Responsible AI and Model Governance Policy

ACME Corporation • Sample Document • Last updated 2025-11-09

PURPOSE

Ensure that AI capabilities in ACME Horizon, including the Apex assistant and predictive models, are designed and operated responsibly.

USE CASES AND RISK RATINGS

- Low impact - assistive suggestions with clear user control
- Medium impact - workflow automation with audit and rollback
- High impact - decisions affecting entitlements require human review

CONTROLS

- Data minimization and access boundaries
- Bias testing with representative synthetic datasets
- Model cards that document training data sources, limitations, and evaluation methods
- Continuous monitoring for drift and anomalous prompts

SECURITY CONSIDERATIONS

- Prompt and output filtering to reduce injection risk
- Isolation of model contexts per tenant
- Signed model artifacts and supply chain controls