ARCHEION
2026, T. CXXVII, s. 0–0
ISSN 0066-6041
e-ISSN 2658-1264

DOI 10.4467/26581264ARC.26.006.23513

STUDIA I MATERIAŁY

AI Lifecycle Records and Paradata: Reconceptualizing Documentation Requirements for Transparency and Accountability

Patricia C. Franks

San José State University (USA)
patricia.franks@sjsu.edu, ORCID 0000-0002-9917-6924

ABSTRACT
Artificial intelligence (AI) governance frameworks are emerging in response to growing demands for transparency, accountability, and regulatory oversight. However, many lack integrated recordkeeping requirements needed to document the AI system lifecycle. This study examines how AI documentation, documentation artifacts, and paradata – information describing the procedures, tools, methods, and decisions associated with AI processes – support transparent, accountable, and trustworthy AI governance. Using a multi-phase qualitative research design, the study combined a survey, literature analysis, regulatory review, and case studies to identify documentation requirements across the AI lifecycle. AI laws and governance frameworks from multiple jurisdictions were analyzed alongside case studies from the Saint Louis Zoo, the Bank of Canada, healthcare governance initiatives, and the NATO Archives. The findings show that existing AI governance frameworks remain fragmented and insufficient to support lifecycle accountability. The study identifies the documentation and paradata required throughout data preparation, model development, deployment, monitoring, and retirement, and highlights the critical role of Records and Information Management professionals in establishing governance policies, metadata standards, retention requirements, and contextual documentation to support transparent, auditable, and legally defensible AI systems.

KEYWORDS
Artificial Intelligence, AI Records, Documentation, Explainable AI, Metadata, Paradata

Dokumentacja w całym cyklu życia AI i paradane: Rekonceptualizacja wymogów dokumentacyjnych na rzecz przejrzystości i rozliczalności

STRESZCZENIE
Ramowy model zarządzania sztuczną inteligencją (AI) powstaje w odpowiedzi na rosnące potrzeby przejrzystości, odpowiedzialności oraz nadzoru regulacyjnego. Obecnie brakuje wymagań dotyczących prowadzenia dokumentacji niezbędnej do dokumentowania cyklu życia systemów AI. Niniejsze badanie analizuje, w jaki sposób dokumentacja AI, jej poszczególne elementy oraz tzw. paradane – informacje opisujące procedury, narzędzia, metody i decyzje związane z procesami AI – wspierają

SŁOWA KLUCZOWE
sztuczna inteligencja, rejestry AI, dokumentacja, wyjaśnialna sztuczna inteligencja, metadane, paradane

przejrzyste, odpowiedzialne i godne zaufania zarządzanie tą technologią. W ramach wieloetapowego badania jakościowego połączono ankietę, analizę literatury, przegląd regulacji oraz studia przypadków, aby zidentyfikować dokumentację powstającą na poszczególnych etapach cyklu życia AI. Przeanalizowano przepisy i ramy zarządzania AI z różnych instytucji oraz studia przypadków obejmujące: ogród zoologiczny w St. Louis, Bank Kanady, inicjatywy w zakresie zarządzania w ochronie zdrowia oraz archiwum NATO. Wyniki wskazują, że obecne ramy zarządzania AI są nadal niespójne i niewystarczające, by zapewnić odpowiedzialność w całym cyklu życia systemu. W badaniu określono wymogi dotyczące dokumentacji i paradanych na etapach przygotowania, tworzenia, wdrażania, monitorowania oraz wycofywania systemu z eksploatacji, a także podkreślono kluczową rolę specjalistów ds. zarządzania dokumentacją w kształtowaniu polityk zarządzania, standardów metadanych, wymogów dotyczących okresów przechowywania oraz dokumentacji kontekstowej, co jest niezbędne dla zapewnienia przejrzystości, możliwości audytu oraz zgodności systemów AI z wymogami prawnymi.

Introduction

Artificial intelligence (AI) refers to computational systems designed to perform tasks that typically require human intelligence, such as learning, reasoning, perception, language processing, and decision support. The term AI in this article refers to systems that employ multiple technologies (machine learning, generative AI, expert systems, natural language processing, computer vision, and more) to generate predictions, recommendations, decisions, or content influencing organizational activities. Because these systems affect business processes and create records throughout their lifecycle, they require governance mechanisms that ensure transparency, accountability, and the systematic creation and preservation of documentary evidence.

Artificial intelligence transforms the work of archivists and records and information management (RIM) professionals in two important ways. First, AI tools are increasingly integrated into records and information management processes, including the creation, classification, retrieval, retention, disposition, and preservation of information resources. Second, AI technologies – and the laws, regulations, and governance frameworks that accompany them – generate new categories of records, including datasheets, prompts, model cards, audit logs, and risk assessments, that must be created, managed, and preserved throughout the AI lifecycle. This second challenge – the documentation of AI systems throughout their lifecycle – is the focus of this paper.

Although AI governance frameworks increasingly emphasize ethics, risk management, transparency, and explainability, they rarely specify the documentary evidence needed to demonstrate how AI systems were designed, trained, evaluated, deployed, monitored, and retired. This paper argues that AI lifecycle records – the records and documentary evidence generated throughout the design, development, deployment, operation, and retirement of AI systems – constitute a distinct category of governance records comprising records, metadata, documentation artifacts, explainable AI outputs, and paradata that collectively document AI processes throughout the system lifecycle. Central to this framework is paradata, defined in relation to the AI process as "information about the procedure(s) and tools used to create and process information resources, along with information about the persons carrying out those procedures"[1]. Paradata complements data, metadata, documentation artifacts, and explainable AI by documenting the processes, decisions, activities, and contextual information necessary for trustworthy AI (i.e., artificial intelligence systems that are explainable, fair, interpretable, robust, transparent, safe, and secure).

Accordingly, this paper addresses the following central research question: How should AI lifecycle documentation be reconceptualized to support transparency, accountability, and governance throughout the AI system lifecycle?

To address this question, the paper examines three related issues: 1) what records, documentation artifacts, and paradata are required across the AI lifecycle; 2) how paradata complements data, metadata, and explainable AI within an AI lifecycle documentation model; and 3) how Records and Information Management (RIM) professionals can contribute their expertise during each phase of the AI Lifecycle.

Throughout the course of the study, the terminology used to describe information artifacts identified as evidence of AI actions and decisions was flagged and key terms were noted, some of which are defined and sourced in Table 1.

[1] InterPARES Trust AI, Terminology Database (2021–2026): Paradata, https://interparestrustai.org/terminology [access: 28.07.2026].

Table 1. Definition of Key Terms

Term	Definition and Source
Agentic AI	Agentic AI is an artificial intelligence system that can accomplish a specific goal with limited supervision. It consists of AI agents – machine learning models that mimic human decision-making to solve problems in real time. In a multiagent system, each agent performs a specific subtask required to reach the goal and their efforts are coordinated through AI orchestration. IBM, https://www.ibm.com/think/topics/agentic-ai
Artificial Intelligence	A sub-field of computer science research and practice that aims to develop computational models to simulate or approximate human cognition, behavior, decision-making, and reasoning. ITrustAI, https://interparestrustai.org/terminology/term/artificial%20intelligence
AI Features	Any tools, functionalities, or components in the Services that utilize artificial intelligence or machine learning models, including generative models, large language models, and algorithmic decision-making technologies. Mural AI Terms, https://www.mural.co/terms/ai-features
AI Model	A component of an information system that implements AI technology and uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs. NIST SP 800-218A, https://doi.org/10.6028/NIST.SP.800-218A.
AI System	An application that utilizes AI methods, potentially multiple methods, to perform a set of tasks. ITrustAI, https://interparestrustai.org/terminology/term/artificial%20intelligence%20system
AI System Lifecycle	The AI system lifecycle is a structured process that occurs in stages, ensuring the holistic coverage of the AI system from discovery to retirement. Digital.gov.au, https://www.digital.gov.au/policy/ai/AI-technical-standard/technical-standard-governments-use-artificial-intelligence-ai-system-lifecycle
AI Techniques	A set of methods that helps computers solve problems, make decisions, and perform tasks. They work on algorithms that are powered by human intelligence. EICTA, https://www.eicta.iitk.ac.in/knowledge-hub/artificial-intelligence/artificial-intelligence-technique
AI Technologies	An umbrella term that encompasses a broad spectrum of different technologies and applications including machine learning, natural language processing, computer vision, robotics process automation. FINRA, https://www.finra.org/rules-guidance/key-topics/fintech/report/artificial-intelligence-in-the-securities-industry/overview-of-ai-tech
Blackbox AI	Blackbox AI refers to AI systems with internal workings that humans cannot easily interpret. These systems, often built on complex algorithms like deep learning, operate without transparent decision-making processes. Founderz, https://founderz.com/blog/the-intricacies-of-blackbox-ai-explained/
Explainable AI (XAI)	An AI system that provides clear and understandable explanations for their actions and decisions. ITrustAI, https://interparestrustai.org/terminology/term/explainable%20AI%20~paL~XAI~paR~

Term	Definition and Source
Generative AI	Machine learning methods that aim to produce outputs that mimic the style and complexity of the training artifacts on which they were trained. ITrustAI, https://interparestrustai.org/terminology/term/generative%20AI
Paradata	Information about the procedure(s) and tools used to create and process information resources, along with information about the persons carrying out those procedures. ITrustAI, https://interparestrustai.org/terminology/term/paradata
Predictive AI	Predictive AI blends statistical analysis with machine learning algorithms to find data patterns and forecast future outcomes. It extracts insights from historical data to make accurate predictions about the most likely upcoming event, result or trend. IBM, https://www.ibm.com/think/topics/generative-ai-vs-predictive-ai-whats-the-difference.
Record	A document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference. Syn.: archival document. ITrustAI, https://interparestrustai.org/terminology/term/record

Source: Author's own elaboration.

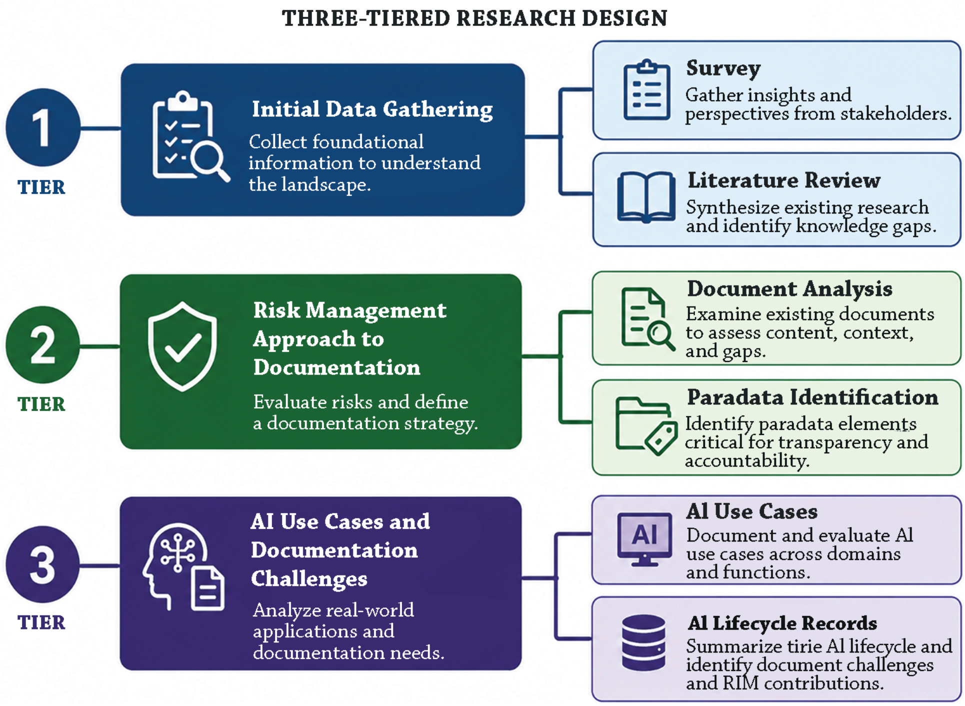
Research Methods

A qualitative, iterative approach was employed to enable concurrent data collection and analysis over the course of the InterPARES Trust AI study, *Preserving AI Techniques as Paradata,* which ran from September 2021 through March 2026. The comprehensive research design shown in Figure 1 integrated multiple sources of evidence, including a survey, literature reviews, guidance identification and analysis, and AI use cases, to investigate documentation requirements across the AI lifecycle. The findings from this long-term study underpin this paper.

The entire research team met monthly, with subgroups collaborating between meetings to develop a definition of *paradata* for the AI process and to design and implement the research plan. Access the final study report, individual journal articles, and conference presentations using the InterPARES Trust AI (2021–2026) dissemination tab.[2].

[2] InterPARES Trust AI, Dissemination tab. (2021–2026), https://interparestrustai.org/trust/research_dissemination [access: 28.07.2026].

Figure 1. Overview of Comprehensive Research Design for main study, *Preserving AI Techniques as Paradata*



Source: Author's own elaboration.

Theoretical Foundations

Understanding the evolution of AI tools and techniques relevant to archival and records management practice was central to this study. The research was further informed by five theoretical foundations: systems theory, archival theory, ethical theory, risk management theory, and agency theory.

- A *systems approach* was employed to consider the technical, operational, and human issues involved in ensuring accountability for AI implementation and in creating an AI Lifecycle record.
- Foundational constructs from *archival theory* were incorporated, including the terms "record" (e.g. document, information, data) and "trustworthiness" (e.g., accuracy, reliability, and authenticity).
- *Ethical theory* was incorporated when reviewing criteria for algorithms and their implementations. Professional and applied ethical theory was

employed to understand roles and responsibilities for multiple stakeholders and participants in AI development and implementation contexts.

– *Risk management theory* was employed throughout the study to provide a strategic framework for identifying, assessing, prioritizing, and addressing potential threats to organizational objectives based on the use of AI.
– *Agency theory* was examined in relation to issues of responsibility and accountability (ethical and legal). Traditionally defined as the capability of an entity (i.e., a human to act), the term is now also applied to AI programs (i.e., artificial agents, agentic AI).

Although this paper does not explicitly discuss the theoretical foundations of AI governance, the underlying concepts are reflected throughout the analysis. For example, the Bank of Canada AI use case incorporates principles of ethical theory through its emphasis on bias mitigation efforts within the framework for identifying and capturing metadata and paradata. Likewise, risk management theory informs the discussion of two foundational AI governance frameworks: the European Union Artificial Intelligence Act (EU AI Act) and the U.S. National Institute of Standards and Technology Artificial Intelligence Risk Management Framework (NIST AI RMF).

Three-Phase Research Study Underpinning Paper

The *Preserving AI Techniques as Paradata* research study was carried out in three phases: Initial Data Gathering, Risk Management Approach to Documentation, and AI Use Cases and Documentation Challenges.

Initial Data Gathering

Initial data gathering, conducted between 2021 and 2022, consisted of a survey and a comprehensive literature review to assess the state of AI process documentation. The literature review examined key concepts, including artificial intelligence, archives, records, documentation, metadata, explainable AI (XAI), and paradata. Findings from this phase were published in *Positioning Paradata: A Conceptual Framework for AI Processual Documentation in Archives and Recordkeeping Contexts*, which defined AI processual documentation as paradata, clarified its relationship to metadata and XAI, and argued that paradata provides a user-centered approach to explainability while strengthening AI accountability, transparency, and impartiality[3].

Survey of Information Professionals

A 2022 survey of 106 information professionals found that archival and records management practitioners were only beginning to adopt AI technologies. Among the sixteen respondents using AI in daily operations, only two reported documenting AI processes. One respondent used machine learning to identify born-digital moving image records for preservation and retained supporting documentation, including manuals, process diagrams, and technical specifications. The other employed machine learning for personally identifiable information (PII) analysis and automated metadata extraction, preserving comprehensive procedural documentation, digitization and file management practices, system administration manuals, and IT security documentation.

Review of Literature Related to AI, Records, and Paradata

An early literature review of AI processual documentation found that although the term *paradata* appeared in disciplines such as statistics, the humanities, archaeology, and cultural heritage, it was largely absent from archival literature. To address this gap, the research team examined paradata from three perspectives: 1) its use in the social sciences, cultural heritage, and documentation fields to describe information about the creation or curation of data and datasets; 2) its relationship to explainable AI (XAI) and approaches for addressing the "black box" problem; and 3) its role, alongside XAI, in archival literature on accountability and transparency[4]. This review identified a clear need for further research on AI documentation practices and the potential of paradata to address a significant gap in AI governance.

---

[3] S. Cameron, P. Franks, B. Hamidzadeh, *Positioning Paradata: A Conceptual Frame for AI Processual Documentation in Archives and Recordkeeping Contexts*, "ACM Journal on Computing and Cultural Heritage" 2023, vol. 16(4), article 75, pp. 1–19, https://doi-org.libaccess.sjlibrary.org/10.1145/3594728.

[4] Ibidem.

Risk Management Approach to Documentation

From 2023 through 2024, specific recommendations to document the AI process were investigated through a risk management lens. A sample of AI laws, regulations, standards, and governance frameworks from the European Union, United Kingdom, United States, Canada, and Singapore was analyzed to identify common documentation requirements. The initial goal was to develop a consolidated, lifecycle-based framework grounded in traceability, accountability, and continuous oversight. However, the analysis revealed that although many frameworks share common requirements – such as data provenance, model documentation, and impact assessments – no single framework can satisfy the needs of every organization or AI system. Instead, organizations must evaluate the laws, regulations, standards, and governance frameworks applicable to their AI use cases and develop documentation strategies tailored to their specific legal, operational, and governance requirements.

A risk-based documentation strategy aligned with emerging AI governance frameworks, such as the NIST AI Risk Management Framework and the EU AI Act, is recommended to support traceability, accountability, and continuous oversight. Such a strategy should address governance, technical documentation and paradata, and ongoing monitoring of AI systems. To ensure defensible AI documentation, organizations should develop a crosswalk between each phase of the AI lifecycle and records and information management (RIM) requirements, identifying the records needed to support retention, disposition, preservation, and access.

In *Navigating Accountability: The Role of Paradata in AI Documentation and Governance*, the authors examined how authoritative AI governance instruments address recordkeeping challenges and the role of paradata in supporting accountability, transparency, and governance[5]. Their analysis distinguished between explicit documentation requirements (e.g., audit trails, risk management plans, model cards, factsheets, and algorithmic impact assessments) and implicit requirements, such as documenting change management, benchmarking, traceability, and model training, validation, and testing. The study concluded that documenting AI systems requires an integrated body of records, metadata,

5 S. Cameron, P. Franks, I. Huvila, N. Mooradian, *Navigating accountability: the role of paradata in AI documentation and governance*, "Journal of Documentation" 2025, vol. 81(4), pp. 906–926, https://doi-org.libaccess.sjlibrary.org/10.1108/JD-01-2025-0009.

paradata, explainable AI (XAI), and related artifacts (e.g., links to external data sources), collectively constituting the AI Lifecycle Record.

Building on these findings, the researchers clarified the complementary roles of XAI, metadata, and paradata. As noted by The National Archives (UK), explainable AI depends not only on algorithms but also on the organizational and operational context in which AI systems function. XAI explains how an AI system produces a particular output from a given set of inputs, whereas paradata documents why, how, and under what conditions the system is developed, implemented, and used. Metadata describes an information resource for purposes of identification, management, preservation, and access, while paradata documents the processes, decisions, and contexts that produce or transform that resource[6]. Together, these forms of documentation provide the processual evidence necessary to support AI transparency, accountability, and governance.

Table 2 presents representative examples of technical and organizational paradata generated throughout the AI lifecycle. Because documentation categories frequently overlap, some examples may also function as records (e.g., AI policies) or supporting documents (e.g., vendor documentation).

Table 2. Examples of Technical and Organizational Paradata to document the AI process that may be created, captured, managed, and preserved

Technical Paradata	Organizational Paradata
AI Model (tested and selected)	AI policy
Evaluation and performance metrics	Design plans
Logs generated	Employee training
Model training data set	Ethical considerations
Training parameters for model	Impact assessments
Vendor documentation	Implementing process
Versioning information	Regulatory requirements

Source: P. Franks, *In the Pursuit of Archival Accountability: Positioning Paradata as AI Processual Documentation* [in:] Society of American Archivists – 2023 Research Forum, p. 6, https://www2.archivists.org/sites/all/files/Franks_In%20the%20Pursuit%20of%20Archival%20Accountability.pdf [access: 31.05.2026].

6 AEOLIAN, The National Archives (UK), Case Study by L. Jaillant, K. Aske, A. Caputo, https://www.aeolian-network.net/wp-content/uploads/2021/11/AEOLIAN-Case-Study-1-The-National-Archives-UK.pdf [access: 28.07.2026].

Technical paradata is system related (e.g., audit logs) and should be, when possible, automatically captured by the AI system. Operational paradata can reflect the context and environment surrounding the AI process, such as the environmental factors and external inputs (e.g., laws and regulations) that influenced the AI's performance. Increasingly organizations employ real-time dashboards to provide a single point of access to everything related to AI, including AI policies, principles, values, vocabulary, taxonomy, and more.

AI Use Cases and Documentation Challenges

During the final phase of the study, 2025 through 2026, four AI use cases were developed to understand how AI techniques could be identified, captured, and preserved as part of the recordkeeping process: 1) Automated Data Recognition and Extraction for Indexing Digitized Images, St. Louis Zoo; 2) Bank of Canada, Architecting Accountability in AI; 3) Governing Clinical AI as Information Infrastructure: A Lifecycle Documentation Architecture of Metadata, Provenance, and Paradata; and 4) Information Framework for Documenting Case Studies. Each case is briefly described here; details can be found in the *Preserving AI Techniques as Paradata Final Report* (2026)[7].

Use Case #1. Automated Data Recognition and Extraction for Indexing Digitized Images, Saint Louis Zoo

The Saint Louis Zoo employs an Electronic Document Management System (EDMS), *DocuWare*, for workflows, active, inactive, and legacy animal records. The Zoo had forty years of legacy records (created from 1969 to 2009) that were microfilmed. The microfilm rolls were then digitized and saved as PDF files with the assistance of graduate-level interns. These records constitute four distinct record series and are classified for permanent retention under the zoo-specific records retention schedule. The anticipated benefits of using AI features were saving employee time and streamlining procedures for interns to assist staff in completing the indexing of legacy animal records.

7 Ibidem.

Testing was conducted on twenty sample animal transaction PDF files in conjunction with the vendor to evaluate the *Intelligent Indexing* features that use machine learning algorithms to classify documents into different types by searching for relevant index terms automatically and making suggestions to the user. The user then confirms the suggestion or suggests an improvement. This information is stored with the training results in a model space and used to index new documents received. Hard-coded rules are also used for the indexing suggestions.

Among the challenges identified during testing was that the Intelligent Indexing feature had difficulties reading or interpreting handwritten information due to image quality and poor handwriting. In addition, splitting PDF files (extracting) to gather individual file information was complicated by the feature's difficulty in recognizing the record series number, whether typed or handwritten to indicate when one record in the PDF file ended and the next one began. In addition, the background interference inherent with the transformation workflow – migration of paper to microfilm and of microfilm to digitized PDF image – presented limitations to using Intelligent Indexing for this project.

Because the AI pilot was not successful, documentation of the use case was categorized under "project files" and retained to improve the evaluation dataset and to prevent repeating the same mistakes. Among the technical paradata captured are AI model, testing dataset, vendor documentation, and evaluation and performance metrics. The operational paradata captured included design plans, employee training, regulatory requirements, and impact assessments. The AI Record will be preserved according to the records retention schedule for project files and disposed of when no longer useful to the organization or required by a records authority.

Although the AI process investigated for this use case was not practical for the intended purpose, the individuals involved reported that the information gained from the evaluation and testing of modern AI tools integrated into the DocuWare EDMS may be useful in the future if applied to born digital records to enhance automation, capture, classification, metadata validation, and general quality control.

Use Case #2. Bank of Canada, Architecting Accountability in AI

The Bank of Canada uses Large Language Models (LLMs) to assist in the analysis of shifting prices in international oil market data. Bank-developed algorithms were combined with “new AI applications to show that the technology could strengthen the Bank’s ability to understand changes in commodities markets and apply these capabilities to policy recommendations. As part of this work, the Bank developed a framework for the application of paradata – a key component in ensuring the »explainability« and ethical application of the AI used in formulating market analysis”[8].

Explainable AI (XAI), for the bank, means making the AI tools understandable and interpretable. Therefore, paradata was deployed to document relationships between AI implementations and their results in order to make them understandable in terms of the systems and individuals responsible and accountable for them. The Bank faced the following key challenges:

1. The Bank cannot build and deploy an LLM that operates within a Blackbox. A mechanism to resolve AI explainability challenges must be devised.
2. A minimal set of paradata to be created for each LLM is needed. Because Paradata covers a wide scope of functionality in AI applications, a specific mandatory subset was necessary.
3. Capturing paradata can be time-consuming. Methods of automating the capture and maintenance of paradata were necessary.

The Bank developed a draft recommendation for both metadata and paradata to be captured for each LLM base in the machine learning (ML) lifecycle. The framework consists of four categories:

1. Core Descriptive Metadata: Includes unique identifiers, versioning, creators, and creation dates.
2. Training Data Paradata: Tracks specific data sources, sizes, date ranges, licensing, and detailed preprocessing steps like cleaning and tokenization.
3. Architecture and Training Paradata: Documents model types (e.g., Transformers), hyperparameters, hardware (GPUs/TPUs), and software frameworks used.

---

8 Ibidem.

4. Evaluation Paradata and Ethical Considerations: Records performance metrics (F1-score, accuracy), benchmarking against other models, bias mitigation efforts, and known limitations.

The benefit of the application of a standard set of paradata is the assurance that key aspects of AI explainability are supported through both technical and contextual paradata. In terms of explainability, an example of contextual paradata would be the answer to the question: How did economists determine the initial definitions for the labels used in the sentiment analysis of oil-market news articles?

A limited number of news articles from the *Daily Oil Market Bulletin* was used for the initial pilot test. A team of data scientists and economists developed the LLM, and the Bank explored the types of paradata needed to ensure explainability of the AI application. As of this writing, the Bank is still determining the exact technical components to include in the framework, and methods to automate the capture of paradata are continuing. To reduce manual burden and ensure consistency, the Bank emphasizes automating paradata collection by integrating with tools like Git, MLflow, and DVC. Furthermore, Prompt Engineering – designing and refining input prompts to effectively guide the behavior of AI models to produce desired outputs – is treated with the same rigor as code; prompts are version-controlled and associated with metadata regarding their intent, task, constraints, and target model.

Use Case #3. Governing Clinical AI as Information Infrastructure: A Lifecycle Documentation Architecture of Metadata, Provenance, and Paradata

To explore the concept of AI in clinical practice, frameworks that identify documentation requirements were examined. An information-science framework-integration approach was employed to design a lifecycle documentation architecture for governance-relevant information artifacts in clinical AI systems. The primary objective was to specify how documentation – spanning metadata, provenance, and paradata – can be systematically organized to operationalize lifecycle governance.

Seven widely used clinical AI reporting frameworks (CONSORT-AI, SPIRIT-AI, TRIPOD-AI, STARD-AI, DECIDE-AI, MI-CLAIM, MINIMAR) were treated as sources from which documentation requirements were extracted and reorganized into a coherent information infrastructure model. Particular emphasis is placed on paradata as an infrastructural trace layer that documents interaction, monitoring, and change over time – a layer largely absent from existing reporting standards. Part of the process required extracting 296 reporting elements and consolidating them through deductive mapping to the National Institute of Standards and Technology's *Research Data Framework (RDaF 2.0)*, followed by iterative framework integration. The resulting architecture comprises 113 distinct documentation elements organized across five lifecycle domains and organized as a three-tier, role-specific architecture corresponding to governance and accountability requirements, scientific and technical reporting, and contextual and operational paradata documentation. Examples of the reporting phases identified, along with documentation requirements for AI applications in clinical healthcare, are shown in Figure 2.

Figure 2. Segment of the checklist for AI applications in healthcare illustrating the reporting phases and elements for comprehensive metadata and paradata. Checklist author: Nancy Powell

Reporting Phases and Examples of Metadata and Paradata for AI Applications in Clinical Healthcare

Study Reporting	Data Collection & Processing	Model Development	Model Training, Testing, & Validation	Model Evaluation
Title and Abstract	Data Sources	Model Details	Model Inputs & Outputs	Explainability
Authorship	Demographics	Trial/Study Design	Outcomes	Generalizability
Objective	Eligibility	Sample Size	Benchmarking	Transparency
Funding Sources	Recruitment	Index Text Rationale	Model Optimization	Ethics
Data Monitoring	Study Setting	Defining Predictors & Outcomes	Participant Analysis	Risks, Limitations, & Biases
Governance	Biological Specimen	Defining Inputs & Outputs	Diagnostic Accuracy	Data & Code Sharing
Auditing	Data Management	Allocation Sequence Generation	Replicability (internal validation)	Usability
Roles and Responsibilities	Data Preparation	Randomization	Adverse Events	Replicability (external validation)

Source: InterPARES Trust AI, Preserving AI Techniques as Paradata Final Report, 2026, RP04-PARADATA-FINALREPORT-2026(3-30-26).pdf [access: 28.07.2026][9].

"Governing Clinical AI as Information Infrastructure: A Lifecycle Documentation Architecture of Metadata, Provenance, and Paradata"[10], an article in the development stage, will provide details on this recommendation for a consolidated framework for AI documentation in clinical healthcare and include the complete checklist.

9 Creative Commons Attribution: All results of the research studies undertaken as part of InterPARES Trust (hereinafter ITrust) are made freely available for use through a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International Public License (https://creativecommons.org/licenses/by-nc-sa/4.0/).

10 N. Powell, P. Franks, S. Ghosh, *Governing Clinical AI as Information Infrastructure: A Lifecycle Documentation Architecture of Metadata, Provenance, and Paradata* [internal working document].

Use Case # 4. Information Framework for Documenting Case Studies

The fourth use case emerged from a collaboration between the Paradata team and a second InterPARES Trust AI team (hereafter CU05), which operated from 2021 to 2026. The project investigated the implementation of AI in archival functions, particularly records classification and metadata enrichment. During the second phase of the project, they endeavored to understand how "AI projects could be described and documented by using (but also, if necessary, integrating) the Paradata Framework developed by the study on *Preserving AI Techniques as Paradata*[11].

CU05 reviewed literature on paradata and analyzed additional information elements to provide insights into the process and methodology used in AI implementation, as well as the decisions made and challenges faced. They developed a documentation framework based on two primary sources of information:

1. Seven phases of the Machine Learning life cycle[12], which encompasses the process activities involved in developing and deploying AI models: 1. Identification and preparation of datasets; 2. Producing AI model; 3. Training AI model with datasets prepared; 4. Evaluating AI model performance; 5. Implementing AI model; 6. Improving AI model with new data; and 7. Monitoring operations.
2. Three phases of data collection based on the Consultative Committee for Space Data Systems (CCSDS) Information Preparation to Enable Long-Term Use (IPELTU)[13] recommendations: 1. Activity Initiation and Planning (Establishing the project's objectives, timelines, and requirements); 2. Activity Executing (Carrying out the activities involved in developing and

11 InterPARES Trust AI, The role of AI for records management: findings from case studies by S. Allegrezza, G. Bezzi, M. Mata Caravaca, M. Grandi, M. Guercio, B. La Sorda, F. Magnoni, M. Tascone, February 2026, p. 6, https://interparestrustai.org/assets/public/dissemination/CU05_TheroleofAIforrecordsmanagement_Findingsfromcasestudies_FINAL_v2.pdf [access: 27.07.2026].

12 S. Cameron, B. Hamidzadeh, *Preserving paradata for accountability of semi-autonomous AI agents in dynamic environments: An archival perspective*, "Telematics and Informatics Reports" 2024, vol. 14, https://doi.org/10.1016/j.teler.2024.100135.

13 The Consultative Committee for Space Data Systems (CCSDS), Information Preparation to Enable Long-Term Use: Recommended Practice, December 2024, https://ccsds.org/Pubs/653x0m1.pdf [access: 27.07.2026].

deploying the AI model); 3. Activity Closing (Concluding the project once all requirements, phases, or contractual obligations have been completed).

An example of the first phase of the documentation framework from the CU05 report illustrating the structure of the Paradata Framework with the main categories of information elements is shown in Figure 3.

Figure 3. Excerpt from the *Information Framework for Documenting AI Projects* displaying the documentation requirements for phase 1 of the seven-phase life cycle, Identification and preparation of datasets

INFORMATION FRAMEWORK FOR DOCUMENTING AI PROJECTS			
(Based on the AI application life cycle and project phases; Adaptation by InterPARES Trust AI, CU05 Study, 2025)			
Lifecycle of AI Application (From: InterPARES Study Trust AI, RP04 Study 2024)	**Phases of Data Collection** (From: IPELTU 2025)		
	1. Planning	**2. Executing**	**3. Closing**
1. Identification and preparation of datasets	1) Information about the purposes and goals to be achieved by getting the datasets processed by AI model 2) Information about the selection and preparation of the datasets 3) Information about the agents (e.g. organizations, other IT systems, people, public authority) that have made available the datasets 4) Information on the constraints (deriving from copyright, privacy, contractual, other legal, technology issues, etc.) bearing on the datasets	1) Information about the datasets to be processed by the AI model (structure; type of data of the datasets; history of the datasets) 2) Information about the control, cleaning and adjustment of the datasets 3) Logs of any operation performed on datasets	1) Outcome of the quality checks performed on the datasets to be processed by the AI model 2) Compliance statements (e.g. statements where the fact that the datasets are deemed to be fit for the purpose is attested to; statements that the datasets to be processed comply with the relevant legislation) 3) Information on the roles and offices involved in the identification and preparation of the datasets

Source: InterPARES Trust AI, CU05 Final Report...

The CU05 team next applied the developed schema to a *NATO Archives-RecordPoint* case study to learn how an AI project can be documented from practitioners' perspectives and to test the framework in a real-world case scenario. Because this attempt occurred ex-post, some of the identified "additional information" of the framework were not available. Among the information identified and collected were the initial project proposal, a log of the records used for the study, the logs produced by the machine learning model exercises, when available, and the machine learning report written by the vendor, RecordPoint, with the aim of mapping the information against the identified AI project phases and stages. In addition, the team drafted the various stages of the project with the available information and its processes in a descriptive form. The complete *Information Framework*, along with a detailed description of the CU05 project, can be found in the CU05 final report, *The role of AI for records management: findings from case studies*.

The use cases described present not only four different uses of AI to solve specific operational challenges but also three major categories of paradata generation and identification cited by Juneström and Huvila (2026) – prospective, in situ, and retrospective[14]. Researchers selected the approach that best suited their needs. Those involved in the *Automated Data Recognition and Extraction for Indexing Digitized Images, St. Louis Zoo* use case took an in situ approach by documenting as they completed each step of the AI process; researchers involved with the *Bank of Canada, Architecting Accountability in AI* and the *Governing Clinical AI as Information Infrastructure: A Lifecycle Documentation Architecture of Metadata, Provenance, and Paradata* use cases took a prospective approach by identifying the paradata to be created and captured prior to implementing an AI project; and researchers in the *Information Framework for Documenting Case Studies* use case described their approach to documenting paradata as ex-post (retrospective).

AI Lifecycle Records

AI Lifecycle Records comprise the records, metadata, paradata, and related documentation generated by and about AI systems throughout their lifecycle. The concept of AI Lifecycle Records is central to records and information management (RIM) in the era of artificial intelligence. AI systems generate both records created through AI-assisted activities (e.g., transcripts of virtual meetings) and records that document the design, development, deployment, and operation of AI systems (e.g., audit logs, model documentation, and other forms of paradata). Together, these records provide the evidence needed to support transparency, accountability, compliance, and governance throughout the AI lifecycle. The following sections examine four interrelated topics that underpin AI Lifecycle Records: AI Lifecycle Models, AI Documentation Guidance, the AI System Lifecycle, and AI-related RIM Responsibilities.

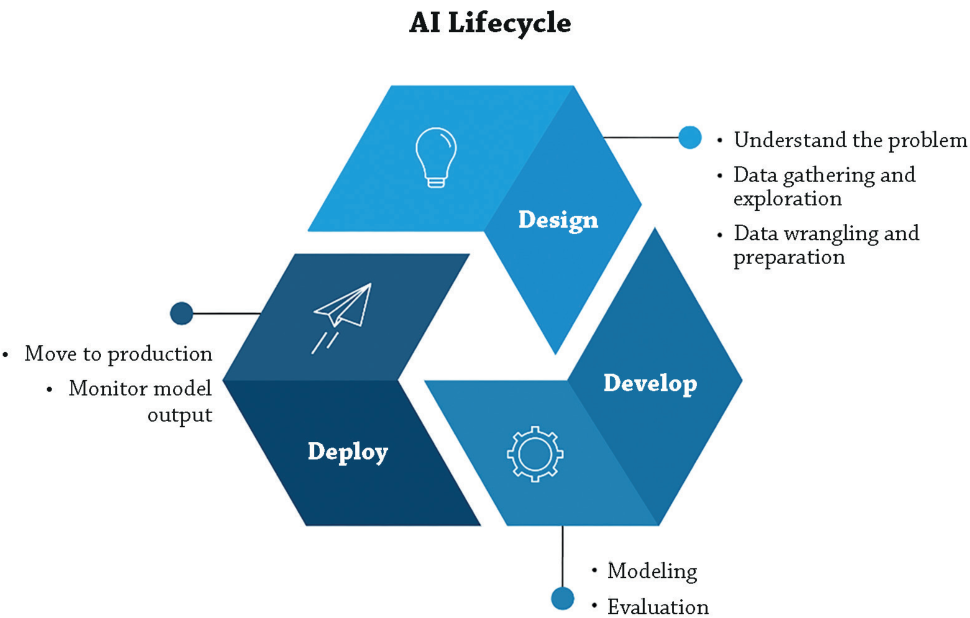
---

[14] A. Juneström, I. Huvila, *Categorizing methods and approaches for generating and identifying paradata*, "Journal of Librarianship and Information Science" 2026, vol. 58(2), https://doi.org/10.1177/09610006251342811.

AI Lifecycle Models

According to the U.S. AI Guide for Government, the "AI lifecycle is the iterative process of moving from a business problem to an AI solution that solves that problem. Each of the steps in the lifecycle (shown in Figure 4) is revisited multiple times throughout the design, development, and deployment phases"[15].

Figure 4. AI Lifecycle with three primary phases: Design, Develop, and Deploy



Source: Center of Excellence, U.S. General Services Administration, Understanding and managing the AI lifecycle, https://coe.gsa.gov/coe/ai-guide-for-government/understanding-managing-ai-lifecycle/ [access: 28.07.2026].

The three major phases – Design, Develop, and Deploy – can be divided into subcategories to more clearly define the process. The actions taken and decisions made throughout the lifecycle must be documented. However, this figure is not the only AI Lifecycle Model in use. An organization may wish to adapt another existing model to suit its needs or design its own model based on the AI process they have developed. Some of the published models are specific to the

[15] Centers of Excellence, U.S. General Services Administration, Introduction to the AI Guide for Government, https://coe.gsa.gov/coe/ai-guide-for-government/introduction/index.html [access: 28.07.2026].

type of AI tool employed (e.g., Machine Learning Lifecycle, Agentic AI Lifecycle, Generative AI Lifecycle). Lifecycle models can be designed for specific industries. For example, the U.S. Digital Health Center of Excellence (DHCoE) mapped the phases of a traditional Software Development Life Cycle (SDLC) to the specifics of AI software development to create its own seven phase AI lifecycle (AILC). Their initial mapping also identified key activities, compiled from literature and reviews of consensus standards, covering each phase of the AILC[16].

AI Documentation Guidance

Although the study examined a broad range of AI laws, regulations, and governance frameworks, two have emerged as foundational references for enterprise AI governance worldwide: the European Union Artificial Intelligence Act (EU AI Act) and the U.S. National Institute of Standards and Technology (NIST) AI Risk Management Framework (AI RMF). Together, they represent the dominant approaches to governing AI systems and have significantly influenced organizational AI governance strategies.

While both frameworks adopt a risk-based approach to AI governance, they differ in purpose and implementation. The EU AI Act establishes legally binding requirements for organizations developing or deploying AI systems within its scope, whereas the NIST AI RMF is a voluntary framework that provides organizations with guidance for identifying, assessing, and managing AI-related risks. Despite these differences, both emphasize comprehensive documentation, lifecycle oversight, and accountability, making them valuable foundations for developing defensible AI documentation and governance practices.

The European Union

The European Union Artificial Intelligence Act (EU AI Act) was adopted on March 13, 2024, with most provisions taking effect by August 2, 2026. The Act applies to organizations established in the European Union and to organizations

---

[16] U.S. Food & Drug Administration, Blog: A Lifecycle Management Approach toward Delivering Safe, Effective AI-enabled Health Care by T. Tazbaz, J. Nicol, 25 July 2024, https://www.fda.gov/medical-devices/digital-health-center-excellence/blog-lifecycle-management-approach-toward-delivering-safe-effective-ai-enabled-health-care [access: 28.07.2026].

outside the EU when AI systems are placed on the EU market or their outputs are used within the EU[17]. It establishes a risk-based regulatory framework comprising four categories, each with documentation obligations that increase according to the level of regulatory risk:

1. Unacceptable risk – AI applications that threaten fundamental rights, including systems that exploit vulnerabilities related to age, disability, or socioeconomic status, are prohibited.
2. High risk – AI systems deployed in regulated domains, including critical infrastructure, education, employment, essential services, law enforcement, and biometric identification, require comprehensive technical documentation demonstrating compliance with the Act.
3. Limited risk – AI systems posing transparency risks, such as chatbots and AI-generated content, must disclose AI involvement and clearly label specified synthetic content, including deepfakes.
4. Minimal risk – most AI applications, such as spam filters and AI-enabled video games, are subject to minimal regulatory requirements.

The Act also establishes transparency requirements for General-Purpose AI (GPAI) models, which underpin many generative AI applications. To support compliance, the European Commission published a voluntary GPAI Code of Practice and a Model Documentation Form (June 10, 2025) specifying the information providers should document under the Act's transparency provisions. The eight documentation categories derived from this guidance are summarized in Table 3.

Table 3. Documentation categories and examples of each

Documentation Categories	Examples of Supporting Documentation (Paradata)
General Information	Model provider, model name, model authenticity (secure hash or URL endpoint), release date (any distribution channel), union market release date, dependencies (version number and model name of previous model or N/A)
Model Properties	Model Architecture (e.g., transformer architecture), design specifications (100 words max), input modalities and maximum size (e.g., text, images, audio, video), total model size (parameters)

[17] European Commission, AI Act, 2024, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai [access: 28.07.2026].

Documentation Categories	Examples of Supporting Documentation (Paradata)
Method of Distribution and Licenses	Distribution channels for each listed method of distribution. Include a link to information about how the model can be accessed, where available, and the level of model access (e.g., weights-level access, black-box access); A link to or copy of model license(s), type of license (free open-source or proprietary), and a list of additional assets (e.g., training data, data processing code, model training code, model inference code, model evaluation code), if any, that are made available with a description of how each can be accessed and what licenses, if any, relate to their use
Use	Acceptable use policy, intended uses, types and nature of AI systems in which the general-purpose AI model can be integraed, technical means for model integration, required hardware, required software
Training Process	Design specifications of training process (training methodologies and techniques), decision rationale (why were decisions made – 200 words max)
Information on the Data used for Training, Testing, and Validation	– Data type/modality (e.g., text, images, audio video) – Data Provenance: Web crawling; private, non-publicly available data sets from third party; user data; publicly available data sets, synthetic data that is not publicly accessible (when created directly by or on behalf of the provider), and data collected through other means – How the data was obtained and selected – Number of data points – Scope and main characteristics – Data curation methodologies – Measures to detect unsuitability of data sources – Measures to detect identifiable biases
Computational Resources	Training time, amount of computation used for training, measurement methodology
Energy Consumption	Amount of energy used for training and measurement methodology. Benchmarked amount of computation used for inference and measurement methodology

Source: European Commission, EU Model Documentation Form, https://ec.europa.eu/newsroom/dae/redirection/document/118118 [access 31.05.2026].

The United States

By contrast, the United States lacks a federal statute; therefore, it depends upon an eclectic mix of state-level laws (e.g., a bill to establish a uniform definition for artificial intelligence (AI) in California; a real estate responsibility law in South

Carolina that makes licensed real estate professionals responsible for any and all work product produced with the assistance of artificial intelligence, machine learning, or similar programs; and an amendment to the Illinois's Human Rights Act passed in 2024 that requires employers to fully disclose how they use AI tools in Human Resources decisions)[18]. On the Federal level, Presidential Executive Orders and Actions influence the development and use of AI. On December 11, 2025, President Donald J. Trump declared that until a national AI standard is in place, his Administration would take "action to check the most onerous and excessive laws emerging from the States that threaten to stymie innovation"[19]. Those responsible for documenting the use of AI are required to be compliant with the laws and regulations in place at the time of design and implementation and to keep abreast of any changes that require modification to their documentation strategy.

Currently, the most popular framework to guide documentation for enterprise AI in the U.S. is the AI Risk Management Framework, released by the National Institute of Standards and Technology on January 26, 2023[20]. It provides guidance to mitigate risks associated with the development and implementation of AI. It is intended to improve the ability to incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems. This framework is supported by a living document, the *AI RMF Playbook*, that suggests actions for achieving the outcomes laid out in the framework[21]. On July 26, 2024, NIST released NIST-AI-600-1, *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile* to help organizations identify unique risks posed by generative AI and proposes actions for generative AI risk management that best aligns with their goals and priorities[22]. On April 7, 2026, NIST released a concept note for an

---

18 Orrick, AI Law Center, https://ai-law-center.orrick.com/ [access: 28.07.2026].

19 The Whitehouse, Ensuring a National Policy Framework for Artificial Intelligence, 11 December 2025, https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/ [access: 28.07.2026].

20 NIST (National Institute of Standards and Technologies), Artificial Intelligence Risk Management Framework (AI RMF 1.0), https://doi.org/10.6028/NIST.AI.100-1 [access: 28.07.2026].

21 Ibidem, NIST AI RMF Playbook, https://airc.nist.gov/airmf-resources/playbook/ [access: 28.07.2026].

22 Ibidem, NIST AI 600-1 Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, https://doi.org/10.6028/NIST.AI.600-1.

AI RMF Profile on Trustworthy AI in Critical Infrastructure[23]. The profile will guide critical infrastructure operators towards specific risk management practices to consider when engaging AI-enabled capabilities. The NIST AI Risk Management Framework introduces four core areas of responsibility: Govern, Map, Measure, and Manage.

The Govern Core function provides guidelines for the implementation of structures, systems, processes, and the human factor[24]. It is further divided into six categories and nineteen subcategories. Examples of the type of documentation recommended within the Govern function include applicable laws and regulations, AI policies, risk management policies, AI system documentation, compliance reviews, impact assessments, data governance practices, and employee training programs. Each of the can be divided into specific entities more clearly recognized as paradata. For example, AI system documentation may include AI system inventory, data dictionaries, and incident response plans.

The Map function is process focused. It requires categorizing AI systems, understanding their capabilities and use, identifying risks and benefits, and characterizing impacts to individuals, groups, communities, organizations, and society[25]. It consists of five categories and eighteen subcategories. Examples of documentation that supports the Map function are AI goals and mission, information disclosure plan, risk profile (tolerance), AI system profile (tasks and implementation methods), context maps (to laws, norms, expectations, settings), datasheets for datasets, trustworthiness assessments (e.g., security, privacy), and TEVV (testing, evaluation, validation, verification) metrics.

The Measure function employs quantitative, qualitative, or mixed-method tools, techniques, and methodologies to analyze, assess, benchmark, and monitor AI risk and related impacts. The knowledge relevant to AI risk from the Map function informs the Measure function[26]. This function consists of four categories and twenty-two subcategories. Examples of documentation that supports the Measure function include system test results to identify limitation and errors, usability testing, stakeholder engagement plans, upstream and

---

23 Ibidem, Concept Note: Artificial Intelligence Risk Management Framework: Trustworthy AI in Critical Infrastructure Profile by R. Sheh and M. Stanley, https://www.nist.gov/document/concept-note-artificial-intelligence-risk-management-framework-trustworthy-ai-critical [access: 28.07.2026].

24 S. Cameron, P. Franks, I. Huvila, N. Mooradian, *Navigating accountability*..., pp. 906–926.

25 Ibidem.

26 Ibidem.

downstream dependencies, algorithmic methodology, and metrics to monitor, characterize, and track external inputs, including third-party tools[27].

During the Manage function, risks that have been mapped and measured, are prioritized based on their impact, likelihood of occurrence, and the resources or methods available for mitigation[28]. This function consists of four categories and thirteen subcategories. Examples of the type of documentation recommended to support the Manage function are impact assessments, incident response plans, model cards and factsheets, datasheets for datasets, AI governance structure, procedures for retirement and decommissioning of the AI system, post-deployment performance evaluation (testing methods, metrics, and performance outcomes), and third-party contracts and terms of service as well as third party documentation (including those addressing explainability or interpretability of models, source of datasets, potential vulnerabilities, risks or biases of systems).

Regardless of the AI tools implemented and the requirements (mandatory or voluntary) governing the use of AI, records and information management responsibilities will manifest themselves across each of the major AI lifecycle phases.

AI System Lifecycle

The outer most ring of the AI System lifecycle (Figure 5) contains the three major phases introduced by the U.S. Federal Government, GSA, Centers of Excellence. The six steps of the AI System lifecycle shown in the intermediate ring of the model represents the AI System Lifecycle as described in the OECD AI Principles[29].

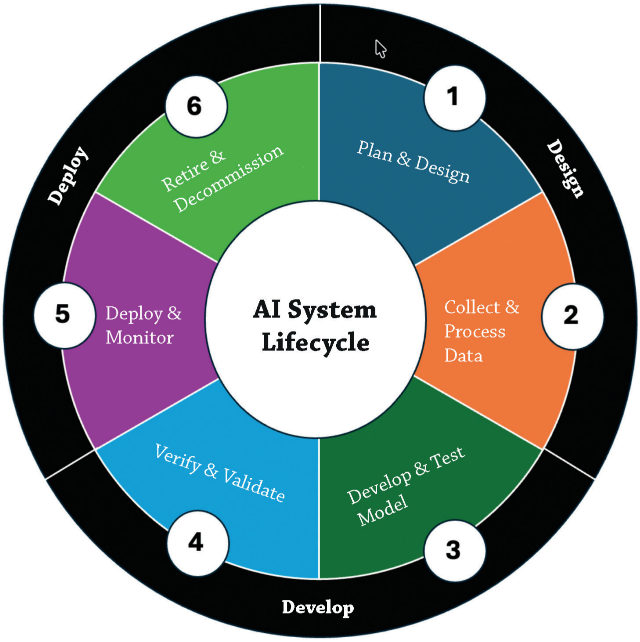
A list of questions to be asked for each phase of the AI process, along with examples of the type of documentation necessary to provide evidence of the actions taken and decisions made, is provided in Table 4. The list is not exhaustive, and the questions and examples of evidence have been gathered from across the frameworks investigated.

27 Ibidem.

28 Ibidem.

29 OECD, OECD AI Principles, https://clairk.digitalpolicyalert.org/documents/oecd-ai-principles-3-may-2024-version/raw [access: 28.07.2026].

Figure 5. The actions performed throughout the AI system lifecycle within the design, development, and deployment steps of the AI process



Source: Author's own elaboration.

Table 4. Examples of questions to be asked and documentation required for each of the six phases of the AI Lifecycle

AI Lifecycle Phase	Questions to be Asked	Documentation as Evidence
Plan and Design	Which AI use case has been identified? What retention schedules and classification standards will be applied to AI records generated?	– AI concept models, prompts, expected outcomes, and retention and disposition schedule – IBM Factsheets
	What compliance, transparency, and risk controls can be embedded from the start?	– AI Risk Management Policy – Review of the AI system for compliance with applicable laws, regulations, standards, and guidance
Collect and Process Data	What training dataset was used, and where is it accessible or preserved?	– Datasheets for Datasets – Data Retention Schedule (could be part of Records Retention and Disposition Schedule)

AI Lifecycle Phase	Questions to be Asked	Documentation as Evidence
	What limitations or qualifications are associated with this dataset, and what risks may emerge from using this dataset in this context?	– Data Privacy Impact Assessment – Datasheets for Datasets
Develop and Test Model	What AI model or tool was used, and can this tool be preserved or documented?	– AI Model Registry – Google Model Cards – Machine Learning Security Operations (MLSecOps)
	Which levels of model evaluation – model testing, red teaming, large-scale field testing – were utilized? (Complex Discovery 2024).	– Completed templates, validated measurement instruments, and automated analysis tools – Automated common T&E activities using tailorable tools
Verify and Validate	What activities were employed to review and evaluate software code, architecture, and documentation to identify errors, inconsistencies, or non-compliance with programming standards?	– Records of software verification activities include Code Review by developers and testers, unit testing, integration testing, static analysis, and design verification
	How well does the software meet the needs and expectations of the end user, that is, perform the functions for which it was designed accurately and completely?	– Records of various types of testing for software validation: functional testing, requirements validation, regulatory compliance, and documentation validation
Deploy and Monitor	How can implementers monitor the AI tool's use over time to ensure its ongoing efficacy and avoid causing harm?	– Identify and document relevant performance metrics to track and assess the tool's effectiveness in relation to organizational objectives and desired outcomes over time – Reports on integration with legacy systems, compliance testing and validation – AI Impact Assessment
	What processual documentation or logs are generated by the AI Tools employed? How are they made available for administrators and researchers for collections management and transparency purposes?	– Access logs, documentation of corrective actions taken – AI hub for AI processual documentation with access based on roles

AI Lifecycle Phase	Questions to be Asked	Documentation as Evidence
Retire and Decommission	How are archival records preserved, and sensitive data securely destroyed when no longer necessary?	– Records Retention Schedule – Decommissioning Plan
	What mechanism will be employed to document decommissioning actions and approvals for future reference?	– AI system compliance records, including those generated during retirement

Source: Author's own elaboration based on results of the InterPARES Trust AI, RP04 FINAL Report: Preserving AI Techniques as Paradata…

AI-related RIM Responsibilities

Although AI Governance structures, including stakeholders, will vary across industries and organizations, Records and Information Management (RIM) professionals can contribute in countless ways during each phase of the AI Lifecycle.

1. During the *Plan and Design* stage, RIM can ensure accountability, transparency, and compliance requirements are defined; govern data, information and records throughout the entire AI initiative; ensure records support evidentiary, accountability, and audit requirements; identify legal, regulatory, risk, and compliance obligations; develop retention schedules for both data and records including new AI-related records; establish access classifications and privacy restrictions; design metadata schemas to support discoverability, traceability, and contextual understanding; manage internal and external documentation, including third-party vendor records; and capture evidence of how AI systems are developed and operate in real time.
2. RIM can contribute to the *Collect and Process Data* stage by collaborating with data stewards, providers, and consumers in data selection, collection, and processing activities; ensuring high quality data characteristics are maintained; defining and implementing retention requirements for AI-related data and records; developing and applying classification schemes based on business activities or record categories; defining retention and access controls for training data, prompt logs, model checkpoints, and AI

agent action logs; reducing ROT (redundant, obsolete, trivial data); applying data minimization principles; ensuring records and data practices are legally defensible and auditable; and managing AI-related records (e.g., including datasets, logs, checkpoints, and other AI artifacts).

3. During the *Develop and Test* phase, RIM can contribute by ensuring that all evidence of AI model development and testing is properly documented and preserved; maintaining documentation of data provenance (origin, lineage, and history of data); managing human-written and AI-generated code as records; supporting configuration management activities including drift detection and compliance automation; preserving validation and testing records; ensuring metadata compliance; ensuring AI-related content, datasets, code, and testing records follow established classification schemes and standards; and supporting documentation practices that enhance explainability, traceability, and trust in AI systems, especially where Explainable AI (XAI) is used.
4. RIM can assist with the *Verify and Validate* phase by ensuring that all impactful actions and decisions throughout verification and validation are documented; updating and maintaining model cards created earlier in the AI lifecycle; capturing, retaining, and making available performance assessments, including bias testing, hallucination analysis, and red-team penetration testing; and managing and preserving AI-related records associated with risk assessments and risk mitigation.
5. During the *Deploy and Monitor* phase RIM responsibilities may include capturing, securing, classifying, and maintaining operational records generated during AI deployment and inferencing; managing records in accordance with established data and records retention protocols; preserving operational records such as decision and action logs, user prompts, autonomous tasks executed, and error and event logs; ensuring access rights and permissions align with established security classifications; supporting eDiscovery readiness for AI-related records, particularly those documenting AI agent decisions and actions that significantly affect outcomes; and preserving evidence of AI system performance, operational activities, and accountability during deployment and production use.
6. RIM responsibilities during the *Retire and Decommissioning* phase include managing the disposition of AI-related records in accordance with retention requirements; ensuring the secure destruction of records no longer required;

preserving historically significant records, such as system documentation, version histories, and audit logs; documenting decommissioning decisions and rationale; ensuring record portability for successor systems, audits, and compliance activities; and maintaining comprehensive lifecycle documentation, including data provenance, model configurations, decision rationales, modifications, datasheets, and model cards. RIM can also ensure that all required AI documentation is appropriately captured, preserved, and responsibly disposed of when no longer needed.

Conclusion

Over the five-year period examined in this study, Artificial Intelligence evolved from a subject of discussion to an accepted and increasingly inevitable component of organizational operations. Besides offering potential benefits such as making organizations more productive and helping solve complex problems, the use of AI poses new legal and ethical issues. Government entities across the globe are struggling to understand how to facilitate innovation through the use of AI and at the same time minimize the resulting risks. Some governments enact laws, such as the EU AI Act, which directly impacts twenty-seven member countries. Others, such as the United States, lack overarching federal legislation and rely on guidance documents, such as the NIST AI Risk Management Framework, along with AI laws enacted in member states.

The use cases examined in this study demonstrate the diverse motivations for organizational AI adoption, the workflows developed to achieve project objectives, and the documentation practices employed throughout the AI lifecycle. While one use case encountered challenges due to the poor quality of test records, it provided valuable lessons regarding the application of AI for legacy records. A second use case focused on developing a framework to document AI activities, producing a consolidated documentation framework that identified essential project records and integrated them into a broader paradata framework. This framework may serve as a model for future AI documentation initiatives.

Two additional use cases from the healthcare and banking sectors illustrate the applicability of the framework in different industrial contexts. The healthcare project undertook the complex task of consolidating documentation requirements from multiple sources into a unified framework for AI documentation in clinical settings. Ongoing efforts will be to further evaluate and refine this framework through practical application. In the banking sector, AI was employed to analyze commodity market trends and support policy recommendations. This project relied on several categories of paradata to document the AI process, although additional refinement and organizational approval of the paradata schema are required before the initiative can progress beyond the pilot stage. Collectively, these use cases demonstrate the value of structured documentation, metadata, and paradata in supporting transparency, accountability, and governance across the AI lifecycle.

The creation, capture, management, and preservation of records, metadata, paradata, and related evidence are essential for documenting the design, development, deployment, and use of AI systems. These requirements reinforce the strategic importance of Records and Information Management (RIM) as a foundational component of AI governance, accountability, and organizational memory. As artificial intelligence continues to reshape the information landscape, robust recordkeeping practices are no longer ancillary administrative functions but critical mechanisms for ensuring transparency, explainability, compliance, and trust.

At the outset of this study, several assumptions guided the research: that standardized paradata could address AI documentation challenges; that AI documentation would be primarily technical in nature; that manual documentation, even when performed retrospectively, would be feasible; that AI functioned principally as a tool; and that paradata served only as a supplementary component of the AI record. The findings challenge these assumptions. Rather than relying solely on standardized approaches, effective documentation requires contextual and domain-specific implementation. Organizational and human factors are as critical as technical considerations, while automation and workflow integration are necessary to ensure sustainable documentation practices. AI systems are sociotechnical systems, and paradata is not merely supplementary but fundamental to transparency and accountability.

The findings establish paradata as "both conceptually essential and operationally indispensable for documenting AI systems and related technologies. Paradata provides a framework for identifying, structuring, managing, and preserving the evidence needed to satisfy accountability, legal, archival, and governance requirements. More importantly, it positions documentation as an integral lifecycle function embedded throughout the design, development, deployment, and use of AI systems, rather than as a retrospective compliance activity"[30].

A fundamental shift in Records and Information Management (RIM) practice is required. Archivists and records managers must recognize AI documentation and paradata capture as core recordkeeping responsibilities. Achieving this objective will require increased AI literacy, the integration of paradata requirements into governance and procurement frameworks, an expanded conception of records and evidence, and clearly defined accountability for documentation throughout the AI lifecycle.

Future research should focus on automating AI documentation processes, developing centralized mechanisms for capturing and managing AI-related records, establishing role-based access and accountability models, and revising data and records retention schedules to explicitly address AI-generated and AI-related records. Additional attention is needed to ensure that AI-related records, metadata, and paradata are retained, preserved, and disposed of in ways that support transparency, accountability, regulatory compliance, and the enduring evidentiary value of AI systems.

Bibliography

AEOLIAN, The National Archives (UK), Case Study by L. Jaillant, K. Aske, A. Caputo, https://www.aeolian-network.net/wp-content/uploads/2021/11/AEOLIAN-Case-Study-1-The-National-Archives-UK.pdf [access: 28.07.2026].

Cameron S., Franks P., Hamidzadeh B., *Positioning Paradata: A Conceptual Frame for AI Processual Documentation in Archives and Recordkeeping Contexts*, "ACM Journal on Computing and Cultural Heritage" 2023, vol. 16(4), article 75, pp. 1–19, https://doi-org.libaccess.sjlibrary.org/10.1145/3594728.

Cameron S., Franks P., Huvila I., Mooradian N., *Navigating accountability: the role of paradata in AI documentation and governance*, "Journal of Documentation" 2025, vol. 81(4), pp. 906–926, doi: https://doi-org.libaccess.sjlibrary.org/10.1108/JD-01-2025-0009.

Cameron S., Hamidzadeh B., *Preserving paradata for accountability of semi-autonomous AI agents in dynamic environments: An archival perspective*, "Telematics and Informatics Reports" 2024, vol. 14, https://doi.org/10.1016/j.teler.2024.100135.

[30] InterPARES Trust AI, RP04 FINAL Report: Preserving AI Techniques as Paradata…

Centers of Excellence, U.S. General Services Administration, Introduction to the AI Guide for Government, https://coe.gsa.gov/coe/ai-guide-for-government/introduction/index.html [access: 28.07.2026].

Center of Excellence, U.S. General Services Administration, Understanding and managing the AI lifecycle, https://coe.gsa.gov/coe/ai-guide-for-government/understanding-managing-ai-lifecycle/ [access: 28.07.2026].

European Commission, AI Act, 2024, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai [access: 28.07.2026].

Franks P., *In the Pursuit of Archival Accountability: Positioning Paradata as AI Processual Documentation* [in:] Society of American Archivists – 2023 Research Forum, https://www2.archivists.org/sites/all/files/Franks_In%20the%20Pursuit%20of%20Archival%20Accountability.pdf [access: 31.05.2026].

InterPARES Trust AI, Dissemination tab. (2021–2026), https://interparestrustai.org/trust/research_dissemination [access: 28.07.2026].

InterPARES Trust AI, RP04 FINAL Report: Preserving AI Techniques as Paradata by S. Cameron, P. Franks (lead), K. Fewster, I. Huvila, N. Mooradian, 30 March 2026, https://interparestrustai.org/assets/public/dissemination/RP04-PARADATA-FINALREPORT-2026(3-30-26).pdf [access: 28.07.2026].

InterPARES Trust AI, Terminology Database (2021–2026): Paradata, https://interparestrustai.org/terminology [access: 28.07.2026].

InterPARES Trust AI, The role of AI for records management: findings from case studies by S. Allegrezza, G. Bezzi, M. Mata Caravaca, M. Grandi, M. Guercio, B. La Sorda, F. Magnoni, M. Tascone, February 2026, https://interparestrustai.org/assets/public/dissemination/CU05_TheroleofAIforrecordsmanagement_Findingsfromcasestudies_FINAL_v2.pdf [access: 28.07.2026].

Juneström A., Huvila I., *Categorizing methods and approaches for generating and identifying paradata*, "Journal of Librarianship and Information Science" 2026, vol. 58(2), pp. 1042–1060, https://doi.org/10.1177/09610006251342811.

NIST (National Institute of Standards and Technologies), Artificial Intelligence Risk Management Framework (AI RMF 1.0), https://doi.org/10.6028/NIST.AI.100-1.

NIST (National Institute of Standards and Technologies), Concept Note: Artificial Intelligence Risk Management Framework: Trustworthy AI in Critical Infrastructure Profile by R. Sheh and M. Stanley, https://www.nist.gov/document/concept-note-artificial-intelligence-risk-management-framework-trustworthy-ai-critical [access: 28.07.2026].

NIST (National Institute of Standards and Technologies), NIST AI 600-1 Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, https://doi.org/10.6028/NIST.AI.600-1.

NIST (National Institute of Standards and Technologies), NIST AI RMF Playbook, https://airc.nist.gov/airmf-resources/playbook/ [access: 28.07.2026].

OECD, OECD AI Principles, https://clairk.digitalpolicyalert.org/documents/oecd-ai-principles-3-may-2024-version/raw [access: 28.07.2026].

Orrick, AI Law Center, https://ai-law-center.orrick.com/ [access: 28.07.2026].

Powell N., Franks P., Ghosh S., *Governing Clinical AI as Information Infrastructure: A Lifecycle Documentation Architecture of Metadata, Provenance, and Paradata* [internal working document].

The Consultative Committee for Space Data Systems (CCSDS), Information Preparation to Enable Long-Term Use: Recommended Practice, December 2024, https://ccsds.org/Pubs/653x0m1.pdf [access: 27.07.2026].

U.S. Food & Drug Administration, Blog: A Lifecycle Management Approach toward Delivering Safe, Effective AI-enabled Health Care by T. Tazbaz, J. Nicol, 25 July 2024, https://www.fda.gov/medical-devices/digital-health-center-excellence/blog-lifecycle-management-approach-toward-delivering-safe-effective-ai-enabled-health-care [access: 28.07.2026].

The Whitehouse, Ensuring a National Policy Framework for Artificial Intelligence, 11 December 2025, https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/ [access: 28.07.2026].