AI DEPLOYMENT AND ETHICS QUICK KIT

JULY 2025

engineering
new zealand
te ao rangahau

Artificial intelligence (AI) systems are increasingly used in engineering to improve efficiency, support decision-making and automate complex tasks. As these systems transition from research and prototyping into real-world operation, it becomes critical to ensure that deployment is ethical, explainable and robust.

This quick kit provides practical tools for engineering professionals, regulators and project teams in Aotearoa New Zealand. It is designed to support responsible deployment of AI in civil infrastructure, utilities, transport, geotechnical systems, environmental modelling and other engineering domains.

The contents reflect both international best practice and the specific expectations of New Zealand's regulatory, cultural and environmental landscape.

These templates are intended for use at design handover, commissioning and whenever models are reviewed or re-tuned in production.

Contents

Ethical checklist

Use this checklist before deploying any AI-based system in an operational environment. It is also recommended as a prompt for internal reviews or pre-audit preparation.

- **Is your AI system explainable?**
  Can domain users or affected stakeholders understand how the system produces its outputs? Provide simple summaries, visualisations, or decision trees where applicable.

- **Have you checked for bias or unfair outcomes?**
  Review your training data and outputs across different groups (eg age, gender, ethnicity, geographic location). Have Māori and other underrepresented communities been considered?

- **Have you assessed alignment with Te Ao Māori perspectives?**
  If the AI interacts with natural resources (eg land, water), cultural indicators, or taonga species, have local iwi and hapū been consulted or informed?

- **Is personal data collected or used with consent?**
  Does the system comply with the Privacy Act 2020? Have you documented how data is stored, accessed and deleted?

- **Is there an identified decision-maker or escalation path?**
  Can someone override or investigate the AI's decision if needed? Is a human-in-the-loop model being used where safety is critical?

- **Can the system be paused, disabled, or reset safely?**
  Ensure there is a clear plan for dealing with failure modes or safety breaches without harming people, infrastructure or the environment.

- **Is the model's performance evaluated in real-world conditions?**
  Lab accuracy may not reflect live conditions. Has the system been trialled under representative scenarios?

- **Are outputs traceable and auditable?**
  Ensure logs, metadata and model versions are captured and that decisions can be reviewed retrospectively.

- **Has risk been assessed conservatively?**
  For high-stakes systems, incorporate buffers, alert thresholds and conservative assumptions to reduce downstream harm.

AI model summary sheet template

This is a structured summary of each AI model to support internal documentation, peer review and regulatory engagement.

Section	Description
Model name	eg SlopeWatch_v1.3
Purpose	eg Predicts slope instability from LIDAR and rainfall data
Developer	eg Wellington Geotechnics, in partnership with NIWA
Deployment date	eg 20/04/2025
Version	eg 1.3 (retrained March 2025)
Model type	eg Gradient boosting, ensemble classification
Inputs	eg 72hr rainfall, slope angle, soil type, NDVI from satellite
Outputs	eg Binary landslide risk (low/high) with confidence score
Training data	Summary of dataset(s), collection dates, and licensing status. Include known gaps (eg under-sampled alpine regions).
Limitations	eg Reduced accuracy in high-altitude catchments or after forest clearing. Not validated for landslip-prone clay soils.
Bias or fairness considerations	eg Model may underestimate events on Māori land blocks due to lack of training data from these zones.
Retraining or update policy	eg Monthly updates with additional incident data; formal review annually.
Review history	Date and outcome of ethics or regulatory reviews.
Contact	Technical lead or responsible officer: name, phone, email.

Risk mitigation plan template

Section	Description
System name	[Enter the name of the AI-enabled system]
Deployment context	[Brief description of where and how the system is used, eg embedded in SCADA system, accessed via dashboard]
Criticality rating	[Low / Moderate / High – based on potential impact of system failure or incorrect predictions]
Key operational risks	[List of known risks such as false positives, failure to alert, data drift, performance degradation]
Monitoring and alerting	[Specify how the system is monitored (eg logs, dashboards, thresholds), and how alerts are triggered]
Fallback or override mechanisms	[Describe manual or automated processes that apply if the system fails or behaves unexpectedly]
Performance KPIs	[Examples: prediction accuracy > 90%, downtime < 4 hrs/month, alert latency < 5 minutes]
Review cycle	[How frequently performance and risk are reviewed – eg monthly, post-incident, quarterly audit]
Responsible party	[Name and contact of the person or role responsible for monitoring and managing risks]

When to use these tools

These tools are designed to support responsible engineering practice across the AI system lifecycle. Use them at the following stages:

- **New deployments:** Before any AI system is commissioned for operational use, apply the ethical checklist and complete a model card to document its function, limitations and oversight structure. This helps ensure the model is ready for live environments.
- **Third-party systems:** Where AI systems are supplied by external vendors or consultants, the model card template can be used to clarify what the system does, what data it uses, and what monitoring or accountability mechanisms are in place.
- **Critical infrastructure and public services:** For AI used in water, energy, transport or emergency systems, apply all three tools (checklist, model card, risk plan) before deployment. These sectors typically have heightened safety and equity requirements.
- **Regulatory engagement:** Use these documents when preparing reports for councils, certification bodies, or environmental auditors. They demonstrate due diligence and can help resolve compliance questions.
- **Ongoing operations and reviews:** For AI that is retrained, updated, or periodically re-evaluated, revisit these tools at each major update. The risk mitigation template is particularly useful for planning and documenting live monitoring arrangements.
- **Incident response and audit preparation:** If a system fails, underperforms, or is subject to public or internal scrutiny, these temp lates provide a record of responsible practice and a starting point for investigation and improvement.

Disclaimer

The information provided in this document is for general informational purposes only. While we have made every attempt to ensure that the information contained herein is accurate and up to date as of July 2025, Engineering New Zealand makes no representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability, or availability with respect to the information contained in this document. As always, professional judgement and discretion must be applied and exercised.