

Field	Value
Document ID	GOV-EAI-001
Version	1.0
Effective Date	December 2024
Document Owner	Chief AI Officer
Classification	Internal Use
Review Cycle	Annual

Table of Contents

[Table of Contents will be auto-generated]

Ethics and Responsible AI Principles

1. Purpose

This document establishes comprehensive ethical principles and responsible AI governance framework to ensure all artificial intelligence systems deployed within the organization align with human-centric values, maintain fairness and safety standards, and operate under clear accountability mechanisms. The framework provides mandatory review triggers, assessment criteria, and oversight structures to embed ethical considerations throughout the AI lifecycle from conception to decommissioning.

The principles outlined herein serve as the foundational governance layer for all AI-related activities, ensuring compliance with emerging regulatory requirements while fostering innovation within ethical boundaries. This framework directly supports organizational risk management, regulatory compliance, and stakeholder trust through transparent, auditable AI governance processes.

2. Scope

2.1 Organizational Coverage

This policy applies to all organizational entities, subsidiaries, joint ventures, and third-party partnerships involving AI system development, deployment, or operation. Coverage includes:

- Internal AI development teams and data science organizations
- Business units procuring or implementing AI solutions
- IT operations teams managing AI infrastructure
- Vendor relationships involving AI services or platforms
- Research and development activities incorporating AI technologies
- Customer-facing AI applications and services

2.2 AI System Classifications

AI System Type	Risk Level	Review Requirements	Oversight Frequency
High-Risk AI	Critical	Full ethics review	Quarterly

Systems		panel assessment	monitoring
Moderate-Risk AI Systems	Medium	Departmental ethics review	Semi-annual monitoring
Low-Risk AI Systems	Low	Self-assessment checklist	Annual monitoring
Experimental/ Research AI	Variable	Research ethics committee review	Project-based

2.3 Core Coverage Areas

- Algorithmic bias detection, measurement, and mitigation
- AI system explainability and interpretability requirements
- Digital accessibility and inclusive design standards
- Human oversight and control mechanisms
- Data privacy and protection in AI contexts
- Third-party AI vendor assessment and monitoring
- AI incident response and remediation procedures

3. Foundational Principles

3.1 Human-Centric Design

All AI systems must prioritize human welfare, dignity, and agency. This principle requires:

- Design processes that center human needs and capabilities
- Preservation of meaningful human decision-making authority
- Protection of human autonomy and self-determination
- Enhancement rather than replacement of human capabilities where feasible
- Consideration of diverse human perspectives and experiences in system design

3.2 Fairness and Non-Discrimination

AI systems must operate equitably across all demographic groups and use cases:

- Proactive identification and mitigation of algorithmic bias
- Regular assessment of disparate impact across protected categories
- Implementation of fairness metrics appropriate to system context
- Transparent documentation of fairness testing methodologies
- Corrective action protocols for identified bias or discrimination

3.3 Transparency and Explainability

Organizations must maintain appropriate levels of AI system transparency:

Transparency Level	System Category	Requirements
High Transparency	Customer-facing decisions	Full explanation of decision factors
Medium Transparency	Internal process optimization	Summary explanation of key factors