CONCEPT

AI Risk Assessment, Model Validation and AI Assurance

AI Risk Assessment, Model Validation and AI Assurance

Technical depth: Practitioner · **Practical mode:** Case study

Introduction

AI systems are moving into credit decisions, customer service, recruitment, fraud detection and operations faster than organisations can assess them. Models are procured or built without an inventory entry, tested only for accuracy and left unmonitored as data drifts, while generative AI adds confabulation, prompt injection and data leakage. This Core Concept course equips risk and assurance practitioners to assess AI risk and test models against recognised standards, from impact assessment to validation and ongoing monitoring. Participants work through case material from several sectors and leave with an AI Risk and Assurance Pack for a system in their own organisation.

Course Objectives

- Build and maintain an AI system inventory with risk tiering aligned to ISO/IEC 42001 and the NIST AI RMF
- Conduct AI risk and impact assessments using ISO/IEC 23894 and ISO/IEC 42005
- Evaluate model performance, bias, robustness and explainability evidence against defined acceptance criteria
- Assess generative AI and third-party model risks, including prompt injection, data leakage and confabulation
- Design monitoring, drift detection and revalidation triggers for models in production
- Produce an AI Risk and Assurance Pack with an evidence-based assurance opinion

Target Audience

- Risk managers responsible for adding AI to the enterprise risk profile
- Model validation and model risk managers testing models before and after deployment
- Internal audit managers planning assurance engagements over AI systems
- Data science and machine learning leads accountable for model documentation and controls
- Compliance and privacy managers assessing the impact of AI on individuals
- Technology and procurement managers evaluating third-party AI products and services

Course Outline

Day 1: AI Risk Landscape and Current-State Inventory

- AI System Types and Life Cycle Stages in ISO/IEC 5338
- AI Incident Case Review: Bias, Hallucination and Automation Failures
- OECD AI Principles and Trustworthy AI Characteristics
- AI System Inventory Build and Use-Case Risk Tiering
- AI Governance Current-State Gap Assessment

Day 2: AI Risk and Management System Standards

- ISO/IEC 42001 AI Management System Requirements and Annex A Controls
- ISO/IEC 23894 AI Risk Management Guidance Mapped to ISO 31000
- NIST AI RMF Core Functions: Govern, Map, Measure and Manage
- NIST AI 600-1 Generative AI Profile Risk Categories
- ISO/IEC 25059 Quality Model for AI Systems

Day 3: Assessing Risk and Validating Models

- AI System Impact Assessment Using ISO/IEC 42005
- Model Card and Datasheet Review for Documentation Completeness
- Performance Validation: Holdout Testing, Confusion Matrix and Threshold Selection
- Bias Measurement Using ISO/IEC TR 24027 Fairness Metrics
- Explainability Evidence Review with SHAP and LIME Outputs

Day 4: Advanced Risks, Controls and Monitoring

- Generative AI Threat Review with the OWASP Top 10 for LLM Applications
- Adversarial Attack Mapping with MITRE ATLAS
- Third-Party and Foundation Model Due Diligence Questionnaire
- Data Drift and Concept Drift Monitoring with the Population Stability Index
- Human Oversight Design and Model Revalidation Triggers

Day 5: Case Work and the AI Risk and Assurance Pack

- Credit Scoring Model Case Study: Validation Findings and Challenge
- Customer Service Chatbot Case Study: Generative AI Red-Team Results Review
- AI Risk Register and Control Mapping for an Own System
- AI Risk and Assurance Pack Drafting
- Assurance Opinion Defence Before a Peer Model Risk Committee

Skills You Will Gain

- AI Use-Case Risk Tiering
- Model Validation
- Algorithmic Bias Assessment
- Explainability Review
- Generative AI Threat Analysis
- AI Control Mapping
- AI Vendor Due Diligence
- AI Assurance Reporting

Why Attend This Course

- Return to work with an AI Risk and Assurance Pack for a real system, already defended before a peer committee
- Ask data science teams and vendors the questions that expose weak validation or missing controls
- Give senior management and the board a clear view of which AI systems carry the most risk and why
- Compare AI assurance practice with risk and technology professionals from other sectors and countries

Conclusion

AI can be trusted only when its risks are assessed before deployment and its behaviour is tested and watched afterwards. This course moves from the AI risk landscape and the governing standards, through impact assessment and model validation, to generative AI threats, third-party models and drift. The final day turns that material into an AI Risk and Assurance Pack that participants take back to their organisation, giving them a repeatable, standards-based method for every AI system they are asked to assess.