Top 10 Artificial Intelligence Recommendations

Speak to your organisation's AI specialists to better understand the below recommendations and risks.

1. Develop an AI framework and strategy

- Set a clear vision, guiding principles and risk appetite to ensure it aligns with your organisation's objectives.
- Create a cross-functional working group to steer the program and resolve issues. Consider including teams from cyber security, risk, privacy, legal, procurement, innovation, data and business owners.

2. Create and enforce an organisation-wide AI policy

- Translate the framework into a single, mandatory policy that sets out roles, responsibilities and approval gates.
- Publish concise, easy-to-follow guidelines and decision trees so staff understand when and how they may use AI tools.
- Embed the policy in induction, supplier contracts and project governance to ensure consistent adoption.

3. Train the workforce

- Offer tailored training for developers, risk owners and end-users to improve adoption and maturity on safe prompting, data handling, and escalation routes.
- Create prompt libraries to help guide users in reaching outcomes that are safe and successful.
- Foster a culture where staff feel confident reporting AI concerns.

4. Build and maintain an AI register

- Catalogue every AI model, tool or embedded feature you use along with the information required to manage. It should be like an asset or third-party register where you can identify and classify based on risk, and should be updated regularly.

5. Introduce risk-based impact assessments

- Screen each AI capability before deployment for ethical, privacy, safety and cyber risks. Maintain a consistent approach to your test scenarios that focuses on your highest risk.
- Apply deeper scrutiny to higher-risk applications such as public-facing generative tools.

6. Strengthen data governance

- Confirm that training and input data are accurate, representative and legally sourced.
- Use Privacy Impact Assessments and de-identification techniques for personal information.

7. Embed human oversight

- Ensure humans can review, override or shut down automated outputs.
- Assign clear accountability for performance, bias monitoring and incident response. Expect that things can go wrong and have people and processes in place for fixing it.

8. Secure the supply chain

- Add contractual clauses on AI use that protects your data, including audit rights for transparent assurance.
- Continually monitor your suppliers, ensuring any AI applied to your data or systems sits within your risk appetite.
- Actively monitor and scan open-source components for known vulnerabilities.

9. Monitor in production

- Track metrics for drift, bias, false positives and cyber anomalies.
- Have clear processes to efficiently respond and remediate, otherwise the speed and volume of issues that arise may be unmanageable over time.
- Set automated alerts to trigger AI retraining or roll-back when thresholds are exceeded.

10. Refresh controls continuously

- Review security architecture and controls early in the design phase and throughout change lifecycles where risk has increased.
- Regularly review AI-related guidelines and user training material given the constant change of the AI landscape. Consider leveraging the NIST AI Risk Management Framework to support.
- Conduct tabletop exercises to test your response to prompt injection, model inversion or hallucination events.

VMIA is the Victorian Government's insurer and risk adviser

Level 10 South
161 Collins Street
Melbourne VIC 3000

P (03) 9270 6900
E contact@vmia.vic.gov.au
ABN 39 682 497 841

vmia.vic.gov.au