AI Governance – Principles and Best Practices

Dr Philip Nesci

August 2025

Agenda

- Why AI Governance is Critical Now

- AI Governance vs. Traditional Governance

- Key Challenges in AI Governance

- AI Governance Frameworks

- Best Practice Processes for AI Governance

- Taking the First Governance Steps

- Q and A

Quick Straw Poll

My organisation has one of the following AI governance profiles:

- ☐ **No Governance**

- ☐ **Limited Governance**

- ☐ **Fit for Purpose Governance**

Why AI Governance is Critical Now

- AI adoption is accelerating across all industries

- Growing Impact: AI systems can make decisions that directly affect individuals, businesses and society (e.g., healthcare, finance, employment)

- Significant new risk landscape with AI adoption, unintended consequences

  ○ Regulatory fines (e.g. GDPR, EU AI Act violations)

  ○ Discrimination in automated decisions, 'hallucinations' damaging brand credibility

  ○ Data privacy breaches and IP leakage

  ○ Algorithmic failures causing real-world harm

- Boards are increasingly being held accountable for oversight

- Good governance promotes trust and brand value for an organisation
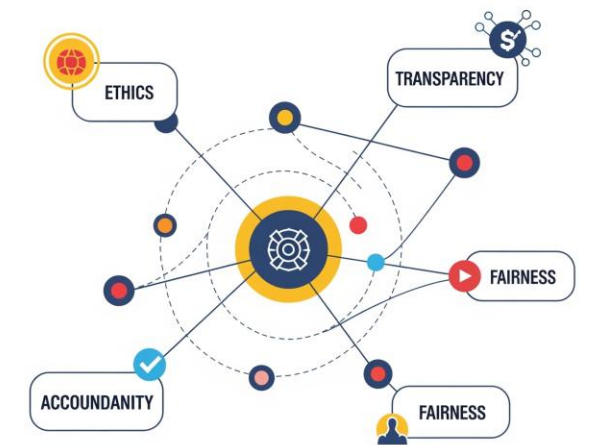
Two Recent Examples

- Bunnings
    - Use of in-store facial recognition to reduce theft and protect staff
        - Is it ethical and legal?
        - How do you ensure accuracy?
        - Bias?
        - Data protection and privacy?
        - How have they gone about engaging and ensuring trust of customers?
- McDonalds
    - Use of third party AI platform to streamline recruitment resulting in data breach
    - Bias? Data breach response? Candidate concerns?

AI Governance vs Traditional Governance

|	Traditional Governance (e.g. IT, Data)	AI Governance
Core Focus	Managing relatively static assets (data, IT infrastructure), compliance with established rules.	Managing dynamic, evolving AI systems Focusing on ethical, societal, and emergent risks. AI evolves after deployment- model drift
System Type	Rule-based, predictable systems.	Decisions tend to be probabilistic. Outputs can be opaque and hard to explain.
Primary Concerns	Data quality, security, compliance, performance, operational efficiency.	Bias, fairness, transparency, accountability, human oversight, societal impact.
Oversight	Periodic audits, established controls.	Continuous monitoring, real-time assessment of model behaviour and impact.
Regulatory Landscape	Established regulations (GDPR, SOX, HIPAA).	Evolving, AI-specific regulations (EU AI Act, NIST AI RMF), often principles-based.

Navigating AI Governance is Not Easy

- **Pace of Innovation:** AI technology evolves faster than governance frameworks can adapt.

- **Complexity and Opacity:** Understanding the inner workings of complex AI models can be challenging.

- **Lack of Expertise:** Shortage of individuals with combined AI technical knowledge and governance expertise.

- **Defining Accountability:** Difficulty in assigning clear responsibility for AI-driven outcomes.

- **Ethical Dilemmas:** Navigating subjective ethical considerations (e.g., fairness metrics).

- **Cultural Resistance:** Fear of job displacement, lack of trust in AI and inertia to change within organisations.

- **Data Quality and Bias:** Ensuring training data is representative, unbiased and of high quality.

- **Measuring Impact:** Quantifying the social and ethical impact of AI systems, business benefits.

Challenges of AI Governance
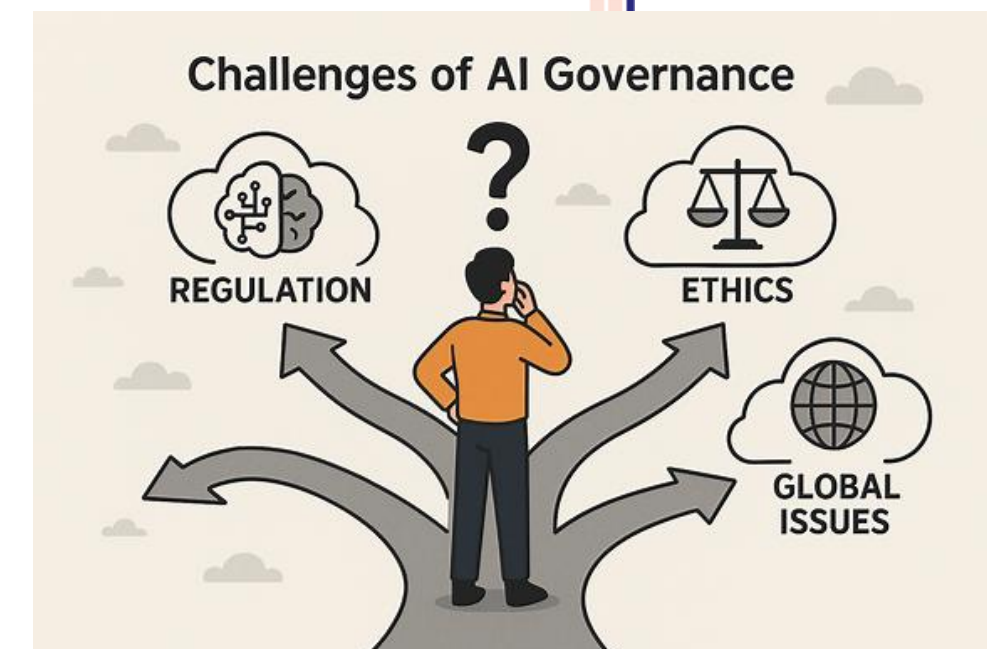
Governance Frameworks – What Are They?

AI governance frameworks provide structured approaches to guide the responsible use and oversight of AI across its lifecycle. They typically include:

- **Principles:** Overarching values (e.g., fairness, transparency, accountability).

- **Policies:** Specific rules and guidelines for AI development and deployment.

- **Processes:** Methodologies for implementing oversight/governance.

- **Tools:** Technologies and mechanisms to support governance activities.
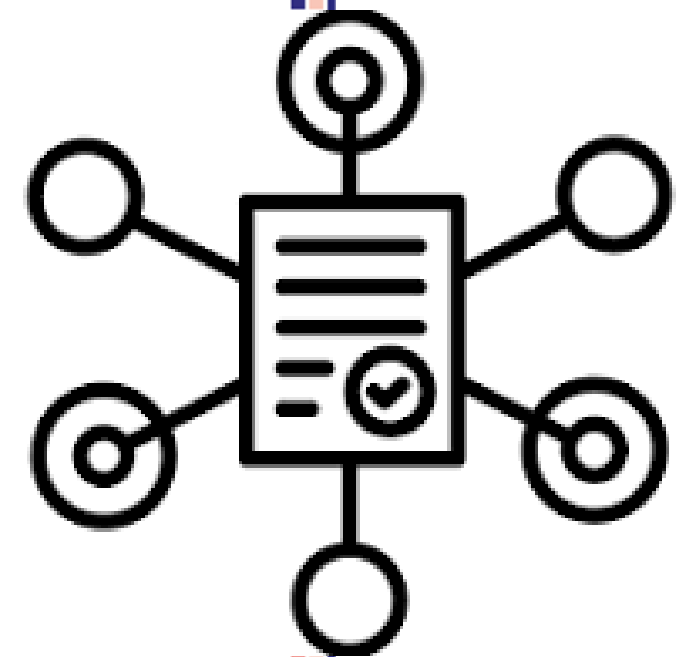
Governance Frameworks and Principles - an Evolving Landscape

- **EU AI Act:** A landmark **regulatory** framework classifying AI systems by risk level, with stricter requirements for 'high-risk' AI.

- **NIST AI Risk Management Framework (AI RMF):** A voluntary framework from the US National Institute of Standards and Technology, offering practical guidance for managing AI risks.

- **ISO/IEC 42001 (AI Management System):** An emerging international standard providing requirements for an AI management system.

- **Microsoft: Responsible AI Standard**

  - Microsoft developed its own **Responsible AI Standard -** aligns with global frameworks like **OECD AI Principles** and **ISO/IEC 42001**.

  - It includes mandatory impact assessments, fairness reviews and human oversight for high-risk systems.

  - ***Microsoft Transparency Report 2025***

- **Australia's 8 voluntary AI Ethics Principles** **- Not a framework** but a set of principles are part of the Australian Government's plan to make Australia a global leader in responsible and ethical AI.

- **OECD AI Principles:** International, non-binding principles for responsible AI stewardship.

Five Best Practices for AI Governance
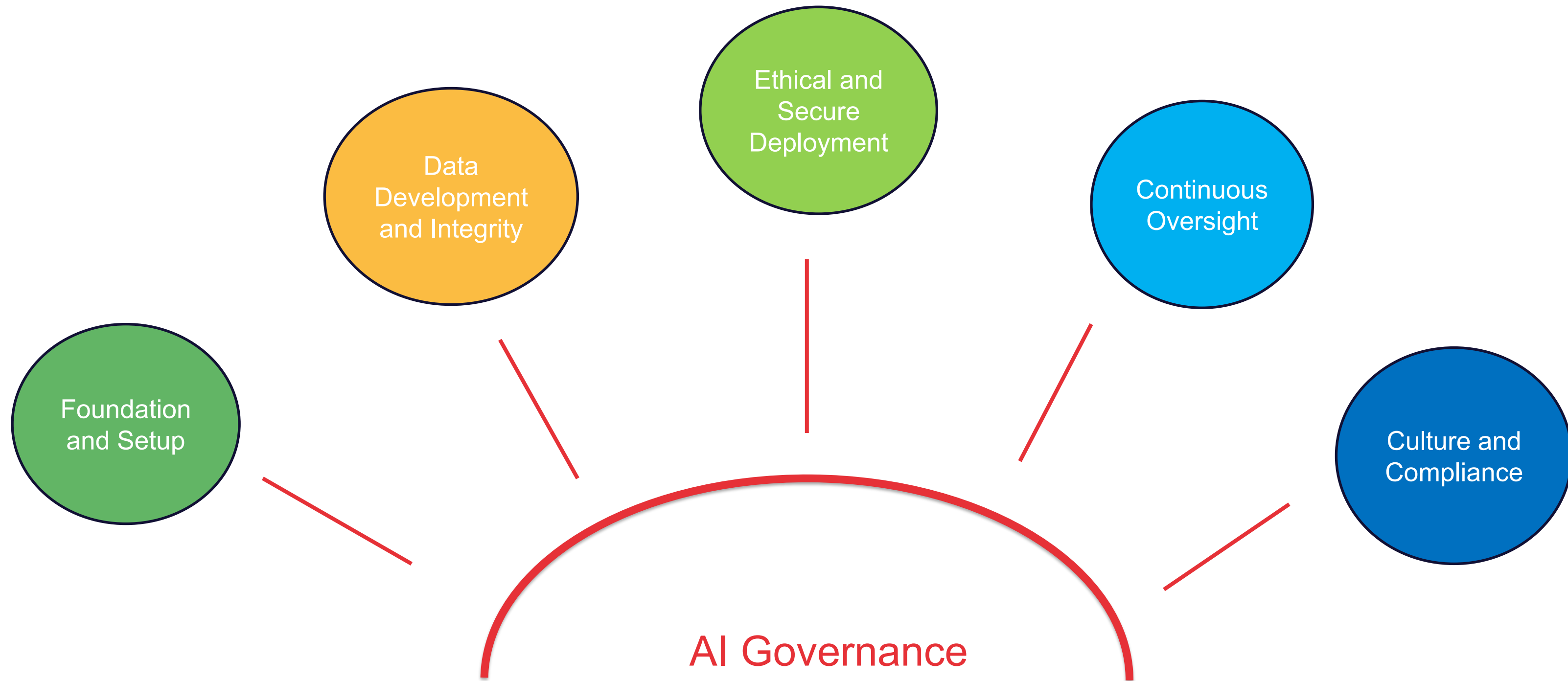
Best Practice – Foundation and Setup

- **Clear Roles and Responsibilities:**

  - Create a **cross-functional AI governance team or committee** with diverse expertise (legal, ethics, data science, IT, business).

  - Define a clear accountability matrix for each stage of the AI lifecycle and ensure **board-level oversight**.

  - Provide ongoing training and education to all relevant personnel on AI ethics and governance.

- **Policy and Ethical Guideline Development:**

  - Establish foundational ethical principles (e.g., fairness, transparency, accountability, human oversight, privacy, safety) that align with organisational values.

  - Develop policies that span the AI lifecycle - from data collection to deployment, emphasising human-centric design

- **Strategic Alignment and Scoping:**

  - Link AI initiatives to your overall business goals.

  - Categorise AI systems by risk level and define the scope of your governance efforts **–** iterative process

Foundation and Setup

Best Practice – Data and Model Development Integrity

- **Data Quality and Management:**

  - Integrate AI data governance with existing enterprise data governance framework. Focus on ensuring high-quality representative training data.

- **Rigorous Model Development and Validation:**

  - Implement comprehensive testing methodologies to assess model performance, reliability and safety.

- **Proactively monitor and mitigate bias** during development.

  - Maintain version control for models and code to ensure reproducibility.

- **Independently validate models** before deployment to confirm they meet both performance and ethical standards.

- **Transparency and Explainability (XAI):**

  - Design AI systems to be understandable and auditable, documenting decision logic and features influencing outcomes.

  - Transparency about the capabilities and limitations of AI systems and inform users when they are interacting with AI.

Data Development and Integrity

Best Practice – Ethical and Secure Development

- **Bias Mitigation and Fairness:**

  - Seek out and use diverse end users to minimise bias.

  - Conduct regular audits for discriminatory outcomes and embed fairness considerations particularly in the initial design phase.

  - Monitor deployed models for bias drift.

- **Robust Security and Privacy:**

  - Implement security and privacy by design principles from the outset of AI development

  - Utilise privacy-enhancing technologies where appropriate.

  - Develop an incident response plan for AI-related security breaches or model corruption

Ethical and Secure Deployment

Best Practice – Continuous Oversight

- **Continuous Monitoring and Auditing**

  - Establish systems for real-time performance monitoring of AI models in production.

  - Conduct periodic internal and external audits to ensure ongoing compliance and effectiveness- establish feedback loops.

- **Stakeholder Engagement and Training**

  - Foster multi-stakeholder collaboration - engage internal teams, customers and external reference groups.

  - Provide internal training on AI literacy and ethical considerations.

  - Maintain open communication and transparency about the organisation's AI approach, create feedback channels for stakeholders.

Continuous
Oversight

Best Practice – Culture and Compliance

- **Proactive Regulatory Compliance:**

  - Monitor evolving AI regulations and standards

  - Maintain documentation and audit trails to demonstrate adherence.

- **Cultivating a Responsible AI Culture:**

  - Foster a culture of innovation with responsible AI behaviours.

  - Requires strong leadership buy-in and an ethics-first mindset embedded throughout the organisation.

Culture and Compliance

A Pragmatic Approach – Start Simple

- Treat AI governance as a board-level priority
  - Board and executive engagement and leadership is critical

- Don't delegate by default to IT!

- Start with identifying high-risk systems and expand

- Integrate AI risks into enterprise risk management framework

- Establish some principles from frameworks like NIST and Microsoft into operational practices

- Invest in talent, training and tooling

Q & A

IBRS

Submit an inquiry or *schedule
a whiteboard session*

IBRS

IBRS | ANALYSIS INSIGHT JUDGEMENT

IBRS is a boutique Australian ICT Advisory Company.

We help our clients mitigate risk and validate their strategic decisions by providing independent and pragmatic advice while taking the time to understand their specific business issues.

https://ibrs.com.au   info@ibrs.com.au   02 4758 9111   PO Box 519, Hazelbrook NSW 2779, Australia