## Overview

This document provides an **Enterprise Grade AI Risk Register** aligned to NIST AI RMF, ISO/IEC 42001 and MITRE ATLAS adversarial threat standards. It includes risk catalog, scoring model, RAG classification, and recommended mitigation controls.

## 1. AI Risk Scoring Model

**Risk Score = Impact × Likelihood × Detectability**

- **Impact: Severity of harm**
- **Likelihood: Probability of occurrence**
- **Detectability: Ease of detecting failure**

**Scores map to: Red (40–100), Amber (20–39), Green (1–19)**

## 2. RISK CATEGORIES

- **Fairness & Bias**
- **Explainability**
- **Privacy & Data Protection**
- **Security & MITRE ATLAS Threats**
- **Performance & Reliability**
- **Ethical & Societal**
- **Compliance & Legal**
- **Operational**
- **Reputational**

## 3. MITRE ATLAS Mappings

**Examples:**

- **AT1003 – Membership Inference**
- **AT1018 – Model Inversion**
- **AT1021 – Prompt Injection**
- **AT1005 – Data Poisoning**
- **AT1028 – Adversarial Input Manipulation**
- **AT1025 – Bias Exploitation**
- **AT1029 – Misuse of AI Capabilities**

## 4. Risk Register Structure

- **Risk ID**
- **Risk Category**
- **Description**
- **Impact, Likelihood, Detectability**
- **Risk Score + RAG Rating**
- **MITRE ATLAS mapping**
- **Existing & Recommended Controls**
- **Risk Owner + Status**

## 5. Heatmap Dashboard:

**Impact × Likelihood matrix with color-coded RAG scores to visualize high risk zones.**