ACME CORPORATION

Responsible AI Development Policy

Version 2.1 | Effective Date: January 1, 2025

SECTION 1: PURPOSE AND SCOPE

This policy establishes mandatory requirements for all AI systems developed, deployed, or procured by ACME Corporation. All employees, contractors, and third-party vendors involved in AI development must comply with these requirements.

SECTION 2: DATA GOVERNANCE REQUIREMENTS

2.1 Data Collection and Use

All AI systems must adhere to the following data governance principles:

- Personal data shall only be collected with explicit user consent
- Data minimization: collect only data necessary for the specified purpose
- Purpose limitation: data shall not be repurposed without additional consent
- Retention limits: personal data must be deleted after 12 months unless legally required otherwise
- Data quality: implement validation checks to ensure accuracy and completeness

2.2 Prohibited Data Sources

The following data sources are strictly prohibited for AI model training:

- Customer data without explicit opt-in consent
- Employee surveillance data (except where legally mandated)
- Scraped data from websites without proper licensing
- Any data that includes protected health information (PHI) unless HIPAA-compliant
- Biometric data collected without specific biometric consent

SECTION 3: MODEL DEVELOPMENT REQUIREMENTS

3.1 Fairness and Bias Testing

All high-impact AI models must undergo bias testing before deployment:

- Test for disparate impact across protected characteristics (race, gender, age, disability)
- Document bias testing methodology and results
- Implement mitigation strategies where bias is detected
- Quarterly re-testing of production models

3.2 Explainability Requirements

AI systems that make decisions affecting individuals must provide:

- Human-readable explanations for individual decisions
- Documentation of model architecture and key features
- Ability to trace decisions to specific data inputs
- Regular model interpretability audits

SECTION 4: TRANSPARENCY AND DOCUMENTATION

4.1 AI System Registry

All AI systems must be registered in the central AI inventory within 30 days of development start:

- System name and purpose
- Data sources used
- Model type and architecture
- Risk classification (low, medium, high)
- Responsible team and point of contact

4.2 Audit Logs

High-risk AI systems must maintain comprehensive audit logs:

- All model predictions and decisions
- Input data used for each decision
- Timestamp and user context
- Minimum retention period: 5 years

SECTION 5: HUMAN OVERSIGHT REQUIREMENTS

5.1 Human-in-the-Loop

Critical AI decisions require human review:

- Employment decisions (hiring, termination, promotion)
- Credit decisions above $10,000
- Healthcare treatment recommendations
- Legal risk assessments

5.2 Override Capability

All AI systems must include:

- Mechanism for human operators to override AI decisions
- Documentation requirement when override is exercised
- Escalation path for contested AI decisions

SECTION 6: THIRD-PARTY AI SYSTEMS

6.1 Vendor Due Diligence

Before procuring external AI systems, teams must:

- Conduct vendor AI responsibility assessment
- Review vendor's data practices and privacy policy
- Verify compliance with relevant regulations
- Obtain contractual guarantees regarding data use and model behavior

6.2 API and Model Risks

When using third-party AI APIs or foundation models:

- Conduct monthly monitoring for model drift or degradation
- Implement fallback systems for API failures
- Review vendor security certifications annually
- Prohibit sending sensitive data to external APIs without encryption and approval

SECTION 7: INCIDENT RESPONSE

7.1 AI Incident Definition

An AI incident includes:

- Discriminatory outcomes detected in production
- Privacy breach involving AI system data
- Significant model performance degradation
- Unintended harmful outputs or behaviors

7.2 Reporting Requirements

All AI incidents must be reported within 24 hours to:

- AI Ethics Committee
- Legal Department
- Data Protection Officer
- Affected business unit leadership

SECTION 8: COMPLIANCE AND ENFORCEMENT

Violation of this policy may result in:

- Mandatory retraining
- Project suspension or termination
- Disciplinary action up to and including termination
- Potential legal liability

This policy will be reviewed and updated annually or as regulatory requirements evolve.