



Croatian State Archives

EU AI ACT

AI IN DIGITAL ARCHIVES

Exploring the Power of AI in Digital Archives

dr. sc. Lana Žaja

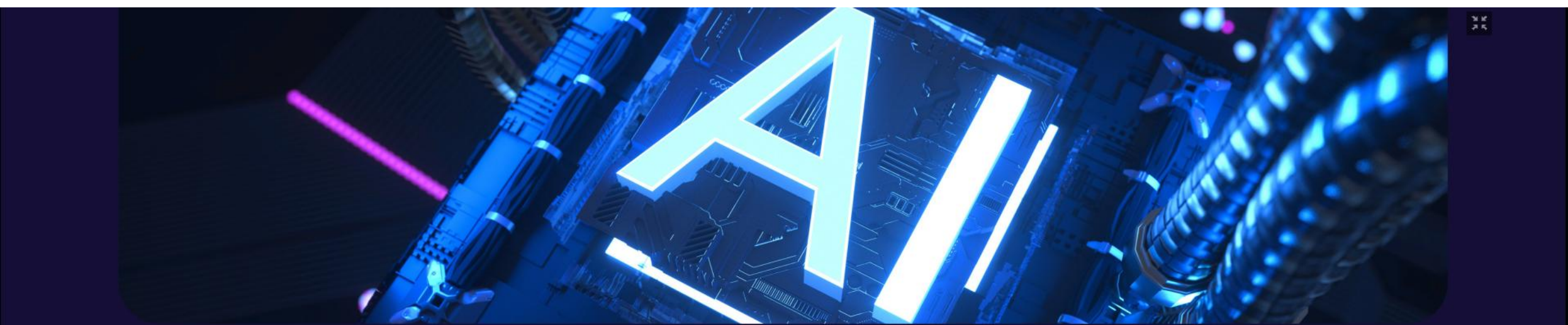


What is AI in Digital Archives?

Artificial Intelligence in digital archives refers to the use of intelligent algorithms to automate classification, enhance searchability, detect patterns, and support long-term data preservation and access.

- **Data Integrity**
- **High-Risk AI Systems**
- **Cybersecurity**





Key Points for Digital Archives

Long-Term Preservation

Digital archives must ensure the integrity and accessibility of data over time, regardless of changes in technology or file formats.

Metadata and Indexing

Accurate and standardized metadata is essential for efficient retrieval, context, and validation of archived content.



**Authenticity and
Provenance**



**Access and Security
Balance**



Digital Archives & the EU AI Act

Key Connections

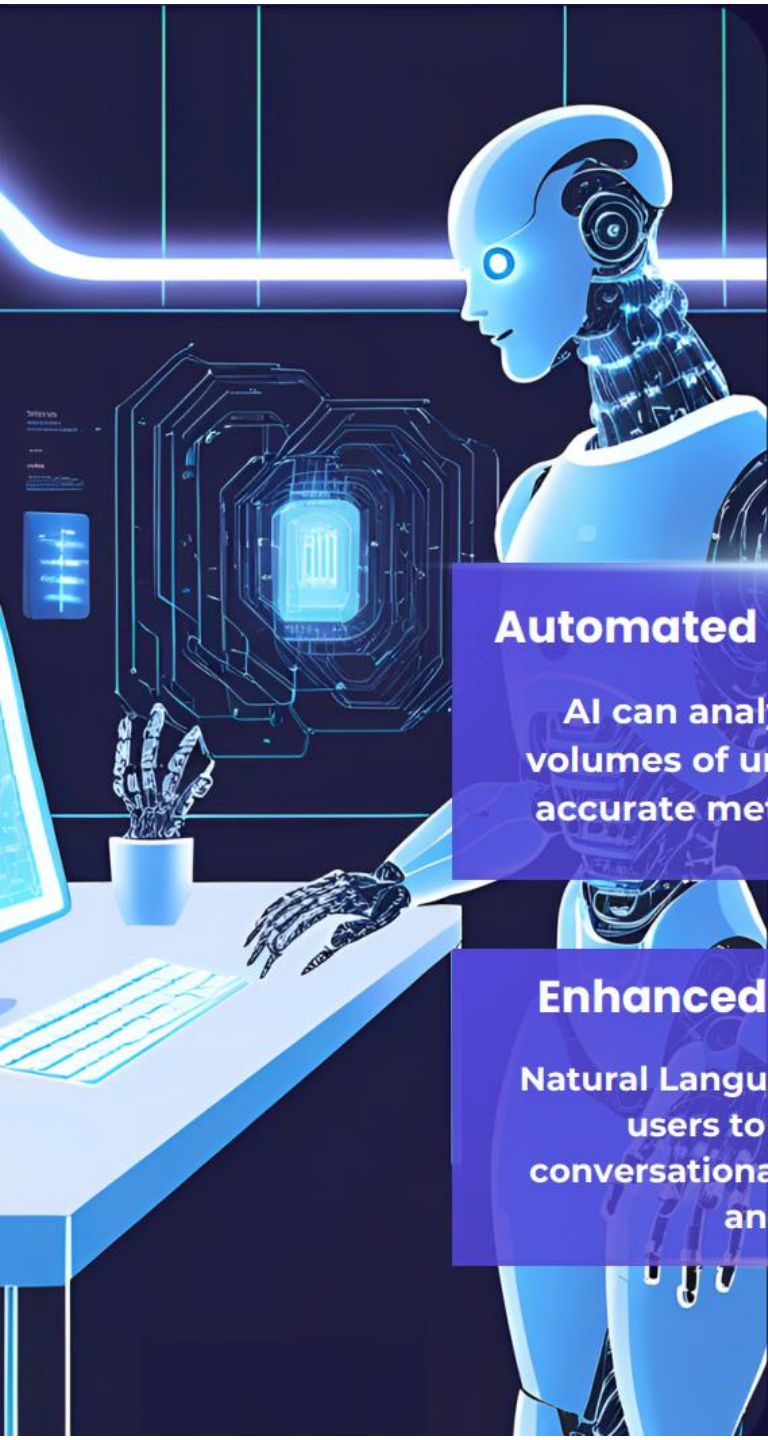
Data Governance & Traceability

Digital archives must ensure traceable data provenance to comply with the AI Act's transparency and accountability requirements

High-Risk AI Systems

Archives involved in training, storing, or providing datasets for AI models must assess if their content supports high-risk AI applications (e.g. biometric identification, legal document processing)





How AI is Transforming Digital Archives

Automated Classification & Tagging

AI can analyze and categorize large volumes of unstructured data, applying accurate metadata faster than manual methods

Digital Preservation & Anomaly Detection

AI monitors file integrity over time and detects signs of corruption or loss, supporting sustainable digital preservation.

Enhanced Search & Retrieval

Natural Language Processing (NLP) allows users to search archives using conversational queries, improving access and discoverability

Pattern Recognition & Historical Insights

Machine learning identifies hidden connections and trends across documents, enabling new interpretations of archival content.



How to Implement AI in Digital Archives



1. Assess Data Structure and Metadata Quality

Begin by evaluating existing digital collections to ensure data is well-structured and metadata is consistent, as AI tools rely heavily on clean, organized input.

2. Integrate AI Tools for Classification and Search

Use machine learning and natural language processing (NLP) to automate document tagging, extract entities, and enable semantic search across large datasets.

3. Establish Ethical and Compliance Frameworks

Implement guidelines that ensure transparency, data privacy, and alignment with standards like the EU AI Act, especially if archives are used in high-risk AI systems.

Digital Archives in Business Processes



Digital archives in business processes are structured digital repositories that securely store, manage, and preserve organizational records to support compliance, efficiency, and long-term knowledge retention

Centralized Information Management

Digital archives provide a structured and secure environment for storing and managing essential business documents across departments

Regulatory Compliance and Audit Readiness

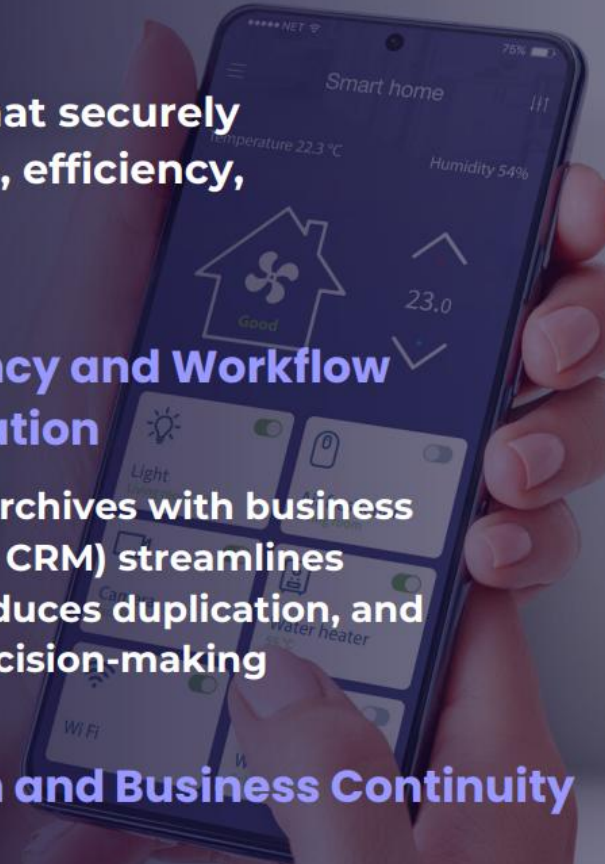
Archival systems ensure legal and regulatory compliance by maintaining accurate, time-stamped records with controlled access and traceability

Operational Efficiency and Workflow Integration

Integration of digital archives with business systems (e.g., ERP, CRM) streamlines document retrieval, reduces duplication, and accelerates decision-making

Knowledge Retention and Business Continuity

Archives preserve institutional memory and critical knowledge, ensuring continuity during staff turnover or organizational change





AI Legislation in Digital Archives

AI legislation in digital archives refers to the legal frameworks and regulatory requirements—such as the EU AI Act—that govern how artificial intelligence is applied to the storage, processing, and retrieval of digital archival content

Transparency and Accountability

Digital archives must support traceability of AI-driven decisions by preserving original datasets, metadata, and system documentation

Data Governance and Risk Management

Institutions must ensure high-quality, unbiased, and ethically sourced data in archives, particularly when used in high-risk AI systems

Compliance with the EU AI Act

Archives must align with the AI Act's obligations for human oversight, documentation, and cybersecurity, especially when AI is used for public records or sensitive content

Conclusion

As artificial intelligence becomes increasingly integrated into digital archiving, it is essential to ensure that AI tools enhance—not compromise—the integrity, accessibility, and ethical use of archival content



Ethical Data Architecture

Archives must be built on a foundation of transparent, inclusive, and bias-aware data structures to ensure fairness in AI processing and decision-making

Human Oversight by Design

AI implementations should include mechanisms for human validation and interpretability, ensuring that automated processes remain accountable and reversible



Thank You for Your Attention



00 385 99 339 3393



lanazaja@gmail.com



www.arhiv.hr

