CONCEPT

AI Data Readiness: Data Quality, Governance and Preparation for AI Use Cases

AI Data Readiness: Data Quality, Governance and Preparation for AI Use Cases

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

Introduction

Most AI initiatives stall not because the model is weak but because the data behind it is incomplete, inconsistent, poorly documented or legally unusable for the purpose intended. Organisations commission pilots, then discover that the records needed are scattered across systems, owned by nobody and unfit for training or retrieval. This Core Concept course equips data and business managers to assess and prepare organisational data for specific AI use cases using the ISO/IEC 5259 series and established data management practice. Participants leave with an AI Data Readiness Assessment and Remediation Plan for a priority use case.

Course Objectives

- Assess the fitness of organisational data for a defined AI use case against the ISO/IEC 5259-2 data quality characteristics
- Map the data sources, lineage and ownership behind an AI use case and identify the gaps that block it
- Establish data governance roles, policies and quality controls for AI data using DAMA-DMBOK2 and ISO/IEC 38505-1
- Specify the labelling, documentation and de-identification steps a dataset needs before it is used for training or retrieval
- Identify bias, privacy and drift risks in AI datasets and define the controls that monitor them in production
- Produce an AI Data Readiness Assessment and Remediation Plan that a sponsor can fund and track

Target Audience

- Data and analytics managers responsible for supplying data to AI and machine learning initiatives
- Data governance, data quality and master data leads accountable for data standards and stewardship
- Business function managers who own the processes and records an AI use case depends on
- Digital transformation and AI programme managers preparing use cases for delivery
- IT and enterprise architecture managers responsible for data platforms and integration
- Risk, privacy and information management leads who approve the use of data for new purposes

Course Outline

Day 1: AI Data Readiness Foundations and Current State

- AI Data Lifecycle Concepts Under ISO/IEC 22989
- Use-Case Data Requirements Canvas: Inputs, Labels, Volumes and Refresh Rates
- Structured, Unstructured and Document Data for Generative AI Retrieval
- Data Maturity Baseline Using the EDM Council DCAM
- AI Data Readiness Scorecard Across Quality, Access, Governance and Volume

Day 2: Data Quality and Data Governance Frameworks

- ISO/IEC 5259 Series Structure and Terminology for Analytics and ML Data
- ISO/IEC 5259-2 Data Quality Measures and ISO/IEC 25012 Characteristics
- ISO/IEC 5259-3 Data Quality Management Requirements for AI Data
- DAMA-DMBOK2 Knowledge Areas: Metadata, Master Data and Data Quality
- ISO/IEC 38505-1 Governance of Data and Accountability for Data Use

Day 3: Preparing Data for AI Use Cases

- Data Source Inventory and Lineage Mapping for a Use Case
- Data Profiling and Quality Rule Definition with Open-Source Profiling Tools
- Labelling Guidelines, Annotation Workflows and Inter-Annotator Agreement
- Dataset Documentation with Datasheets for Datasets and FAIR Principles
- Document Chunking, Metadata Tagging and Access Controls for Retrieval-Augmented Generation

Day 4: Data Risk, Bias, Privacy and Ongoing Control

- Representativeness and Sampling Bias Checks on Training Data
- De-identification Techniques Under ISO/IEC 20889
- Purpose Limitation and Consent Review for Repurposed Data
- Data Drift Monitoring and Data Quality Service Levels in Production
- AI Data Risks Mapped to NIST AI RMF and ISO/IEC 42001 Controls

Day 5: Case Work and the Readiness Plan

- Customer Service Chatbot Case Study: Knowledge Base Readiness Review
- Predictive Maintenance Case Study: Sensor and Work-Order Data Gap Analysis
- Readiness Scorecard Completion for an Own Use Case
- AI Data Readiness Assessment and Remediation Plan Drafting
- Sponsor Review Panel and Plan Defence

Skills You Will Gain

- Data Fitness Assessment
- Data Lineage Mapping
- Data Quality Rule Design
- Data Stewardship
- Dataset Documentation
- Data Bias Detection
- Privacy-Preserving Data Preparation
- Data Remediation Planning

Why Attend This Course

- Return to work with an AI Data Readiness Assessment and Remediation Plan for a use case your organisation is already considering
- Tell early whether an AI proposal is blocked by data before budget is committed to building it
- Speak a common data quality language with data scientists, vendors and business owners, based on international standards
- Compare how organisations in banking, healthcare, utilities and the public sector prepare their data for AI

Conclusion

AI delivers value only when the data behind it is fit for the purpose, documented, governed and lawful to use. This course moves from diagnosing a use case's data requirements, through the ISO/IEC 5259 quality model and data governance frameworks, to the practical preparation, bias, privacy and monitoring work that makes data usable in production. The final day brings that together in a readiness plan that shows sponsors what must be fixed, by whom and in what order before an AI initiative can proceed.