OCTOBER 2026 · APPROVED ISSUE

Responsible AI and AI Ethics Policy

Explainability, auditability and human oversight across the lifecycle

Organisation	UnifAI Technology Limited · Company no. 12301340
Version	2026.2
Owner	Chief Technology Officer; Board oversight
Approved by	Dan Byles, Chief Executive Officer
Effective date	29 October 2026
Next review	29 October 2027, or sooner after a material change
Contact	info@unifaitechnology.com

1. Purpose and scope

UnifAI Technology Limited develops and uses AI to support real-world monitoring, prediction, interpretation and decisions. This policy sets requirements for responsible development, procurement, deployment and use. It applies to our predictive systems, Digital Reader applications, third-party models and internal use of generative AI.

The policy is aligned with the values of the OECD AI Principles. Requirements are applied proportionately to the intended use, affected people and potential harm.

2. Ownership, intended use and risk

Each application must have a named accountable owner and a documented intended purpose, users, operational context and allocation of responsibilities between UnifAI, customers and partners. Define what the output means, the decisions it may support and the circumstances in which it must not be used.

Before a material commitment or release, assess foreseeable harm, misuse, bias, security, privacy, environmental implications and the consequences of error or unavailability. Clinical, safety-critical and other high-impact uses require the appropriate domain expertise, regulatory assessment and independent review where warranted. Escalate unresolved significant risks to the Board.

3. Data provenance, rights and quality

Establish the right to use data and its provenance, relevant consent or other lawful basis, licences and customer restrictions. Separate training, development, evaluation and operational data appropriately. Protect personal and confidential data and limit access and retention to the agreed purpose.

Assess completeness, representativeness, labelling, measurement quality and relevant sources of bias. Identify where a dataset or sensor context differs from the proposed operating setting. Document significant limitations and the reference evidence used. Do not treat a large dataset as sufficient evidence of suitability by itself.

4. Objective-led modelling and evaluation

Define success and acceptance criteria before selecting a configuration. Consider the practical costs of errors, missed events, false alerts, delay and resource use, rather than relying on a single accuracy score. Record material operating constraints and the reasons for trade-offs.

ANNA’s automated experimentation must operate within agreed objectives and bounds. Evaluation should distinguish development performance from appropriately separate validation evidence. Assess the model together with thresholds and response rules, including relevant subgroups, sites, devices and unusual or degraded conditions.

Claims must identify the application and evidence supporting them. A strong result for a historical test or site must not be presented as a guarantee for a different assay, population, site or future deployment.

5. Explainability, transparency and auditability

Determine the explanation needed by users, operators and reviewers. Explainable AI may involve inspectable signals, feature behaviour, model-level methods, operational rules or a combination appropriate to the model and use. Recording a configuration is useful for audit, but does not by itself explain every prediction.

Retain proportionate records of data provenance, model and software versions, training or optimisation configuration, validation results, approved decision rules, release decisions and significant changes. Protect these records from unauthorised alteration and manage access and retention.

Where Digital Reader workflows collect image evidence, line intensity, confidence or contextual metadata, the approved configuration must specify which fields are captured, how they support review and how privacy and quality are protected. Do not claim that optional fields are collected in every deployment.

6. Human oversight and operational guardrails

Define who may review, challenge, override, pause or stop the system, and what action follows an uncertain result or system failure. Users need information and training suitable for their role. Human oversight must be meaningful, not merely a nominal approval of an output that cannot be examined.

Guardrails should specify permitted inputs, operating ranges, decision thresholds, escalation conditions and actions the system is authorised to take. A change to a material rule or intended use requires review. Automated action requires explicit agreement about authority, safeguards and fallback behaviour.

7. Security, resilience and release

Use security controls appropriate to the application, including access management, separation of environments, dependency management, change review and protection of model and data assets. Assess relevant attacks, data poisoning, manipulation of inputs and misuse of interfaces.

Release only after the required acceptance, quality, regulatory and customer approvals. Record the configuration, responsibilities, deployment plan, fallback arrangements and rollback route. Clinical product obligations must be assessed for the particular intended use; a historical authorisation does not automatically cover a new application.

8. Monitoring and controlled improvement

Agree monitoring appropriate to the risks and available evidence, including data quality, performance, drift, operational outcomes and incidents. Define review frequency, thresholds for investigation and responsibilities for responding.

Retraining or optimisation must not silently bypass validation or authorisation. A candidate configuration must meet the agreed criteria before replacing an operational version, with a recoverable record of the change. The requirements for automated replacement depend on the approved deployment; the wider managed lifecycle remains a development direction where it is not yet implemented.

Review significant incidents and complaints, correct or suspend unsafe or unsuitable use and communicate material effects to the appropriate parties. Retirement must include access removal, data disposition and records needed for continuing obligations.

9. Third-party and generative AI

Use approved tools and assess supplier terms, data use, retention, international access, security and intellectual-property implications before providing protected information. Do not use customer or personal data to train a general-purpose model without express authorisation and a valid legal basis.

Check generated text, code and analysis before relying on it. Use additional verification for consequential claims, security-sensitive code and decisions affecting people. Do not invent sources, evidence, permissions or completed actions. Disclose material AI involvement where required for transparency or an agreed deliverable.

10. Assurance and review

The Chief Technology Officer is responsible for technical implementation, with the Chief Executive Officer responsible for commercial and organisational commitments and the Board overseeing significant risks. Project records must show the actual allocation of responsibilities and required approvals.

Review the policy annually, following significant incidents and when legal, technical or deployment requirements change. Compliance with this policy must be evidenced through project records and corrective actions, not inferred from a public statement of principles. Exceptions require a documented decision by an appropriately authorised owner and must never waive legal obligations.

Reference framework

OECD AI Principles, updated 2024

ICO: AI and data protection guidance

UnifAI Privacy Notice and Supplier Code - Related policies, version 2026.2