

Human Technology Institute



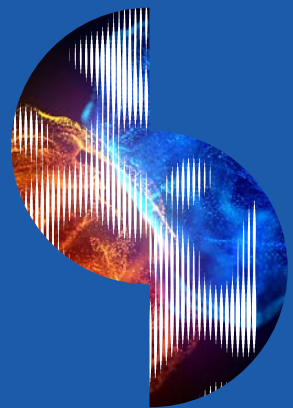
Essential Components of AI Governance

AI governance snapshot #1

Shaping our Future

Symposium 2024

31 January 2024



Essential Components of AI Governance

AI governance snapshot #1

This document is the first in a series of Artificial Intelligence (AI) Governance Snapshots from the UTS Human Technology Institute (HTI). HTI has prepared these governance snapshots as part of the *AI Corporate Governance Program* to help build greater shared understanding and awareness of the practical steps that organisations seeking to transform AI governance practice can take.

These snapshots are living documents which evolve to incorporate the latest research and emerging trends in responsible AI governance transformation.

What is AI governance?

AI governance is the system by which an organisation manages its development and use of AI systems, and includes a wide range of governance structures, policies, skills, practices, and other mechanisms that guide AI use, monitoring, and management.

AI is increasingly becoming central to how organisations operate. But, as HTI outlined in *The State of AI Governance in Australia*, effectively managing the unique characteristics of AI systems requires special governance attention.

AI governance is an emerging field that explores ways that organisations can evolve existing corporate governance approaches and introduce new strategies to realise the benefits of AI systems while addressing the risk of AI harms to the organisation, individuals, other organisations, and society at large. It encompasses organisation-wide governance approaches as well as project-specific and system-level interventions designed to ensure that AI systems are accurate, accountable, fair and fit-for-purpose.

Building strong AI governance competency is a vital task for corporate leaders who are making strategic AI investment decisions, looking to secure benefits such as productivity gains and competitive advantage, while mitigating commercial, regulatory and reputational risk.

How does AI governance relate to strategy?

AI governance should be guided by an organisation's overarching strategy. Given the rising importance of AI as a driver of productivity, efficiency and customer value today, all organisations should ensure that their strategy sets expectations for where, how and when AI can and should be leveraged to create sustainable value for stakeholders. In doing so, organisational strategy should also acknowledge the resources, estimated costs, potential harms and range of risks involved, and establish an AI risk appetite as a critical reference point for AI investment, system design and decision-making.

What are the essential components of an AI governance framework?

Based on HTI's investigation of governance practice locally and internationally, we have identified eight essential components of AI governance. These include accountability, oversight, roles and responsibilities; governance structures; people, skills, values, and culture; principles and policies, practices, processes and controls; supporting infrastructure; stakeholder engagement, co-design and impact assessment; monitoring, reporting and evaluation.

These components and their interactions may be documented in an enterprise governance framework. HTI's work with corporate leaders across Australia strongly suggests that organisations must take a holistic approach and address all components to create an effective system of AI governance.



Essential Components of AI Governance

AI governance snapshot #1

1. Accountability, oversight, roles and responsibilities

This set of factors – accountability, oversight, roles and responsibilities – deals with how and in what form responsibility for AI systems is vested in individuals or groups of decision-makers across the organisation. This can include the role that directors and boards play at the highest levels of corporate governance, and where management accountability and decision authority lie for specific AI systems. Usually, this will take the form of a hierarchy of responsibilities for AI systems that map to the organisation’s chosen governance structure.

For example, as part of its strategy-aligned AI governance approach, Telstra has ensured that all relevant team members – including project teams, business managers, risk and compliance teams, senior management and relevant board committees – have clear and set roles and responsibilities regarding oversight of the risks and opportunities of AI systems.

Addressing this component may include:

- Establishing the individuals and positions in the organisation best placed to:
 - > make critical decisions about AI system procurement, management and ongoing oversight; whose responsibility it is to ensure that systems operate as intended; and who is held accountable for any resulting harms
 - > assess different forms and sources of harms, risks and liability that could flow from AI systems deployed in line with the organisation’s strategy.

2. Governance structures

AI governance structures aim to formalise AI accountability and allocate responsibilities across organisational functions, positions or specialised groups – often in the form of boards, project teams, advisory groups, councils, committees or other bodies. These individuals or groups are charged with governance-related tasks such as collecting information, reviewing progress and making decisions in relation to AI systems.

AI governance structures should be seen as part of an organisation’s overarching enterprise governance framework and reporting structures, committee(s) structures and charters.

For example, Telstra created the Risk Council of AI & Data (RCAID) committee as just one such structure within its enterprise governance framework. This body is responsible for overseeing and giving practical life to Telstra’s Responsible AI policy, and is charged with reviewing and approving all current and proposed AI systems that have significant stakeholder impact.

Addressing this component may include:

- Understanding and updating the enterprise governance framework to acknowledge the structures responsible for AI systems
- Establishing or updating board and management committees and advisory structures, including ensuring that committee charters and composition are fit-for-purpose.

3. People, skills, values, and culture

AI governance is ultimately achieved by people working together to ensure that all AI systems operate effectively, responsibly and legally in the way that the organisation intends. Effective governance therefore requires an organisation to have the right people with relevant skills, aligned values and a supportive culture across the AI system lifecycle, from design to decommissioning. While many organisations have prioritised the acquisition of technical skills for AI, this component suggests that possessing strategic skills related to AI systems and building an appropriately diverse AI culture are equally important.

This component of AI governance is one of the most wide-ranging and complex, encompassing both competencies and values. Thanks to the proliferation of third-party AI services and the extended ‘value chain’ of AI system, issues related to people, skills, values and culture affect employees at all organisational levels and will often extend to contractors, suppliers and technology partners. This area also has special relevance for the composition, skills and culture of company directors.



Essential Components of AI Governance

AI governance snapshot #1

For example, to ensure that effective guardrails were embedded throughout KPMG Australia’s internal generative AI chatbot, KymChat, the development of this AI system was led by a multi-disciplinary team with deep expertise in data governance and privacy. This ensured that KPMG was able to carefully and successfully navigate challenges related to client confidentiality, intellectual property and data quality.

Addressing this component may include:

- Ensuring that accountable and responsible individuals and teams possess the appropriate level of strategic understanding and governance skills relevant to AI systems
- Reviewing performance management and incentive structures and their influence on the use and management of AI systems
- Promoting a culture that emphasises both opportunities and risks related to AI systems and supports the organisation’s AI risk appetite
- Adopting a people strategy which recognises the benefits of a diverse workforce in effectively identifying and managing AI risk
- Ensuring organisational values support a culture of responsible and accountable use of AI.

4. Principles and policies

AI policies are increasingly used by organisations to define and communicate the guiding principles and overarching approach of the organisation in relation to its development and use of AI systems. AI principles and policies often outline the “why” and the “what” of AI use in an organisation.

HTI’s research indicates that defining high-level principles related to AI systems is a necessary but insufficient step towards AI governance. It is therefore critical that organisations ensure that AI principles are translated into specific policies that support staff to identify “red-lines” that should not be crossed during AI system procurement, design, development or deployment, as well as creating safe spaces for experimentation, innovation and value-creation.

For example, UTS has its *Artificial Intelligence Operations Policy* that guides the use, procurement, development and management of AI at UTS.ⁱ

Addressing this component may include:

- Translating organisational values into clear, useful principles that describe how the organisation makes choices around AI systems
- Developing policies for AI system procurement, design and use, including the conditions under which different responsibilities arise, the role that governance structures play and the ‘safe spaces’ where responsible innovation is encouraged
- Specifying at a high level the chosen risk, compliance and reporting frameworks which give effect to the policies at an organisational and project level.

5. Practices, processes and controls

If principles and policies are the “why” and the “what” of AI governance, practices, processes and controls are the “who”, “when” and “how”. These are very practical tools, steps, design elements or other requirements that – when used appropriately – give effect to the policies and ensure that both people and AI systems behave in line with the overarching principles.

Although AI governance is still an emerging field, organisations can draw on AI-specific governance practices and processes in the form of international standards developed by global experts. These include *ISO/IEC 42001*,ⁱⁱ the recently published standard that specifies requirements for establishing, implementing, maintaining and continually improving an Artificial Intelligence Management System (AIMS), and the *National Institute of Standards and Technology AI Risk Management Framework*.ⁱⁱⁱ

Practices tend to offer practical guidance that helps individuals and teams discharge their duties in line with AI policies. Processes are often more detailed descriptions of system or team behaviour, while controls represent the set of checks and balances that mitigate against the risks of AI systems.

For example, Atlassian has created a *Responsible Technology Review Template*, designed to be used by teams to document critical information related to the purpose, design, risks and stakeholder impact of AI systems.^{iv}



Essential Components of AI Governance

AI governance snapshot #1

Addressing this component may include:

- Developing specific guidance and practices for teams directly involved in the procurement, design, development and use of AI systems. This includes detailing the responsible people, processes and safeguards that give effect to stated policies.
- Detailing how established risk management approaches and standards (for example, ISO 31000-based systems) and compliance frameworks apply to AI systems, and where these need to be augmented by AI-specific practices

6. Supporting infrastructure

Supporting infrastructure is a broad category that encompasses the data sources, data attributes, technology platforms and related systems required to support and deliver the required governance practices. The most common supporting infrastructure will relate to cyber security and data governance and may include ensuring that all data ingested into an AI system has been validated in terms of data lineage, meta-data and tagging.

For example, Telstra has undertaken work to clean up, simplify and modernise its data platforms (including decommissioning and consolidating legacy systems) to reduce the risk of data breaches and so that staff can more easily and safely access the data that they need.

Addressing this component may include:

- Undertaking separate AI system and data inventories to understand what is in use where across the organisation
- Ensuring robust data governance practices across all systems, especially those offered by third parties
- Applying and testing cyber security policies to AI systems to ensure their resilience and robustness.

7. Stakeholder engagement, co-design and impact assessment

Stakeholder engagement in the context of AI governance involves organisations identifying, understanding and responding to individuals and groups that may be impacted by AI in order to ensure that AI systems ultimately serve stakeholder needs. It requires corporate leaders and organisations to recognise their responsibilities to shareholders, customers, the environment, regulators and the broader community, and establish mechanisms for co-design, harm identification, risk management and testing.

For example, to determine the principles that should govern the use of analytics and AI at UTS, UTS undertook a consultation process with UTS students, casual tutors and academics, using the principles of *Deliberative Democracy*.^v

Addressing this component may include:

- Taking a human-centered approach, and consciously designing inclusivity, accessibility and diversity into AI systems from the earliest stages of system development
- Establishing processes to engage customers, employees and other affected individuals with their full consent at both an organisational and AI system level to effectively identify, assess, monitor and respond to potential AI risks
- Engagement with regulators and guidance bodies.



Essential Components of AI Governance

AI governance snapshot #1

8. Monitoring, reporting, and evaluation

When compared to traditional software, ongoing testing, monitoring, reporting and evaluation of AI systems is critical for AI governance.

It is important to ensure that the most relevant measures and metrics related to the desired outcomes and potential harms of AI systems are actively monitored over time, and that relevant individuals and groups respond effectively when signs of poor performance or unintended consequences become evident.

For example, KPMG Australia has evaluated the impact of KymChat through a modified version of the *Microsoft Responsible AI Impact Assessment Template*.^{vi} This template provides a structured way to consider the potential benefits and harms of AI systems on stakeholders, including issues like accountability, transparency, reliability, privacy and inclusiveness.

Addressing this component may include:

- Designing live or periodic monitoring and reporting systems, including automated performance assessment
- Establishing internal and external audit systems
- KPI reporting frameworks and frequency.

- i. University of Technology Sydney, Artificial Intelligence Policy (Policy, 9 June 2023) <<https://www.uts.edu.au/about/uts-governance/policies/uts-policy/artificial-intelligence-operations-policy>>.
- ii. International Organisation for Standardization, ISO/IEC DIS 42005 Information Technology Artificial Intelligence AI system impact assessment (Web page) <<https://www.iso.org/standard/44545.html>>.
- iii. National Institute of Standards and Technology (NIST), Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1, January 2023) <<https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf>>.
- iv. Atlassian, Responsible Technology Review Template (Template, 2023) <<https://www.atlassian.com/trust/responsible-tech-principles>>.
- v. Simon Buckingham Shum, The UTS “EdTech Ethics” Deliberative Democracy Consultation: Rationale, Process and Outcomes (Connected Intelligence Centre, University of Technology Sydney Report, 31 January 2022) <<https://cic.uts.edu.au/projects/edtech-ethics>>.
- vi. Microsoft, Responsible AI Impact Assessment Template (Template, June 2022) <<https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2022/06/Microsoft-RAI-Impact-Assessment-Template.pdf>>.

Bringing it all together – improving accountability and building trust

HTI’s early research underscores the importance of organisations integrating all of these key components of AI governance into existing corporate governance mechanisms. To achieve this, corporate leaders must be firmly focused on the practical expression of both organisational strategy and values across AI systems, while carefully managing potential harms and risks in line with the organisation’s risk appetite.

