AI for Records Management and Digital Archives: Auto-Classification, HTR and Redaction

2 – 6 August 2027
Paris - Right Bank business hotel



AI for Records Management and Digital Archives: Auto-Classification, HTR and Redaction

Ref: 931_21487 **Date:** 2 – 6 August 2027 **Location:** Paris - Right Bank business hotel **Fees:** 23500 SAR

Introduction:

AI for records management and digital archives is moving from vendor demonstrations into daily practice, yet many records and archives teams still key metadata by hand, leave digitised handwritten collections unsearchable and file content against the file plan inconsistently. Others switch on auto-classification or redaction tools without accuracy targets, human review or evidence that AI-processed records remain authentic. This Core Concept course gives records and archives practitioners a tool-neutral method to select, test and govern AI across capture, classification, description, access and disposition. Participants produce an AI Adoption Roadmap and Auto-Classification Pilot Design for a case records programme.

Course Objectives:

- Map AI use cases against each stage of the records lifecycle and rank them by value, feasibility and evidential risk
- Specify OCR, handwritten text recognition and intelligent capture workflows with measurable character and field accuracy targets
- Design auto-classification rules and training sets that assign content to file plan classes and retention categories
- Apply entity recognition to enrich descriptive metadata and to detect personal data for redaction before access or release
- Set human review, sampling and audit trail controls that preserve authenticity and chain of custody for AI-processed records
- Produce an AI Adoption Roadmap and Auto-Classification Pilot Design for a records or archives programme

Target Audience:

- Records managers accountable for file plans, retention schedules and disposition decisions
- Archivists responsible for digitised collections, description and researcher access
- Document control leads who oversee capture, indexing and distribution of controlled documents
- Information governance managers who approve new technology touching personal data and evidential records
- Digital archive and repository managers planning automation of ingest, description and search

Course Outline:

Day 1: AI Opportunities Across the Records Lifecycle

- AI Techniques Glossary: Machine Learning, Natural Language Processing, Computer Vision and Large Language Models
- Records Lifecycle Opportunity Map from Capture to Disposition
- ISO 15489 Records Controls as the Baseline for AI-Assisted Processes
- Backlog Diagnostic: Undescribed Collections, Unfiled Shared Drives and Legacy Scans
- AI Use Case Scoring Matrix: Value, Data Availability and Evidential Risk

Day 2: Recognition, Capture and Description Models

- OCR Pipeline Stages: Image Preprocessing, Layout Analysis and Text Output
- Handwritten Text Recognition Model Training on Archival Hands and Scripts
- Character Error Rate and Word Error Rate as HTR Acceptance Measures
- Intelligent Document Capture: Form, Letter and Invoice Field Extraction
- ISO 23081 Metadata Elements Populated by Extraction and Enrichment Models

Day 3: Auto-Classification, Entity Recognition and Archive Search

- Auto-Classification Against File Plan Classes: Rules, Supervised Models and Hybrids
- Labelled Training Set Design from Sample Records and Retention Categories
- Named-Entity Recognition for Persons, Organisations, Places and Dates in Case Files
- Personal Data Detection and Redaction Workflow Before Release or Public Access
- Semantic Search and Conversational Retrieval over Digitised Archive Holdings

Day 4: Disposition Automation, Assurance and Ethics

- Retention Trigger Detection and Disposition Review Queues Driven by Classifier Confidence
- Human-in-the-Loop Review, Confidence Thresholds and Statistical Sampling Plans
- Authenticity, Integrity and Chain of Custody Evidence for AI-Processed Records in OAIS Ingest
- Bias, Misclassification and Access Rights Risks in Archival AI Outputs
- ISO/IEC 42001 AI Management System Controls Applied to Records Operations

Day 5: Case Records Programme: Roadmap and Pilot Design

- Case Pack Review: File Plan Extract, Retention Schedule and Mixed Scanned Holdings
- Classification Pilot Scope, Sample Selection and Success Criteria Definition
- Pilot Evaluation Protocol: Precision, Recall and Reviewer Agreement Measures
- AI Adoption Roadmap Sequencing: Quick Wins, Scale-Up Stages and Decision Gates
- AI Adoption Roadmap and Auto-Classification Pilot Design Presentation and Peer Challenge

Skills You Will Gain:

- AI Use Case Prioritisation for Records
- HTR and OCR Accuracy Evaluation
- Auto-Classification Design
- Metadata Extraction and Enrichment
- Personal Data Redaction
- Classifier Confidence Thresholding
- Chain of Custody Assurance
- AI Pilot Evaluation

Why Attend This Course:

- Leave with an AI Adoption Roadmap and Auto-Classification Pilot Design ready to adapt to a live records or archives programme
- Question AI vendor claims with accuracy measures, sampling plans and pilot success criteria of your own
- Open digitised handwritten and printed collections to search and description without losing evidential value
- Compare AI adoption experience with records, archives and document control practitioners from public, corporate, legal and heritage settings

Conclusion:

AI helps records and archives only when recognition, classification and redaction outputs are measured, reviewed and recorded as part of the evidential trail. The course moves from lifecycle opportunities and use case scoring, through OCR, handwritten text recognition, capture and metadata models, to auto-classification, entity recognition, redaction and archive search, then disposition automation, assurance and ethics. The final day turns this into an AI Adoption Roadmap and Auto-Classification Pilot Design for a case records programme.