AI Risk & Incident Management Checklist

Identifying, scoring and treating AI risk, and responding to AI incidents against their reporting deadlines

ISO/IEC 42001 · ISO/IEC 23894 · ISO 31000 · EU AI Act · GDPR · POPIA

AI fails in its own ways, but the discipline for managing it is the one you already use for every other risk. This checklist covers AI risk from identification to board reporting, and AI incidents from first report to lessons learned.

A AI Risk Identification

- ☐ Risks are identified per AI system, not only for the AI programme as a whole
- ☐ A shared list of AI failure categories is used, such as bias, fabricated output, data leakage, performance drift and adversarial attack
- ☐ Risk sources include the data, the model, the people using it and the suppliers behind it
- ☐ Each AI risk has a named owner who can act on it

B Scoring & Appetite

- ☐ AI risks are scored with the same likelihood and impact scales as every other business risk
- ☐ Inherent and residual scores are both recorded, so the effect of controls is visible
- ☐ Risks above appetite have a treatment plan with an owner and a deadline
- ☐ A legally prohibited use is never accepted as a risk; it is stopped
- ☐ AI risks appear in the same board and committee reports as operational, financial and cyber risk

C Controls & Human Oversight

- ☐ Each AI risk is linked to the controls that treat it
- ☐ Controls are tested for design, implementation and operating effectiveness, with evidence and a validity date
- ☐ Human oversight, overrides and interventions are recorded as events, not described in a policy
- ☐ Events reported automatically by AI tools are reviewed and confirmed by a person before they count as evidence

D Monitoring

- ☐ Performance and output quality of each AI system are monitored against agreed thresholds
- ☐ Drift in accuracy, fairness or behaviour raises an alert to the risk owner
- ☐ Feedback and complaints from users and affected people are captured and linked to the system
- ☐ Supplier notices about model updates or known issues are tracked and assessed

E AI Incident Response

- [ ] Everyone knows how to report an AI incident, including near misses and harmful output
- [ ] Incidents are classified by type and severity and linked to the AI system and its risks
- [ ] Immediate containment is recorded: system paused, output withdrawn or human review increased
- [ ] Affected people and internal owners are informed according to a defined process

F Regulatory Reporting Clocks

- [ ] For each incident, you check whether a regulator deadline applies: EU AI Act serious incident, GDPR or POPIA breach notification, or sector rules
- [ ] The deadline is calculated from the right starting point and is visible to the incident owner
- [ ] Deployers of high-risk AI systems inform the provider first, as the EU AI Act requires
- [ ] Notifications made, and the reasons for not notifying, are recorded with dates

G Investigation & Lessons Learned

- [ ] Root cause is identified for each significant incident, looking at data, model, use and controls
- [ ] Corrective actions have owners and deadlines, and their effectiveness is checked before closure
- [ ] Risk scores, controls and impact assessments are updated with what the incident showed
- [ ] Incident trends across AI systems are reviewed at management level

H Change-Triggered Review

- [ ] Changes to a model, data source, capability or purpose are logged before they go live
- [ ] Each change is assessed for which reviews it should trigger: risk, impact, controls, oversight or classification
- [ ] Triggered reviews have due dates and owners, and a person decides each outcome
- [ ] Supplier-driven changes to bought-in AI tools are treated the same as internal changes

© XGRC® Software, a Strategix product. Strategix Application Solutions (Pty) Ltd, Reg No. 2015/192960/07. ISO 27001:2022 certified. info@xgrcsoftware.com