# Safe and Responsible AI - Putting it into practice

## How do we put this into practice?

We understand the five key principles that guide responsible AI use. The next step is understanding how to apply these principles within an organisation.

AI governance provides the framework for doing this. AI Governance has multiple building blocks:

* AI Policy & Principles (that we covered in Module 1)
* People & Skills
* Organisational Structure
* Monitoring, Reporting & Evaluation
* Procedures & Controls

Your agency may have its own AI governance in place or be in the process of establishing this.  It is important that you are familiar with your agency’s AI governance policies, processes and structures, or help highlight where these need to be put in place. When establishing AI governance, it should be done in a way that enables innovation (responsibly), not hinder it.

## Guiding Principles for AI - The Public Service AI Framework

In Module 1, we referred to AI as a powerful engine, and gave an overview of the rules of the road – key laws and regulations that govern AI. To keep the passengers of our AI vehicle safe, we need a combination of these (external) laws, and internal AI Governance.

## Consider how each AI governance building block keeps our AI vehicle safe

AI Policy & Principles:

An AI Policy steered by identified principles (OECD) and risk appetite is used to define the ethical objectives guiding the development, deployment and procurement of AI systems.

Clear driving instructions help the driver understand how to operate the vehicle.

People, Skills and Technology:

People with the right skills need to be allocated and/or trained to safely execute against AI ambition – powered with a responsible technical architecture on which to make this happen.

The driver needs to have the right skills and experience before getting behind the wheel.

Organisational Structure:

The operating model for AI needs to be clearly defined – including deciding who makes what decision, and who is involved where in the AI lifecycle

It is helpful to decide before turning on the engine who will drive, who will navigate, and who can make the call to stop the car.

Monitoring, Reporting & Evaluation:

A robust monitoring system is needed to evaluate the health, performance and compliance of AI systems, and review AI governance effectiveness ongoing.

It is important to keep checking the speedometer and keep an eye out for the check engine light.

Procedures & Controls:

An AI Risk Management Framework is used to map, measure and manage risks throughout the AI lifecycle – e.g. initial impact assessment, ongoing risk and controls assessment.

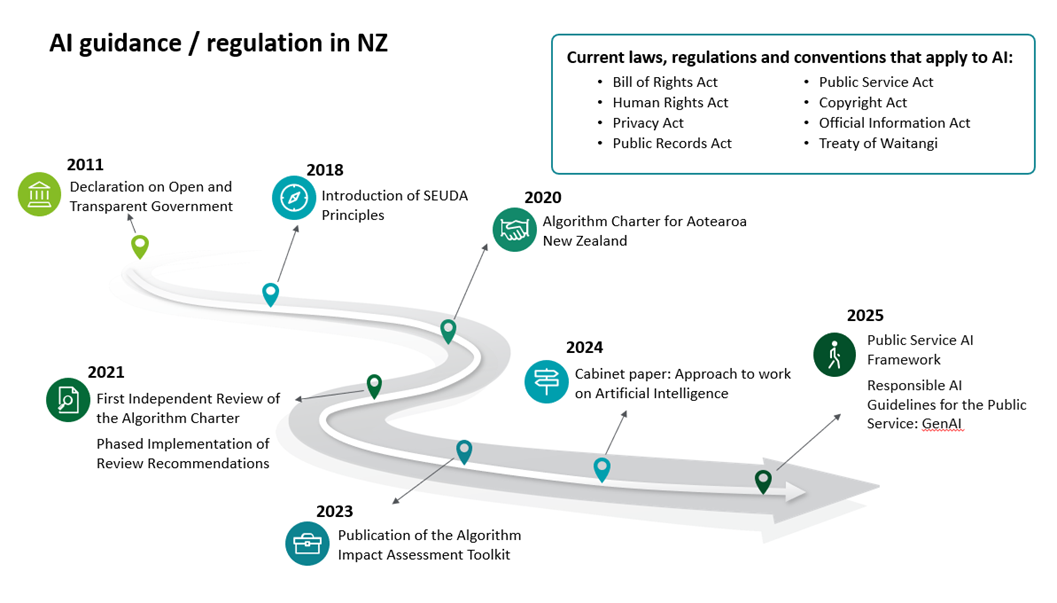
IT is important to evaluate risks at the right part of the journey – check tire pressure before departing, keep an eye on traffic throughout.

## What obligations and guidelines should we keep in mind?

We explored legal and regulatory obligations in Module 1 – covering the current laws, regulations and conventions that apply to AI in New Zealand. In addition to these, the graphic below covers the specific guidance that has emerged as relates to algorithms and AI.

As we continue to integrate AI into public service, it is essential to stay informed, recognising that guidelines and obligations will evolve alongside technological advancements.

## AI guidance/ regulation in NZ



Your agency may be a signatory to the Algorithm Charter, which covers algorithms broader than (but inclusive of) AI systems. You can read more about its six commitments of transparency, partnership, people, data, privacy and human oversight on the Statistics New Zealand website. <https://data.govt.nz/assets/data-ethics/algorithm/Algorithm-Charter-2020_Final-English-1.pdf>

Further direction on responsible AI can be found in Government advice in the Approach to Work on AI July 2024 Cabinet Paper, which mandates agencies to implement the AI principles of the OECD. We cover these principles, and detailed examples related to them, in the sections that follow.

A Public Service AI Framework was published in 2025, as covered in Module 1. The GCDO is leading a Public Service AI work programme to support the implementation of the Framework’s vision. <https://www.digital.govt.nz/standards-and-guidance/technology-and-architecture/artificial-intelligence/public-service-artificial-intelligence-framework>

Specific guidance on Responsible AI use for the Public Service was made available: Responsible AI Guidance for the Public Service: GenAI | NZ Digital government. We do not cover this guidance in detail in this course, as it is specifically focussed on Generative AI. <https://www.digital.govt.nz/standards-and-guidance/technology-and-architecture/artificial-intelligence/responsible-ai-guidance-for-the-public-service-genai>

## Conclusion

You have now explored the key principles for responsible AI use in the public service, with a focus on data and security. This was a high-level introduction to help you start thinking critically about AI risks and responsibilities. As you move towards completing your use case template, apply what you have learned to assess AI solutions in a responsible and informed way