Responsible AI in Healthcare: Managing Medical Data under the GDPR and AI Act

Isabel Piedrahíta-Vélez[1], Ulises Cortés[2] and Atia Cortés[1]

[1]Barcelona Supercomputing Center
[2]Universitat Politècnica de Catalunya

Abstract

Artificial Intelligence (AI) has, in recent years, become increasingly embedded in the medical field. By leveraging patient data, AI promises to improve diagnostic processes and treatment plans, and even achieve personalized treatment. However, the processing of highly sensitive medical data has raised significant privacy concerns.

To alleviate these concerns, different policies and regulations have been set in place, and technical controls, governance practices, and internal policies have been proposed. Based on these, the Medical Data Lifecycle (MDLC) emerges as a potential domain-specific methodology for developing privacy-conscious medical AI systems.

While the MDLC promotes privacy-aware development, the UMAPER auditing framework complements it by providing a structured approach to internal system audits. Drawing on established regulatory frameworks such as the GDPR and the AI Act, and informed by real-world audit practices, UMAPER offers a practical, multi-stage process for planning and executing internal audits in medical AI systems. By using internal assessments to verify the actual implementation of privacy-protecting intentions, UMAPER contributes to the ongoing discourse on responsible AI development in high-stakes domains like healthcare.

Keywords

Artificial Intelligence, AI in Healthcare Systems, Medical Data Privacy, Data Protection, Privacy-Centric Data Lifecycle, Auditing, Responsible AI Development, Internal Auditing.

1. Introduction

As artificial intelligence (AI) becomes increasingly embedded in healthcare systems, the conversation has begun to shift from *Can we build it?* to *Should we trust it?* AI-based systems promise improved diagnostics, personalised treatment, and more efficient workflows, but these benefits hinge on the responsible use of sensitive medical data. The potential for the wrongful collection, unauthorised circulation, or misuse of medical data during the design and development process of AI-based solutions is a genuine concern that can result in enormous harm to the very patients it should be helping.

Finding ways to analyse compliance with privacy-preserving practices, as well as their trustworthiness and effectiveness, can aid in developing more privacy-conscious AI. In this context, it becomes fundamental to focus on embedding privacy protection into data collection and model development. However, an equally important parallel concern should be assessing whether these protections are truly being implemented and providing benefits, particularly at scale.

The challenge of trustworthy AI in healthcare cannot be addressed solely through robust development practices. It also demands equally robust auditing mechanisms that evaluate whether the systems developed comply with privacy regulations, are ethically aligned, and are safe in practice. In response to this pressing need, we propose UMAPER [1], an auditing framework tailored to evaluate medical AI projects designed under the Medical Data Lifecycle methodology (MDLC) [1].

This paper builds on the foundation laid by MDLC, a lifecycle framework for processing of medical data responsibly throughout an AI-based system's design, development, and deployment. While MDLC offers developers a structured approach to privacy-aware AI, UMAPER takes a step further by empowering stakeholders to evaluate such a process's outcomes critically. It introduces audit stages

HHAI-WS 2025: Workshops at the Fourth International Conference on Hybrid Human-Artificial Intelligence (HHAI), June 9–13, 2025, Pisa, Italy

isabel.piedrahita@bsc.es (I. Piedrahíta-Vélez); ia@cs.upc.edu (U. Cortés); atia.cortes@bsc.es (A. Cortés)

aligned with MDLC milestones, integrating risk assessment, procedural transparency, and system accountability into the pipeline.

Through this dual-lens, privacy-aware development and systematic auditing, this work addresses a core question: How can we ensure that privacy-conscious intentions translate into measurable adherence to privacy-preserving techniques in real-world medical AI-based systems? By focusing on planning and verification, we offer a framework for evaluating the adherence to and potential effectiveness of privacy-preserving techniques.

2. Methodology

This research builds on prior work [1], in which the Medical Data Lifecycle (MDLC), a methodology for embedding privacy and transparency into the development of AI-based medical applications, was proposed. While MDLC focuses on how such systems should be built, this paper addresses a complementary concern: how to conduct audits to evaluate AI-based systems developed following MDLC, in order to promote accountability and trustworthiness.

Through this work, we attempt to address one central question: How can we ensure that privacy-conscious intentions translate into measurable adherence to privacy-preserving techniques in real-world medical AI systems?

To this end, we propose UMAPER, an auditing framework that can be used as a reference during both internal and external audits. It is based on the classical auditing process and structured around the GDPR, the AI Act, and auditing guides and frameworks proposed by public and private entities.

A qualitative research approach was taken to acquire a deeper understanding of the study area, capture contextual nuance, and provide detailed insights into these complex issues. When possible, domain experts were contacted to gain real-life insights that might have otherwise eluded us, particularly concerning how auditing practices are applied or interpreted in healthcare institutions.

For its development, we comprehensively analysed contemporary data protection regulations within the European Union (EU) and auditing guidelines for AI-based systems, focusing on technical, design, and governance aspects. The research design incorporated the following elements: (a) Systematic Review, (b) Document Analysis, and (c) Domain Expert Interviews.

3. A Primer on Data Governance, Privacy and Transparency

Privacy, transparency, and adequate data governance are concerns that must be addressed throughout the entire lifespan of medical data-driven systems to avoid harming data subjects. These concerns have gained even more prominence due to the introduction of regulatory efforts such as the GDPR [2], and more recently, the AI Act [3], which establish a strict legal framework outlining system expectations around privacy, explainability, and transparency, among others. To meet these legal obligations, systems must adopt privacy-aware design, development, and governance practices, such as privacy by design principles and data minimisation.

Transparency, as defined in Recital 58 of the GDPR, refers to the clarity and accessibility of information provided to data subjects. However, transparency should not be limited to external communication. Internal documentation and record-keeping practices can also be essential in ensuring accountability and trustworthiness. Proper documentation creates what has been described as a *transparency trail* [4], enabling developers and auditors to trace how decisions were made, how data was handled, and how responsibilities were assigned. This not only facilitates audits, both internal and external, but also fosters accountability and collective awareness within development teams. To this end, the AI/ML community has produced a number of well-established documentation techniques, such as [5, 6, 7, 8], which include both general and domain-specific approaches.

A core principle of transparency and privacy awareness is the end-to-end management of these concerns. Lifecycle-based management provides a structured means of incorporating privacy-aware and transparency-oriented activities throughout development. A variety of data lifecycle models exist

in the literature [9, 10, 11, 12]. The MDLC, which we will explain in greater detail in §5, is a lifecycle approach for the responsible development of medical data-driven applications, emphasising privacy preservation.

4. Auditing for Trustworthy AI

Audits are tools for evaluating complex processes that often inquire about the level of compliance with company policies, industry standards, or regulations. There has been strong interest in developing strategies for AI auditing in the context of the GDPR and the AI Act. In June 2024, the European Data Protection Board presented the AI Auditing Project, which aims to develop tools to help the Data Protection Agencies of the different Member States understand and assess data protection safeguards related to the AI Act [13].

External auditing is a review carried out by a third party, usually a competent authority or an auditing firm, such as the well-known Big Four (Deloitte, PwC, EY, and KPMG), which holds the audited company accountable.

Internal auditing is the counterpart of external auditing carried out by the organisation itself, instead of an external entity. It is meant to be an objective evaluation of an organisation's internal controls conducted from within the organisation itself to effectively manage risk [14]. One of their goals is to assess the completeness of, and adherence to, internal controls.

Internal controls are the policies, processes, tasks, behaviours, and other aspects of an organisation that promote the quality of internal and external compliance and facilitate an appropriate response to significant business, operational, financial, compliance, and other risks [15].

Both internal and external auditing are necessary, however, they each have their strengths and weaknesses. For instance, one of the main advantages of external auditing is that the auditors do not share an organisational interest with the audited company, leading to a more objective viewpoint during the audit. However, they may be limited by the auditors' ability to access the internal processes of the audited organisation. In contrast, internal auditors usually have unrestricted access to data sources and tools.

Furthermore, external audits are most commonly post-deployment, which implies that although they can capture risks that have not yet materialised, they are still remedial in some capacity. By contrast, internal auditing can span the entire lifespan of the system. This makes internal auditing more than just a practice test for external auditing. Internal auditing is a transversal evaluation of the design, development, and deployment processes that could potentially aid in proactive interventions to ensure privacy rather than simply informing reactive measures.

Internal auditing usually follows a somewhat standard process; however, each audit's specific requirements will differ. In this section, we will illustrate the usual internal audit cycle and also explore two relevant examples: an external auditing guideline by the European Data Protection Board and the SMACTR internal auditing framework.

4.1. The Traditional Internal Audit Structure

Internal audits provide assurance and consulting services, which can focus on performance or control verification. Assurance engagements are determined unilaterally by the internal auditing needs. In contrast, consulting engagements are determined jointly by the internal auditors and the client, with the specific process and steps varying significantly between engagements. In general, an internal audit will have three main stages that focus on planning, performing and communicating the engagement results. [16]

Planning the engagement is the first stage of an internal audit and usually involves tasks such as understanding the context and purpose of the audit, gathering information to understand the system, identifying and assessing risks, and identifying key controls. With this information, the scope and objectives of the audit will be determined, as well as the test plan to outline how each key control will be evaluated. [17][16]

Performing the engagement focuses on executing the test plan, in order to gather evidence in the form of results and documents to support the auditor's conclusions. Test results will be evaluated to reach conclusions regarding the level of mitigation of the underlying risks in the current system. Finally, the auditor will document any deficiencies identified during the audit to facilitate their discussion during the final audit stage. [18][19]

Communicating the engagement results and monitoring process is the final stage of an audit. However, it is important to emphasise that although there are final communications, there are also interim communications that span all stages of the internal audit. Final communication should provide stakeholders with the results of the audit, including its scope and purpose, timeframe, observations and recommendations, conclusion and overall rating of the system, as well as a management action plan to address the observations made by the auditor.[20][21]

4.1.1. The European Data Protection Board Auditing Checklist

The proposal by the EDPB [13] is a checklist auditing methodology defined in five stages: Model card, system map, moments and sources of bias, bias testing, and optionally, adversarial audit. It emphasises the importance of developing an end-to-end approach for algorithmic auditing, as these systems exist in complex social and organisational contexts and depend on data produced by imperfect and complex individuals and societies. However, although it aims to be an end-to-end audit, it focuses mainly on the pre- and post-processing stages of the algorithmic lifecycle. As a final note, it also focuses specifically on bias assessment, which can be very useful in medicine, as it is one of the main concerns for AI-based systems in the field.

The algorithmic audit proposed by the EDPB is intended to be an iterative interaction between auditors and developers. It provides templates and instructions for this interaction, specifying the information auditors will need from the development team to carry out their work.

In the model card stage, a more thorough model card is proposed. It introduces new fields based on GDPR and AI Act principles, which were not explicitly prioritised in the original [6]. These fields are related to risk level assessment according to the AI Act, type and level of human involvement, explainability profiling, and auditability. The auditor is also encouraged to gather any existing documentation on the system, data reuse and utilisation permissions, data sharing agreements, transparency reports, relevant scientific papers, and more.

The system map puts the algorithm in context. It should illustrate the relationship between the algorithmic model, the technical environment of its deployment, and the decision-making process. The auditor proposes the system map with the help of information gathered during the model card stage, and the development team then reviews it. With this information, the auditor will look for evidence to help answer *yes-or-no* questions related to the identification and transparency of the AI-based component of the system, its purpose, adherence to policy, and the involvement of competent authorities in its development.

In the bias testing stage, auditors will define tests to determine if different types of bias are affecting the system. Auditors will harness statistical analysis, developer consultations, and stakeholder engagement to analyse possible system bias comprehensively.

The adversarial audit stage is considered optional but highly recommended. This stage aims to uncover biases, hidden variables, or proxy features that emerge only in real-world deployments. To achieve this, the auditor will need to gather data about the model's impact at scale, either through web scraping, user interviews, crowdsourcing data, or sockpuppeting.

The EDPB Audit Checklist is also a good example of how external auditing guides can inform internal auditing guides, as shown by the Spanish Data Protection Agency's (*Agencia Española de Protección de Datos*, AEPD by its initials in Spanish) guide for auditing requirements for personal data processing in activities involving AI [22]. It is targeted towards the data controllers and processors themselves, making it a guide for internal auditing, and is consistent in its scope and methods with the external auditing guide proposed by the EDPB.

4.1.2. The SMACTR Framework

The Google Responsible AI Impact Lab has also proposed an end-to-end framework for internal algorithmic auditing, the SMACTR framework. This framework incorporates lessons from auditing practices in other safety-critical and regulated industries, such as aerospace and medical device design and manufacturing. These industries have a long history of auditable processes and design controls that have helped improve safety.

The SMACTR framework comprises five stages plus a post-audit stage. It aims to create internal audits that produce a trail of documentation at each stage of development and enable critical reflection on the system's impact. These audits can also serve as internal educational material on ethical awareness.

The goal of the **scoping stage** clarifies the audit's objective. Auditors review the system's motivations and intended impact to anticipate potential harm and social impact sources.

The **mapping stage** reviews the system's current state, searching for internal stakeholders and relevant collaborators. The Failure Modes and Effects Analysis (FMEA) will be started at this stage and updated and expanded through the following stages.

In the **artefact collection stage**, auditors set out to identify and collect all the necessary documentation from the development process to prioritise testing. If necessary, the auditor may enforce retroactive documentation or document some processes themselves, checking the results with the development team. The outcome of this stage includes the datasheets and model cards, as well as a design checklist that keeps track of the availability of all the expected documentation.

During the **testing stage** the auditing team executes a series of tests to assess the system's level of compliance with the organization's ethical values. Depending on the organisational context, the auditing team's approach to conducting tests will vary at this stage. The main output of this stage is the ethical risk analysis chart, in which the importance of each risk is defined in proportion to its severity and likelihood.

The **reflection stage** is reserved for analysing the results of the tests in the context of the ethical expectations defined in the scoping stage. In this stage, the FMEA is finalised, and artefacts such as a mitigation plan can be proposed. A design history file that contains all the documentation related to previous stages and, finally, a summary of the audit can also be included.

The **post-audit stage** is concerned with keeping track of the implementation of action plans proposed by the auditing team, as well as making final decisions regarding the proposed AI-based application, even to the extent of deciding to scrap the project fully.

5. A Summary of the Medical Data Lifecycle Methodology

The Medical Data Lifecycle is a structured framework designed to guide the planning, design, and implementation of large-scale AI-based systems in healthcare while balancing strict data protection standards in Europe. As a development methodology, it aims to clarify concerns about privacy, transparency, and accountability in the design and development process. The MDLC methodology is the backdrop against which the UMAPER auditing framework was created and the basic use case for it.

Because of this, we must provide a summary of the MDLC's general structure, outlining its stages, tasks, and deliverables. In addition, we will examine the auditable outputs and audit focus associated with each stage.

Auditable outputs refer to the deliverables and documentation produced during a stage that serve as useful assets during the audit process. **Audit focus**, on the other hand, describes the aspects of

a given stage that an auditor may wish to verify. For instance, compliance with privacy principles, adherence to procedures, or evidence of risk assessment.

What makes MDLC particularly valuable is its emphasis on traceability and accountability. Each phase introduces artefacts, such as planning documents, risk assessments, data selection justifications, and audit-ready logs, that create an auditable trail of the decisions and safeguards implemented throughout development. This documentation-centric approach aligns with the explicit requirements of the GDPR and AI Act and with the needs of internal and external audits. These documents form an *accountability trail* that enables compliance checks, ethical oversight, and continuous evaluation.

The MDLC is a non-linear iterative data lifecycle with seven stages plus an intermediate risk assessment stage. Its modular, iterative structure also accommodates evolving legal standards and organisational policies, making it well-suited for deployment in real-world, high-stakes healthcare settings. A diagram depicting the basic transitions between stages in MDLC is shown in Figure 1.

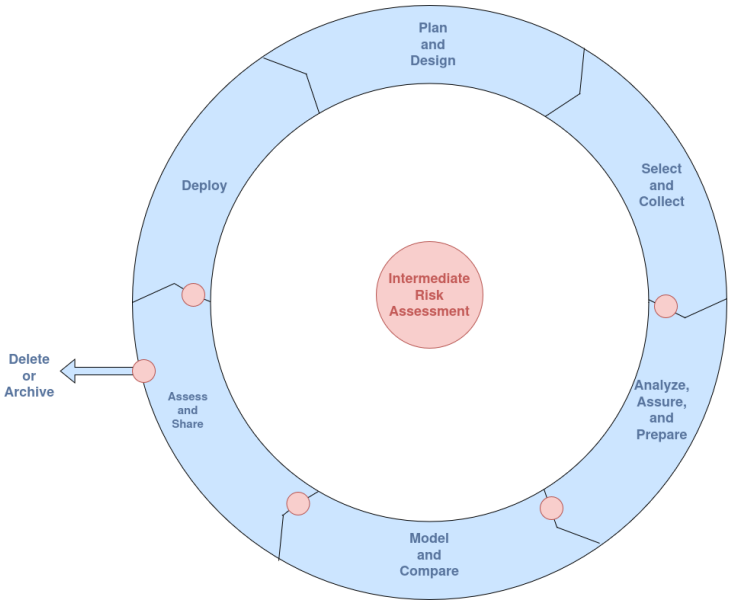

Figure 1: Medical Data Lifecycle Diagram

5.1. Plan and Design

This foundational stage defines the project's scope, objectives, and stakeholders. It emphasises transparency through early documentation of data policies, accepted use cases, and the study of relevant populations. For UMAPER, it establishes the baseline for evaluation. For instance, defining the data policy defines *how* data management will be evaluated. This, however, does not mean that a posterior audit could not point out issues with these plans and policies that would need to be addressed in future developments.

Auditable Outputs: The tasks and outputs of this stage are shown on Figure 2. From these, we can emphasise the usefulness of project scope outputs in evaluating the degree of completion of the project, the organisational structure and policies outputs to guide the verification of adherence to safe data management practices, and the responsible data use foundations as a guide to the concerns identified by the development team.

Audit Focus: The auditor might want to verify the compliance of the proposed data and organizational policies with GDPR and AI Act requirements. They might also want to verify the coherence of project objectives and values.

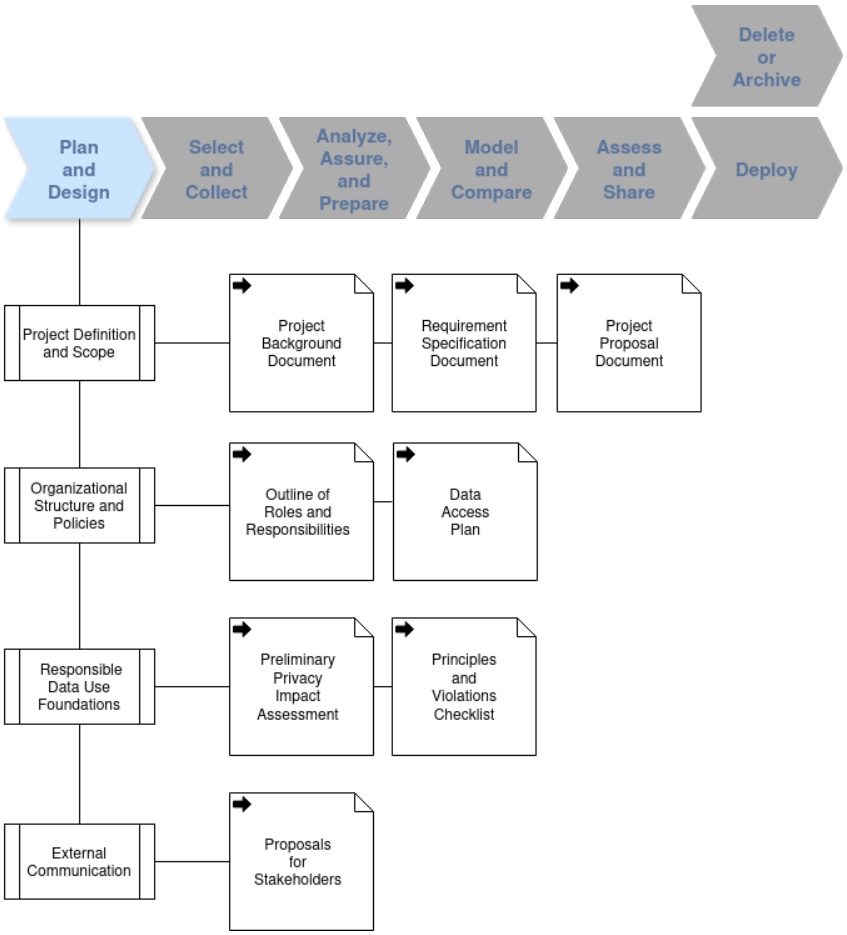

Figure 2: Plan and Design Tasks and Deliverables

5.2. Select and Collect

Data sets are identified and acquired during this stage based on pre-established relevance and privacy criteria. Data selection and data origin and selection reports provide the transparency necessary for internal and external audits regarding data minimisation, data correctness, and verification of informed consent.

Auditable Outputs: The tasks and outputs of this stage are shown in Figure 3. The data selection report details why data was selected, enabling auditors to verify data minimisation. The list of participants will enable auditors to verify whether informed consent exists for all those involved. The outputs of the data collection task will provide the auditor with a base to verify adherence to data policies and task distribution.

Audit Focus: During audits, it will be verified if the data processing conducted in this stage meets compliance and governance expectations, additionally auditors may verify compliance with the internally defined data processing policies during this stage.

5.3. Analyse, Assure, and Prepare

During this stage, data are profiled, cleaned, and evaluated for risks. This stage reinforces privacy guarantees through technical inspections such as risk assessments and de-identification checks.

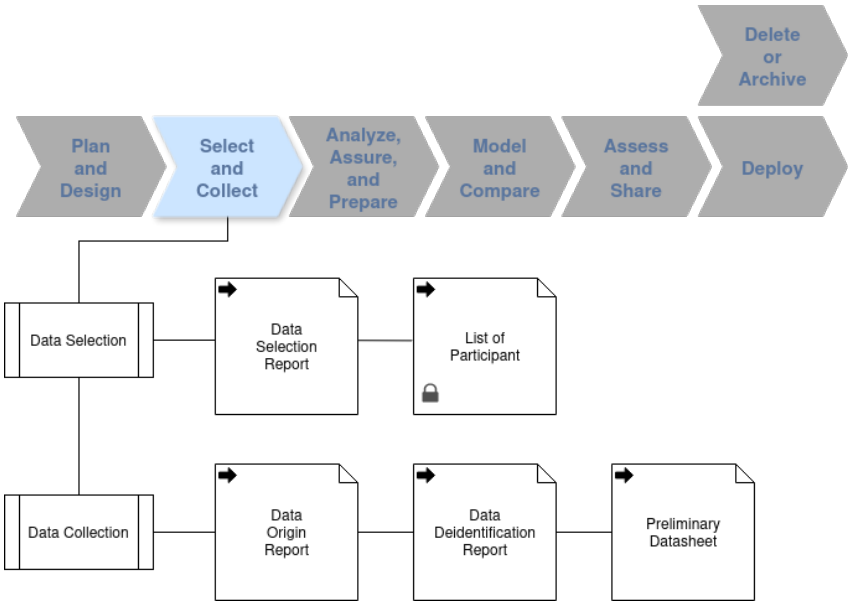

Figure 3: Select and Collect Tasks and Deliverables

Auditable Outputs: The tasks and outputs of this stage are shown in Figure 4. Documents related to data integration can be very relevant during an audit to verify privacy-by-design principles, as centralisation is not encouraged. Additionally, the Datasheet should provide auditors with enough information to understand the created datasets.

Audit Focus: The main concerns of an auditor when reviewing this stage will be related to compliance with de-identification requirements and the sufficiency of data quality, as well as the compliance of data processing with internal and legal requirements.

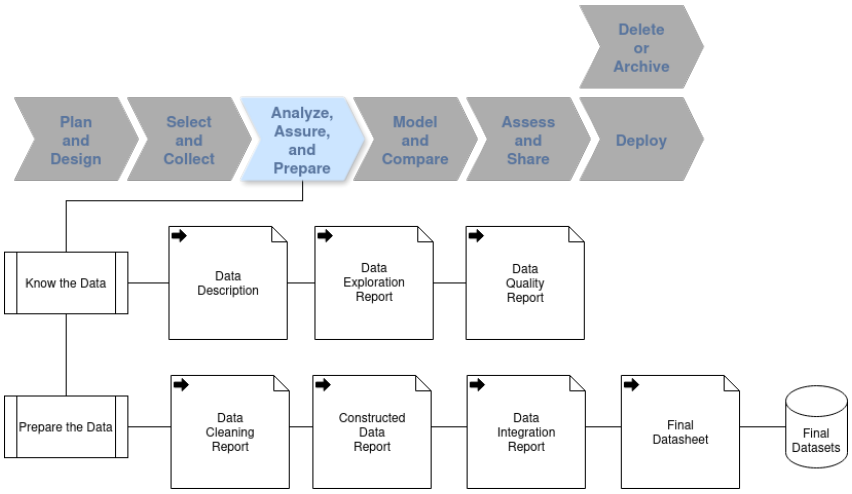

Figure 4: Analyse, Assure, and Prepare Tasks and Deliverables

5.4. Model and Compare

This phase involves developing and evaluating machine learning models. It emphasises reproducibility and fairness by documenting model assumptions, evaluation metrics, and performance across demographic groups.

Auditable Outputs: The tasks and outputs of this stage are shown on Figure 5. Some key documents include the model card, the global and specific performance metrics, and model comparison records which will provide auditors with enough information to assess compliance and performance.

Audit Focus: Auditors might want to verify the degree of performance of the selected model or models. If it has already been deployed, it will also be relevant to check the performance compared to the deployed model. The auditor might also verify the completeness of the development team's testing. Once more, the auditor might want to verify the compliance with internal and legal requirements for the data processing carried out during this stage. Finally, the auditor must verify whether the proposed model falls within the acceptable use cases of AI under the AI Act.

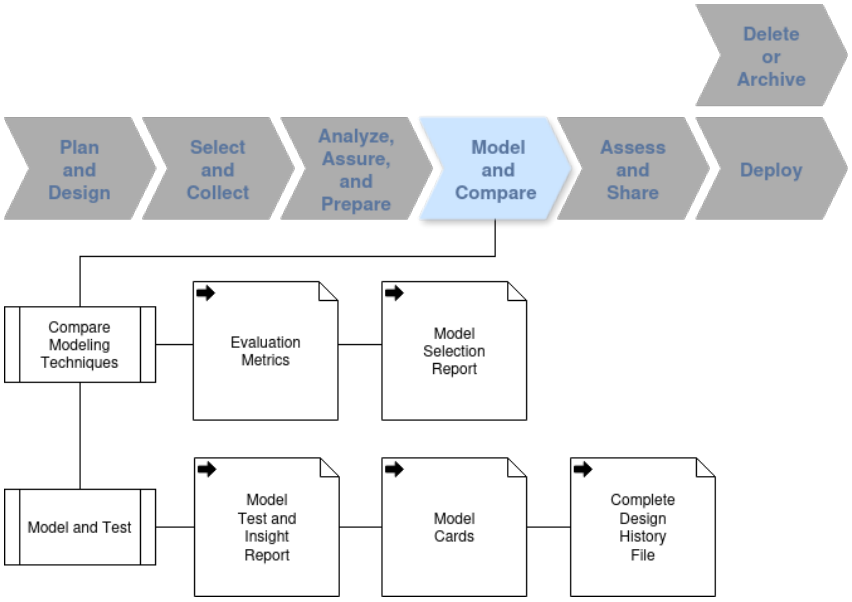

Figure 5: Model and Compare Tasks and Deliverables

5.5. Assess and Share

Before deployment, the system undergoes an internal review to assess legal, ethical, and clinical risks. This aims to assess whether the system meets predefined objectives, adheres to proposed plans and policies, and finally produce fully documented decisions regarding the project's next stages.

Auditable Outputs: The tasks and outputs of this stage are shown in Figure 6. The final project evaluation summaries, compliance self-assessments, and stakeholder feedback can be useful for auditors to verify the development team and stakeholders' perceived adherence to policy. Additionally, suppose the development team or stakeholders set any requirements or constraints for the continuation of the project. In that case, these documents will enable the auditor to review whether or not they were satisfied.

Audit Focus: During an audit, it might be necessary to review the alignment of results with project objectives and development values. Additionally, the level of correctness of compliance self-assessments might also be double-checked through interviews with developers.

5.6. Deploy

In this stage, deployment protocols and access controls are implemented. Deployment logs, system interfaces, and user feedback mechanisms are captured to support post-deployment accountability. This

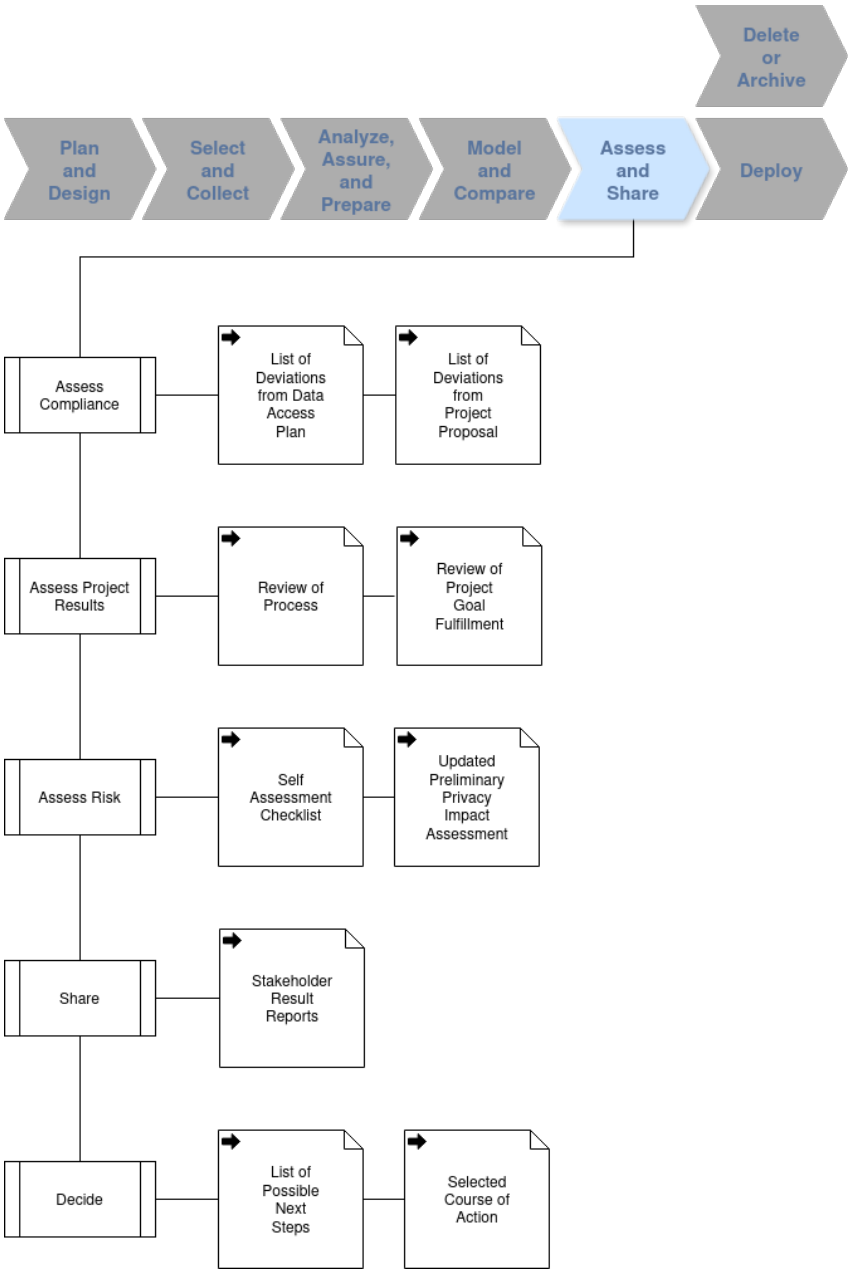

Figure 6: Assess and Share Tasks and Deliverables

stage is vital for UMAPER's evaluation of operational safeguards and user impact assessment.

Auditable Outputs: The tasks and outputs of this stage are shown on Figure 7. The system FactSheet will be a key document capable of giving auditors an overview of the entire system. Additionally, the proposal for monitoring and maintenance, as well as audit triggers, can guide testing during the audit.

Audit Focus: An auditor might want to review the compliance of the deployed system with maintenance and monitoring strategies and audit triggers. Additionally, it is important to verify the compliance of the deployment interface with data protection guidelines.

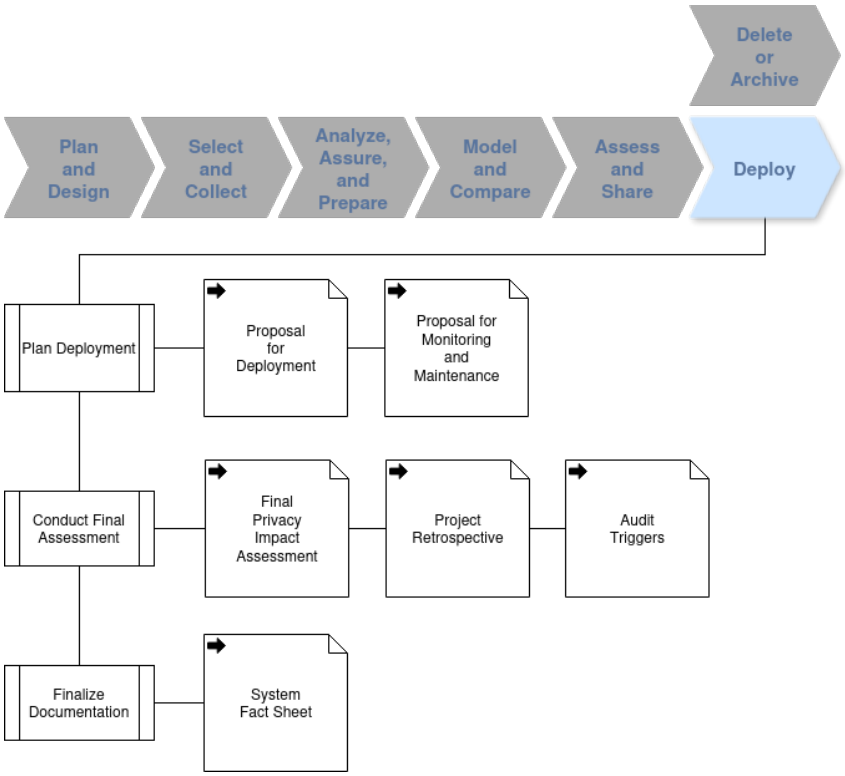

Figure 7: Deploy Tasks and Deliverables

5.7. Delete or Archive

This final stage handles the long-term management of data and models, including secure deletion or archiving practices. For auditing purposes, it ensures that lifecycle closure is verifiable and aligned with both regulatory obligations and user expectations around the "*right to be forgotten*."

Auditable Outputs: The tasks and outputs of this stage are shown in Figure 8. The archival proposal and details will provide sufficient information to verify the regulatory adherence of the archival process. Meanwhile, the deletion notification and report will provide the auditor with the necessary information to verify the regulatory and organisational compliance with the defined policies within the specified timelines.

Audit Focus: During an audit, evaluating if data deletion notifications were sent on time and if data was deleted or archived according to policy and regulatory timelines might be relevant. The access rights to archive files should be verified to ensure they are appropriately restricted.

5.8. Intermediate Risk Assessment

IRA checkpoints are embedded between stages to reassess system risks and compliance dynamically. UMAPER will analyse these documents closely with to verify that risk has been adequately reassessed as the project evolves, not just at its endpoints.

Auditable Outputs: The tasks and outputs of this stage are shown on Figure 9. The history file will provide auditors with a complete trail of the design choices made during development. The document requirement checklist can help guide them through the available documentation, as well as identify which aspects of the system might require retroactive documentation. Finally, the principles and violations checklist and the preliminary privacy impact assessment can clue the auditors into the risks the development team perceives.

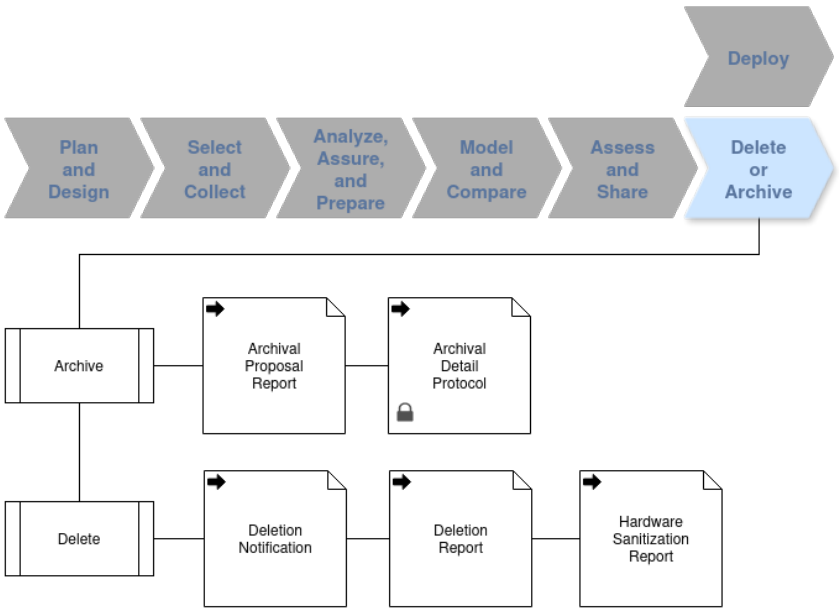

Figure 8: Delete or Archive Tasks and Deliverables

Audit Focus: The auditor might want to verify the adherence to the principles and violations checklist, and the documentation's completeness level.

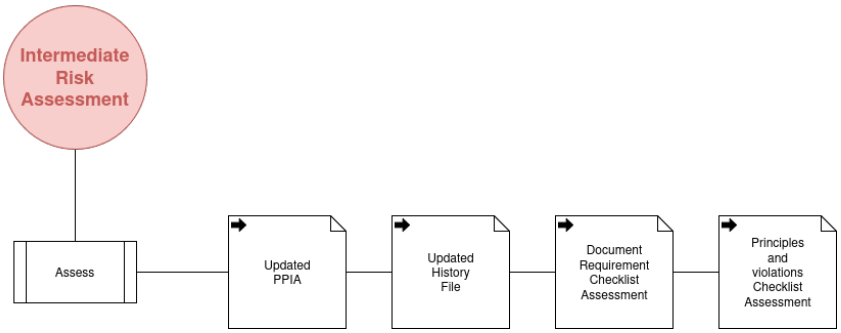

Figure 9: Intermediate Risk Assessment Tasks and Deliverables

6. UMAPER: An Auditing Framework for MDLC Projects

An audit is the activity of verifying processes and quality systems to ensure that "*the organisation's control processes are adequate to mitigate its risks, governance processes are effective and efficient, and organisational goals and objectives are met*" [23]. Auditing gathers evidence to confirm the proper functioning of established internal controls. It is a valuable tool for identifying errors, ensuring regulatory compliance, and highlighting areas with outdated controls.

Internal controls are policies, processes, tasks and behaviours that facilitate effective operations, they ensure quality reporting, and maintain compliance with applicable laws and regulations [15]. Throughout the stages of the MDLC, various internal control mechanisms were introduced to raise awareness of risks and challenging situations, promote and support informed decision-making, and minimise errors.

Healthcare systems are already subject to high levels of scrutiny and frequent audits due to the critical nature of their operations and the need to maintain public trust. Auditing in this domain enhances transparency and accountability, reduces errors, and strengthens compliance with regulations. However,

introducing AI-based systems introduces unique risks, particularly regarding privacy. As addressed in the justification for the MDLC, if AI-based systems ever become widespread in the medical field, they will likely need to be audited. We propose an auditing methodology based on the traditional auditing cycle and the SMACTR methodology [4] to conduct end-to-end audits for projects modelled after the MDLC.

A tailored auditing framework for the MDLC offers several key advantages. First, it provides a structured approach to verifying compliance with internal controls while assessing the performance and safety of AI-based medical systems. This framework focuses on both the process and product levels, ensuring that workflows adhere to the MDLC-defined standards and that the resulting systems meet quality benchmarks. Furthermore, adapting the existing auditing cycle to fit the MDLC model reduces the initial costs and resource demands of conducting audits, making it a more feasible option for organisations.

The benefits of such a framework extend beyond operational efficiency. A well-designed audit methodology promotes continuous improvement by creating a feedback loop between audit findings and system development. This iterative process improves the quality of the audited project, as it might kickstart a new iteration of the MDLC, and it can also improve the quality of future projects through the insights it produces.

Stakeholder trust is a crucial consideration in the adoption of AI-based medical systems. Auditing outcomes are often met with scepticism due to their reliance on human judgment, which can lead to the insights provided by the audit being disregarded. Proposing a specific framework with predefined stages and activities can generate a sense of legitimacy and credibility around the auditing process. In particular, the development of an auditing framework seeks to establish procedural justice in the auditing process, as establishing a framework can help auditors demonstrate adherence to accepted standards and the integrity of the audit.

The proposed auditing framework has six stages: Understand, Map, Assess, Plan, Execute, and Reflect. Each stage is based on the key auditing activities, defined by referencing the SMACTR framework, the standard auditing cycle, and the ISO/IEC 27001:2022 requirements for information security, cybersecurity and privacy protection. An overview of the audit framework is shown in Figure 10

6.1. Understand

This stage aims to evaluate the audit's relevance and specify its objectives by anticipating sources of risk and areas to investigate. To do this, the auditing provider (the entity that provides the auditing service, whether it is internal or external) will collaborate with the audit requester (individual, group, or organisation that requests the audit to be conducted) to gather information such as:

- Why was the audit requested? Is it a routine control, or did a specific concern trigger it?
- What is the relevance of this system in the context of the requester's operations?
- What are the main characteristics and functionalities of the system?
- Has the system already been audited in the past? If so, what were the recommendations made? Did management and the system team follow trough with them?
- Have any significant changes been made to the system between audits?
- What is the expected timeframe for the audit?
- What is the expected scope of the audit?
- Who is the system owner and other key collaborators the auditing team can rely on to gather information about the system?

The auditor can complement this information with the documentation produced by the MDLC at different stages:

- From the Plan and Design stage: The project background document, the requirement specification document, and the principles and violations checklist.

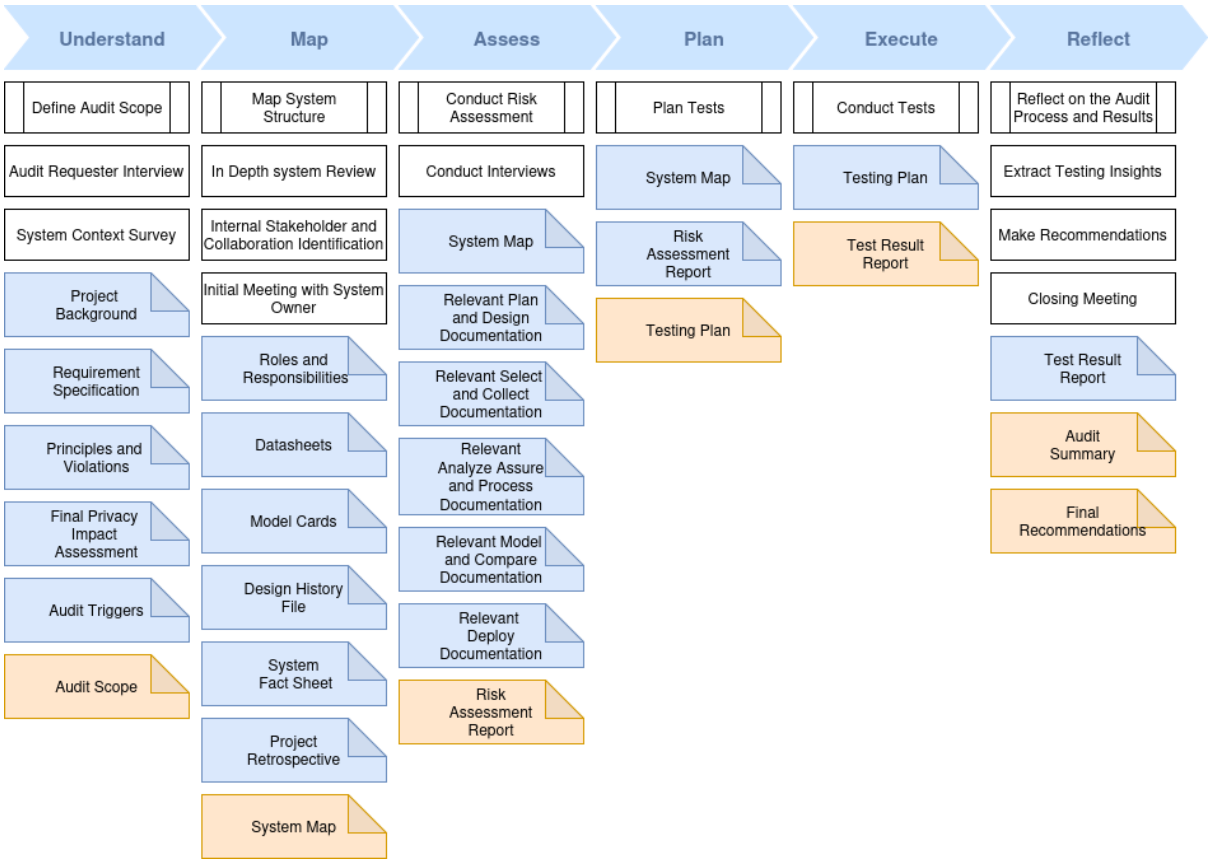

Figure 10: Overview of the UMAPER Framework. Double-lined boxes represent processes, regular boxes represent tasks, blue represents input documents, and orange represents output documents.

- From the Deploy stage: The final privacy impact assessment and the audit triggers.

In short, this stage's main tasks are gathering information on the audit's motivation and the system's organisational, ethical, and regulatory context. The reason for carrying out this preliminary investigation is twofold. First, it will help the auditing team gather initial data to assess the feasibility of the audit. In second place, it will help the auditing team identify the areas on which the audit will focus or the audit scope. The audit scope could potentially be a privacy risk assessment, but this framework is flexible enough to provide a base for other types of audits. The findings from this stage should be reflected in a report to be presented to the requester once it is finished and to the system owner (the individual responsible for the system) before the initial meeting.

The Audit Scope Report: This report should include notes on the requester interview, the reasons for the conduction of the audit, the areas of interest for the audit, the normative references for the audit, the proposed auditing team, and the immediate activities, which will include the notification of the audit to the system owner and the setting of a date for an initial meeting with the system owner so that they may prepare.

6.2. Map

This stage's goal is for the auditing team to learn as much as possible from the system's documentation. This stage occurs before the first meeting with the system owner, and its goal is to minimise the time they are imposed on by becoming as familiar with the system as independently as possible. The knowledge acquired during this stage will later need to be verified with the system owner team. Although the documentation strives to capture the system's reality, it is prone to errors and inaccuracies, especially if the system has had any significant changes recently.

In this stage, the auditing team will review the systems documentation to familiarise itself with the development process and identify internal stakeholders and collaborators. The documentation review aims to identify opportunities for testing by reviewing the development process and assessing the completeness of the documentation by interacting with it. Meanwhile, the internal stakeholder and collaborator identifications help establish relevant contacts that might be necessary during the following stages and identify individual participation in the final outcome, allowing auditors to assess personal accountability for each stage of the development process.

To carry out the tasks of this stage, the auditing team may rely on documents from the MDLC, such as:

- From the Plan and Design stage: The outline of roles and responsibilities.
- From the Analyse, Assure, and Process stage: The final datasheet.
- From the Model and Compare stage: The model cards and the design history file.
- From the Deploy stage: The system fact sheet and project retrospective.

Once these tasks have been completed, the auditing team may move on to the initial meeting with the system owner. During this meeting, the auditing team will share and confirm the information they have acquired about the system, and the system owner will provide any additional resources. After this meeting, the auditing team will create a final system map with the most up-to-date information on the system. It may then move on to the following stages, in which they will begin preparing questionnaires and tests to conduct the audit. It is important to note that it is common practice to have an opening meeting with the system owner before going into detailed planning, as this sets up the stage for a better-informed and prepared audit.

The System Map: This document provides a checklist of the expected documentation, marking it as available or unavailable and pointing out if any errors or inaccuracies came to light during the initial meeting. The system map also provides a summary of the system's structure, a description of each relevant system unit and the relevant stakeholders and contacts.

6.3. Assess

The risk assessment stage is a necessary precursor to the planning stage. The potential risks incurred by the system are necessary to inform test selection, as they are the key to defining relevant tests. This is essential to guarantee that the audit is efficient and effective, focusing on areas of greatest risk.

Although the MDLC innately considers privacy impact assessment, auditors' expertise can greatly enhance the review and complement of this document. Additionally, as an audit's scope is not restricted to privacy, it might be relevant to conduct a human rights impact assessment, system bias risk assessment, or failure modes and effects, among others.

This stage is informed by the system map; however, this information can and should be supplemented by consulting others. Relevant contacts found during the previous stage, domain experts (medical professionals or hospital administrative personnel in this context), external experts, or legal advisors should be consulted to have wider coverage of the potential risks inherent to the system being audited and systems similar to it.

Additionally, documentation produced during the MDLC that is not considered during the system mapping can provide insight into potential risks and mitigatory measures already in place. These documents will be particularly relevant to a privacy audit. Still, they can also be relevant towards informing audits with different objectives, as they cover various relevant aspects regarding the project's development. Relevant documents include:

- From the Plan and Design stage: the data access plan and the principles and violations checklist.
- From the Select and Collect stage: The data selection report, the data origin report and the data deidentification report.
- From the Analyse, Assure, and Propose stage: The data quality report.

- From the Model and Compare stage: The model test and insight report.
- From the Assess and Share stage: The list of data access plan deviations and the self assessment checklist.
- From the Deploy stage: The proposal for monitoring and maintenance, the final privacy impact assessment, the project retrospective and the audit triggers.

The auditing team's work should be registered in a report. This document will be shared with the system owner and the audit requester so that they may review it and make necessary recommendations before proceeding to the next stage.

The Risk Assessment Report: Lists the risks found, their likelihood and severity, and possible mitigatory measures for them. This document will also disclose any communication with relevant contacts that led to discovering these potential risks or their justification based on the system map.

6.4. Plan

During this stage, auditors will use all the information they have gathered regarding the system, its context, and the risks it entails to produce a concrete set of tests to be executed. Tests are one of the main tools auditors use to assess a system's compliance with internal company policies, legal regulations, and even ethical values, as proposed in the SMACTR Framework. These tests should ideally help assess the system's most concerning or most likely risks, which is why the planning is conducted once the auditors thoroughly understand the system.

If this is a routine control and no major changes to the system have occurred between this audit and the last, a list of relevant tests might already exists. If this is the case, auditors should take it as a base, update the timeframe, review the tests to ensure they are all still relevant, and either remove or update them as necessary.

The resulting plan will be delivered to the audit requester and the system owner to inform them of the proposed tests. The auditing team can consider any questions or clarifications the audit requires or the system owner wishes to make. With this information, the system owner can also begin to prepare for the upcoming tests by designing which human, technical, and other resources will be assigned to aid during test execution.

Test selection and satisfaction criteria depend heavily on the system's specific purpose, implementation details, and context. This is why planning will have to be conducted for each system, and results may vary widely. However, auditors may reference testing plans for similar systems during this process. For instance, for a medical image classifier system, auditors may want to take inspiration from clinical audit guides to assess the medical aspect of the system, audit guides for classification systems to assess the algorithmic aspects of the system, GDPR compliance guides to assess the regulatory aspects of the system, and so on.

The Testing Plan: This document will include the tests to be carried out during the execution stage, their justification, the risk they target, the required collaborators to carry them out, their expected start and end date, the member of the auditing team responsible for them, and any additional considerations that are deemed relevant by the auditing team. If this is an update to a previously existing plan, it should highlight the added, removed or edited tests. Each test will also be accompanied by a justified grading rubric detailing the requirements for a satisfactory test result. In addition to the individual test design, the testing plan will include how system satisfaction will be calculated based on the individual test results.

6.5. Execute

During this stage, the audit team and the system team will carry out the established tests in accordance with the test plan. This is where most of the audit time is likely spent, and the audit team will need to

work most closely with the system team. Through this stage, auditors will engage with the system to carry out each test defined in the plan, grade it according to the rubric proposed in the test plan, and point out any especially relevant results.

The test results will be provided to the system owner and the audit requester, who may request a re-test if the results seem inconsistent or unexpected. These re-tests need to be highlighted in the test result report, as they can shed light into aspects of the system that are of special interest for future audits.

The Test Result Report: This document will list the tests conducted; if it was not possible to execute all the tests proposed in the test plan for any reason, it must also be justified in this document. For each test conducted, the auditors responsible will provide the timeframe for execution, a result according to the rubric, and any notes made by themselves or the system team members they collaborated with.

6.6. Reflect

In this stage, test results will be analysed and the auditing process will be reviewed. The audit team and the system team will reflect on the way in which the audit was conducted, the sufficiency of the tests, any insights that can be extracted from the test results, and the system's overall performance.

In this stage, mitigation plans for risks found too high through testing can be proposed. Remedial actions can also be suggested to handle any materialised risks that are discovered during the audit. Finally, based on the audit findings, the risk assessment proposed during the Assess stage can be updated to reflect the system's current state better.

The success of this stage depends on the collaboration between the auditors and the system team. The system team will provide valuable insight into the action plan in light of the audit results. In contrast, the audit team recommends increasing the system's compliance in aspects such as regulation or internal policy. Stakeholder involvement is also crucial in this stage, as their commitment and support are essential to ensure that recommendations are effectively implemented.

The results of this stage will be presented to the system owner and the audit requester at the closing meeting. During this meeting, the auditor may answer questions and receive feedback on the auditing process.

The Audit Summary: Review the audit process, summarise test results and system evaluation, and provide final considerations regarding the extent of test coverage and any unexpected situations or challenges encountered during the audit.

Recommendations: This document formalizes the recommendations made by the auditing team, provides a timeframe for their implementation and proposes a follow-up audit after that timeframe.

The regulatory landscape for the development of AI solutions in critical fields such as healthcare is extensive and dense. It is important to develop frameworks, such as UMAPER, that allow stakeholders and developers to direct their efforts toward compliance and verify the compliance of existing systems. The structure and expectations set by such a framework can enable AI in healthcare to grow in a more orderly manner and reach further.

By virtue of producing an extensive accountability chain, UMAPER improves compliance with some of the legal requirements for such systems. UMAPER can be considered a continuous, iterative process that is planned and executed to oversee the entire life cycle of an AI system, as required by the AI Act in its Article 9 on risk management.

The use of UMAPER will also inherently aid in complying with documentation requirements set out in Recitals 12 and 82 and Articles 30 and 33 of the GDPR, as well as Articles 11, 17, and 18 of the AI Act. Additionally, auditing will enhance process transparency, facilitating adherence to the requirements for high-risk AI systems defined in Article 13 of the AI Act. Finally, aspects regarding the identification, supervision, and mitigation of risks will also be improved by adopting auditing practices.

Auditing can broadly be considered a form of human oversight, as it ensures that a natural person will periodically oversee the system, explicitly looking for potential risks or existing errors. This minimizes future errors and allows for earlier detection of materialized risks. Such human involvement is a requirement for high-risk AI systems according to Article 14 of the AI Act.

Finally, auditing presents an opportunity to review a variety of relevant issues. It is not beyond the scope of imagination to propose tests to verify the existence of informed consent forms for all data subjects involved (Article 9 of the GDPR, among others), the existence and effectiveness of anonymization and pseudonymization processes (Article 25 of the GDPR, among others), and the collection of post-market performance data to review risk analysis and documentation (Recital 147 of the AI Act; Recital 74 of Regulation 2017/745 of the European Parliament and the Council on medical devices [24]).

7. Conclusions

In healthcare, where decisions made by AI systems can directly impact patient outcomes and institutional credibility, auditing becomes an essential tool to align the level of control within medical AI-based systems with the high standards of the healthcare field. In this paper, we propose UMAPER, a structured auditing framework designed to support primarily internal audits of systems developed under the Medical Data Lifecycle (MDLC). While MDLC focuses on embedding privacy and transparency throughout the development process, UMAPER builds on this foundation by offering a practical pathway for verification and accountability.

The UMAPER framework incorporates key principles from established auditing methodologies, including the standard internal audit cycle, the SMACTR framework, and regulatory considerations, while tailoring them to the specific context of medical AI-based systems. By aligning audit tasks with the lifecycle stages of MDLC, it becomes possible to evaluate whether systems uphold regulatory standards and organisational policies before, during, and after development. By enabling ongoing internal audits throughout the system's lifecycle, not just retrospective external reviews, UMAPER allows for proactive risk mitigation.

A clear auditing structure, such as the one proposed for UMAPER, also supports procedural justice, an essential component in building stakeholder trust. Structured audits with predefined stages, deliverables, and evaluation criteria can lend legitimacy to the process, reduce subjectivity in audit outcomes, and help ensure that findings are taken seriously by system owners, increasing the likelihood of observations and recommendations being implemented instead of being disregarded as unfounded or unfair.

Additionally, it can facilitate adoption of mature and structured control and assessment processes, making auditing more accessible across the board. Which, over time and on average, might help mitigate the risks associated with this type of system as they have for other high-security fields [4].

In the context of responsible AI, the MDLC and UMAPER frameworks emphasise the ethical need to ensure that artificial intelligence systems are developed and implemented in ways that prioritise data privacy, accountability, and transparency. Responsible AI is about minimising harm and proactively fostering trust and equity in systems that influence critical areas such as healthcare. By embedding principles of fairness, accountability, and compliance within MDLC and UMAPER, these frameworks align AI development with both ethical guidelines and regulatory standards, ensuring that AI serves society.

Limitations: UMAPER is a conceptual framework that does not have empirical validation as of now, however, it is in alignment with widely used guidelines and we intend to pursue a case study eventually. Additionally, the UMAPER framework assumes a fairly structured organisational environment where it is reasonable to assume resources will be available for internal auditing. Finally, UMAPER hinges on EU-specific legislation, which may not fully generalize to different legal frameworks.

Acknowledgments

This work is based on the first author's master's thesis.

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 871042 (SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics).

Declaration on Generative AI

While preparing this work, the authors punctually used Grammarly to check grammar and spelling. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

[1] I. Piedrahita-Velez, CRITICAL PRIVACY ISSUES ON MEDICAL DATA, Master's thesis, Universitat Politècnica de Catalunya, 2025. URL: https://drive.google.com/file/d/1XqLeCSUZ_kn1zOeY6QNnfBR9vAOC1qEE/view?usp=sharing.

[2] European Parliament, Council of the European Union, Regulation (EU) 2016/679 of the European Parliament and of the Council, 2016. URL: https://data.europa.eu/eli/reg/2016/679/oj.

[3] European Parliament, Council of the European Union, Regulation (EU) 2023/1234 of the European Parliament and of the Council, 2023. URL: https://artificialintelligenceact.eu/the-act/.

[4] Closing the ai accountability gap: defining an end-to-end framework for internal algorithmic auditin, 2020. URL: https://dl.acm.org/doi/abs/10.1145/3351095.3372873.

[5] T. Gebru, J. Morgenstern, B. Vecchione, J. W. Vaughan, H. Wallach, H. D. III, K. Crawford, Datasheets for datasets, 2021. URL: https://arxiv.org/abs/1803.09010. arXiv:1803.09010.

[6] M. Mitchell, S. Wu, A. Zaldivar, P. Barnes, L. Vasserman, B. Hutchinson, E. Spitzer, I. D. Raji, T. Gebru, Model cards for model reporting, in: Proceedings of the Conference on Fairness, Accountability, and Transparency, FAT* '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 220–229. URL: https://doi.org/10.1145/3287560.3287596. doi:10.1145/3287560.3287596.

[7] J. T. Richards, D. Piorkowski, M. Hind, S. Houde, A. Mojsilovic, K. R. Varshney, A human-centered methodology for creating ai factsheets, IEEE Data Eng. Bull. 44 (2021) 47–58. URL: https://api.semanticscholar.org/CorpusID:246482320.

[8] G. S. Collins, K. G. M. Moons, P. Dhiman, R. D. Riley, A. L. Beam, B. Van Calster, M. Ghassemi, X. Liu, J. B. Reitsma, M. van Smeden, A.-L. Boulesteix, J. C. Camaradou, L. A. Celi, S. Denaxas, A. K. Denniston, B. Glocker, R. M. Golub, H. Harvey, G. Heinze, M. M. Hoffman, A. P. Kengne, E. Lam, N. Lee, E. W. Loder, L. Maier-Hein, B. A. Mateen, M. D. McCradden, L. Oakden-Rayner, J. Ordish, R. Parnell, S. Rose, K. Singh, L. Wynants, P. Logullo, Tripod+ai statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods, BMJ 385 (2024). URL: https://www.bmj.com/content/385/bmj-2023-078378. doi:10.1136/bmj-2023-078378. arXiv:https://www.bmj.com/content/385/bmj-2023-078378.full.pdf.

[9] T. Stobierski, The data life cycle: Processing, 2022. URL: https://online.hbs.edu/blog/post/data-life-cycle#processing.

[10] P. e. a. Chapman, Crisp-dm1.0 step-by-step data mining guide, 2009. Tech. Rep.

[11] C. Strasser, R. Cook, W. Michener, A. Budden, Dataone best practices primer: primer on data management: what you always wanted to know, 2012. 11pp. DOI: http://dx.doi.org/10.25607/OBP-84.

[12] H. M. School, The biomedical data lifecycle, n.d. URL: https://datamanagement.hms.harvard.edu/plan-design/biomedical-data-lifecycle#:~:text=What%20is%20the%20Data%20Lifecycle,collection%20use%2C%20and%20storage.

[13] E. D. P. Board, Ai auditing, 2024. URL: https://www.edpb.europa.eu/our-work-tools/our-documents/support-pool-experts-projects/ai-auditing_en.
[14] ACCA, A brief guide to internal auditing, n.d. URL: https://www.accaglobal.com/gb/en/member/sectors/internal-audit/learn/brief-guide.html, accessed: December 31, 2024.
[15] T. R. Group, Internal control: Guidance for directors on the combined code, 1999. Also known as the Turnbull Report.
[16] U. L. Anderson, M. J. Head, S. Ramamoorti, C. Riddle, M. Salamasick, P. J. Sobel, Information Technology Auditing, The Internal Audit Foundation, 2017.
[17] T. I. of Internal Auditors (The IIA), The international professional practices framework (ippf), 2019. Accessed from www.globaliia.org, pp. 96-99.
[18] T. I. of Internal Auditors (The IIA), The international professional practices framework (ippf), 2019. Accessed from www.globaliia.org, pp. 154-157.
[19] U. L. Anderson, M. J. Head, S. Ramamoorti, C. Riddle, M. Salamasick, P. J. Sobel, Information Technology Auditing, The Internal Audit Foundation, 2017.
[20] T. I. of Internal Auditors (The IIA), The international professional practices framework (ippf), 2019. Accessed from www.globaliia.org, pp. 173-181.
[21] U. L. Anderson, M. J. Head, S. Ramamoorti, C. Riddle, M. Salamasick, P. J. Sobel, Information Technology Auditing, The Internal Audit Foundation, 2017.
[22] A. S. D. P. Agency), Audit requirements for personal data processing activities involving ai, 2021. Accessed from https://www.aepd.es/guides/audit-requirements-for-personal-data-processing-activities-involving-ai.pdf.
[23] The Institute of Internal Auditors (IIA), About internal audit, 2025. URL: https://www.theiia.org/en/about-us/about-internal-audit/, accessed: 2025-01-15.
[24] European Parliament, Council of the European Union, Regulation (EU) 2023/1234 of the European Parliament and of the Council, 2016. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.