Appendix for Framing AI as an Enterprise Risk

Below is a structured overview of the main standards relevant to AI governance, grouped into useful categories for your course.

We can see these as a layered AI governance stack:

1) Regulation (EU Artificial Intelligence Act)
2) Governance frameworks (NIST AI RMF, ISO 42001)
3) Risk taxonomies (MIT AI Risk Repository)
4) Threat models (MITRE ATLAS, OWASP LLM Top 10)

These are based on a simple AI risk taxonomy:

- Strategic risks
- Operational risks
- Cyber/security risks
- Regulatory risks
- Societal/ethical risks

1. CORE AI GOVERNANCE STANDARDS

a. NIST AI Risk Management Framework (AI RMF)

Organisation: National Institute of Standards and Technology

Purpose: provides a framework for identifying, assessing and managing AI risks

Core functions: govern, map, measure, manage

Widely used internationally, aligns well with ERM approaches

b. ISO/IEC 42001 (AI Management System Standard)

Organisation: International Organization for Standardization

Purpose: first formal AI management system standard.

Focus: AI governance structures, risk management, lifecycle oversight, documentation and accountability

Analogous to ISO 27001 (information security) and ISO 9001 (quality management)

c. OECD AI principles

Organisation: Organisation for Economic Co-operation and Development (OECD)

Core principles: transparency, accountability, robustness, human oversightand fairness

These principles influence many national AI governance frameworks.

d. Risk taxonomies:

Define what risks exist

Examples: MIT AI Risk Repository, OECD AI risk taxonomy

2. ENTERPRISE RISK MANAGEMENT STANDARDS

AI governance is usually embedded within enterprise risk management.

a. **COSO Enterprise Risk Management Framework**
   Organisation: Committee of Sponsoring Organizations of the Treadway Commission
   Focus: risk governance, risk appetite, strategic risk integration
   Used heavily in: financial services listed companies

b. **ISO 31000 Risk Management**
   Organisation: International Organization for Standardization
   Purpose: provides general risk management principles and processes
   Core principles: integrated, structured, inclusive, dynamic
   Useful for integrating AI into existing risk governance processes

3. CYBERSECURITY STANDARDS (AS APPLICABLE TO AI)

AI systems introduce new cyber risks, so cyber standards are part of AI governance.

a) **NIST Cybersecurity Framework**
   Organisation: National Institute of Standards and Technology
   Core functions: identify, protect, detect, respond, recover
   AI relevance: protecting training data, securing models and preventing AI attacks

b) **ISO/IEC 27001**
   Organisation: International Organization for Standardization
   Purpose: Information Security Management System
   Important for: data used by AI, access control, monitoring and incident response

c) **AI Security Standards/ frameworks**
   These focus specifically on AI attack surfaces based on the MITRE ATLAS standards covering threat models for adversarial machine learning attacks. Examples: include data poisoning, model extraction, adversarial inputs, prompt injection

4. **REGULATORY AI GOVERNANCE FRAMEWORKS**

These are not standards but regulatory frameworks that influence governance.

a) **EU Artificial Intelligence Act**: risk-based regulation
   Risk categories: unacceptable risk, high risk, limited risk, minimal risk

b) **UK AI Governance Framework**: developed under the UK AI White Paper for sector-specific regulation, pro-innovation approach and regulatory oversight

c) **SR 11-7 Model Risk Management**
   Regulator: Federal Reserve System covering model validation, governance, and monitoring. Used by banks, insurance firms, regulators

d) **OWASP**: Top 10 risks for LLMs