



ETHAAI

Guía ETHAAI para la Adopción Responsable de la Inteligencia Artificial

Marco Estratégico de Gobernanza, Gestión de Riesgo y Supervisión Institucional

Documento Técnico

Versión 1.0

Global Alliance for the Ethical Use of Artificial Intelligence (ETHAAI)

Tabla de Contenidos

1. Resumen Ejecutivo
 2. Introducción
 3. Contexto Global y Tendencias Regulatorias
 4. Riesgos Sistémicos de la Inteligencia Artificial
 5. Marco ETHAAI de Adopción Responsable
 6. Modelo de Madurez ETHAAI
 7. Recomendaciones de Política Organizacional
 8. Consideraciones Estratégicas para América Latina
 9. Implementación Gradual y Hoja de Ruta
 10. Conclusión
 11. Anexos
 12. Referencias
-

1. Resumen Ejecutivo

La inteligencia artificial (IA) ha dejado de ser una tecnología emergente para convertirse en infraestructura estratégica transversal. Su impacto alcanza productividad, servicios públicos, finanzas, salud, educación y defensa.

La aceleración tecnológica supera actualmente la maduración institucional.

Esta brecha genera tres tensiones estructurales:

- Innovación vs. regulación





ETHAAI

- Eficiencia vs. responsabilidad
- Automatización vs. supervisión humana

En ausencia de marcos internos de gobernanza, las organizaciones enfrentan:

- Riesgos reputacionales
- Exposición legal
- Dependencia tecnológica concentrada
- Pérdida de control estratégico

El Marco ETHAAI propone una arquitectura práctica de adopción responsable basada en cuatro pilares:

1. Gobernanza Institucional
2. Evaluación y Gestión de Riesgo
3. Transparencia y Trazabilidad
4. Supervisión y Mejora Continua

Su propósito no es frenar la innovación, sino hacerla sostenible.

2. Introducción

La IA representa un cambio estructural comparable a la electrificación o la digitalización.

A diferencia de ciclos tecnológicos anteriores, la IA:

- Toma decisiones probabilísticas
- Aprende dinámicamente
- Interactúa directamente con usuarios
- Genera contenido indistinguible del humano

Por tanto, no puede gestionarse únicamente como software tradicional.

Requiere gobernanza.

3. Contexto Global y Tendencias Regulatorias

3.1 Convergencia Internacional

Diversas instituciones han desarrollado principios y marcos normativos orientados a IA responsable:





ETHAAI

- OECD – Principios de IA confiable
- National Institute of Standards and Technology – AI Risk Management Framework
- European Commission – AI Act
- World Economic Forum – Modelos de gobernanza de IA

Existe convergencia en cinco elementos:

- Supervisión humana
- Transparencia
- Gestión basada en riesgo
- Seguridad
- Rendición de cuentas

3.2 Implicaciones para Organizaciones

La tendencia global indica que:

- La regulación aumentará
- La trazabilidad será obligatoria
- La responsabilidad institucional no será delegable

Las organizaciones que implementen marcos internos anticipadamente reducirán fricción futura.

4. Riesgos Sistémicos de la Inteligencia Artificial

4.1 Riesgos Operativos

- Model drift (desactualización del entrenamiento del LLM)
- Dependencia de proveedores
- Generación de errores plausibles (hallucinations)

4.2 Riesgos Legales

- Uso de datos no autorizados
- Responsabilidad en decisiones automatizadas
- Falta de documentación

4.3 Riesgos Reputacionales





ETHAAI

- Sesgos detectados públicamente
- Uso engañoso de IA
- Incidentes amplificados en redes sociales

4.4 Riesgos Estratégicos

- Concentración de mercado
- Dependencia geopolítica
- Asimetría tecnológica

Centros como Stanford HAI y AI Now Institute han documentado ampliamente estos fenómenos.

5. Marco ETHAAI de Adopción Responsable

5.1 Arquitectura General

El Marco se compone de cuatro pilares integrados.

PILAR I: Gobernanza Institucional

Objetivo

Formalizar responsabilidad y toma de decisiones.

Componentes

- Comité de Supervisión de IA
- Responsable de IA Ética
- Política Corporativa de IA
- Evaluación contractual de proveedores

Resultado Esperado

Decisiones trazables y supervisadas.

PILAR II: Evaluación y Gestión de Riesgo

Enfoque Basado en Impacto





ETHAAI

Nivel	Impacto	Requisito
Bajo	Interno	Registro
Medio	Usuarios	Evaluación documentada
Alto	Legal/Financiero	Auditoría interna
Crítico	Derechos fundamentales	Supervisión reforzada

Instrumentos

- Al Impact Assessment
 - Análisis de sesgos
 - Evaluación de dependencia tecnológica
-

PILAR III: Transparencia y Trazabilidad

Requisitos

- Documentación técnica
 - Registro de versiones
 - Log de decisiones automatizadas
 - Comunicación responsable
-

PILAR IV: Supervisión y Mejora Continua

Componentes

- Monitoreo post-despliegue
 - Red teaming
 - Protocolo de incidentes
 - Revisión anual
-

Principio Transversal

Supervisión humana significativa en todos los niveles.

La responsabilidad no se delega al algoritmo.





ETHAAI

6. Modelo de Madurez ETHAAI

Nivel 1 – Experimental

Uso informal, sin documentación.

Nivel 2 – Operativo

Casos de uso aislados, evaluación limitada.

Nivel 3 – Gobernado

Política formal, supervisión documentada.

Nivel 4 – Estratégico

IA integrada en gobernanza corporativa.

7. Recomendaciones de Política Organizacional

1. Nombrar responsable formal de IA.
 2. Crear inventario de sistemas IA.
 3. Implementar evaluación de impacto previa.
 4. Integrar IA en auditoría interna.
 5. Exigir transparencia contractual.
-

8. Consideraciones Estratégicas para América Latina

8.1 Oportunidad

- Adopción temprana con gobernanza integrada.
- Ventaja competitiva frente a mercados desorganizados.

8.2 Riesgos

- Dependencia tecnológica externa.
- Brecha regulatoria.
- Falta de capacidades técnicas locales.

La región puede adoptar IA responsable desde el inicio sin cargar con deuda institucional heredada.





ETHAAI

9. Implementación Gradual

Fase 1 (0–90 días)

Diagnóstico institucional.

Fase 2 (6 meses)

Formalización de gobernanza.

Fase 3 (12 meses)

Integración en procesos corporativos.

Fase 4 (18–24 meses)

Auditoría independiente.

10. Conclusión

La IA no es únicamente una herramienta tecnológica.

Es un factor estructural de poder económico y organizacional.

Las organizaciones que gobiernan la IA gobernarán su futuro.

El Marco ETHAAI ofrece un equilibrio:

Innovación con responsabilidad.

Velocidad con control.

Automatización con supervisión.

11. Anexos

Anexo A – Plantilla de Evaluación de Impacto en IA

- Descripción del sistema
- Nivel de riesgo
- Usuarios afectados
- Supervisión humana prevista
- Mitigaciones

Anexo B – Checklist de Debida Diligencia de Proveedores

- ¿El proveedor documenta limitaciones?





ETHAAI

- ¿Existen cláusulas de responsabilidad?
- ¿Se permite auditoría externa?

Anexo C – Inventario Institucional de IA

- Nombre del sistema
- Área responsable
- Nivel de riesgo
- Fecha de revisión

12. Referencias

OECD (2019). *OECD Principles on Artificial Intelligence*.

NIST (2023). *AI Risk Management Framework*.

European Commission (2024). *Artificial Intelligence Act*.

World Economic Forum (2022). *AI Governance Framework*.

Stanford HAI (2023). *AI Index Report*.

AI Now Institute (2023). *AI Governance Publications*.

