港公司治理公會

(Incorporated in Hong Kong with limited liability by guarantee)

The Hong Kong Chartered Governance Institute (HKCGI) is the sole accrediting body in Hong Kong and the Chinese mainland for the globally recognised Chartered Secretary and Chartered Governance Professional qualifications. Formerly known as The Hong Kong Institute of Chartered Secretaries (HKICS), HKCGI is the Hong Kong/China Division of The Chartered Governance Institute (CGI).

With a legacy of over 76 years, HKCGI has established itself as a trusted and reputable professional body in the region. Its influence extends to CGI's global network of around 40,000 members and students, making it one of its fastest-growing divisions. HKCGI's community comprises about 10,000 members, graduates, and students, with significant representation in listed companies and diverse governance roles across various industries.

Guided by the belief that governance leads to better decision-making and a better world, HKCGI is committed to advancing governance in commerce, industry, and public affairs. It achieves this through education, thought leadership, advocacy, and active engagement with its members and the broader community. As a recognised thought leader, HKCGI promotes the highest standards of governance while advocating for an inclusive approach that considers the interests of all stakeholders, and ensures that every voice is heard and valued.

Better Governance. Better Future.
For more information, please visit www.hkcgi.org.hk.

關於香港公司治理公會
(於香港成立的有限擔保公司)

香港公司治理公會（前稱「香港特許秘書公會」）（公會）是特許公司治理公會（國際總會）的中國香港屬會，同時亦是中國內地和香港地區唯一頒授獲國際廣泛認可的「特許秘書」和「公司治理師」專業資格的機構。

公會成立迄今已有逾 76 年的歷史，其專業地位於中國內地及香港地區備受信賴及尊崇。公會的影響力擴展至國際總會的所有屬會，涵蓋全球約 40,000 名會員和學員，亦成為增長最快的屬會之一。公會現擁有約 10,000 名會員、畢業學員及學員，他們在上市公司和各行各業中擔任重要的治理角色。

作為治理領域的思想領導者，公會始終秉承「卓越治理帶來更佳決策，從而創造更美好世界」的理念，致力於通過教育、思維領導、倡導工作，以及與會員和廣泛社會層面的互動交流，提升工商業以及公共事務的治理水平，並促進最高治理標準。同時，公會提倡考慮所有持份者的利益，確保各種寶貴意見及建議都被聽取和重視。

卓越治理 更佳未來

如欲了解更多資訊，請瀏覽：www.hkcgi.org.hk。

CONTACT US

The Hong Kong Chartered Governance Institute 香港公司治理公會
(Incorporated in Hong Kong with limited liability by guarantee)

Hong Kong Office

3/F, Hong Kong Diamond Exchange Building, 8 Duddell Street, Central, Hong Kong
Tel: (852) 2881 6177
Fax: (852) 2881 5050
Email Address: ask@hkcgi.org.hk
Website: www.hkcgi.org.hk

Beijing Representative Office

Room 1220, Jinyu Tower, No. 129, Xuanwumen West Street, Xicheng District, Beijing, 100031, P.R.C
Tel: (86 10) 6641 9368/6641 9190
Email Address: bro@hkcgi.org.hk
Website: www.hkcgi.org.cn

Ellie Pang FCG HKFCG(PE)

Chief Executive
The Hong Kong Chartered Governance Institute
T: +852 2830 6029
E: ellie.pang@hkcgi.org.hk

Mohan Datwani FCG HKFCG(PE)

Deputy Chief Executive
The Hong Kong Chartered Governance Institute
T: +852 2830 6012
E: mohan.datwani@hkcgi.org.hk