Preservica

AI Guidelines - Cheat Sheet

1. Purpose of AI Use

□ The specific tasks or processes where AI is applied (e.g., metadata extraction, document classification, OCR etc).

2. Data Types and Sources

□ What types of archival data can and cannot be processed by AI (e.g., personal data, sensitive records, public archives).

□ The provenance and quality of data used for training or inference.

3. Legal and Ethical Considerations

□ Compliance with relevant laws (e.g., GDPR, UK-GDPR, EU AI Act).

□ Ethical principles such as fairness, transparency, and accountability.

4. Human Oversight

□ When and how a human must review or approve AI-generated outputs (“Human in the Loop” requirements).

5. Risk Assessment

□ Identified risks associated with AI use (e.g., bias, errors, data leakage).

□ Mitigation strategies and acceptable confidence thresholds for AI outputs.

6. Vendor and Third-Party AI

□ Documentation of vendor compliance with laws and standards.

□ Where data is stored and processed, and whether it is used to train external AI models.

7. Transparency and Explainability

□ How AI decisions are documented and explained to users and stakeholders.

□ Records of model versions, training data, and decision logic.

8. Access Controls and Security

□ Who can access AI systems and data.

□ Security measures to protect archival data during AI processing.

9. Training and Awareness

□ Records of staff training on AI risks, ethics, and compliance.

10. Monitoring and Review

□ Procedures for ongoing monitoring of AI systems.

□ Logs of incidents, errors, or policy breaches, and actions taken.