Subdomain 4.1: Responsible AI Lifecycle Management and Use Playbook

Guiding responsible AI adoption

Before You Start

For more about the CHAI AI Governance Framework, methods used to develop these playbooks, and information on how to use them, please refer to the CHAI AI Governance Playbooks: Introduction & Background document. Note that **baseline operational controls** were consensus-defined by Healthcare Delivery Organizations (HDOs) as "must-haves" for health AI governance. The **implementation guidance** that accompanies the baseline operational controls are **suggestions/recommendations** based on known practices. It is assumed and appropriate that implementation will differ based on organizational goals, resources, and maturity.

Disclaimer

The implementation guidance, frameworks, and examples contained herein represent generalized practices and considerations drawn from across the field of responsible AI deployment in healthcare. They are offered as one set of perspectives to inform your thinking not as the "right way" or "only way" to build, govern, or operate an AI program.

Every organization operates within its own unique context: differing patient populations, clinical priorities, regulatory environments, resource constraints, technical infrastructure, cultural norms, and stage of AI maturity. While the specific methods, timelines, and depth of implementation will appropriately vary across organizations, the underlying controls, principles, risks, and considerations raised throughout this playbook are intended to be engaged with thoughtfully and in good faith by any organization deploying AI in healthcare.

Organizations should prioritize state and federal regulatory requirements, especially for AI solutions that meet the definition of a Software as a Medical Device (SaMD), as defined by the FDA.

Not a Standard of Care. This playbook does not establish, define, or reflect a legal, regulatory, professional, or industry standard of care, nor is it intended to serve as one. It is a voluntary, aspirational, and educational resource designed to support organizational learning and dialogue. The field of healthcare AI is rapidly evolving, and reasonable organizations and professionals may differ on how to approach the issues discussed herein. The practices described may not be appropriate, feasible, or advisable for every organization, every use case, or every point in time. Adoption of any particular practice described in this playbook is entirely voluntary, and a decision to adopt, adapt, partially implement, or decline to implement any practice should not be construed as evidence of compliance with, or deviation from, any standard of care, duty, or legal obligation. This playbook is not intended to be used, and should not be used, as a benchmark, checklist, or measuring stick against which the conduct of any

organization, clinician, or other party is evaluated in any litigation, regulatory proceeding, accreditation review, or similar context.

General Informational Purpose. The information in this playbook is provided for general informational and internal operational purposes only. It is intended to support organizations in thinking through the design and stewardship of their AI programs and should be used in conjunction with from clinical leaders, technical experts, compliance officers, ethicists, patient and community representatives, and other advisors familiar with the specifics of your organization.

Not Legal Advice. Nothing in this playbook constitutes legal advice, and it should not be relied upon as a substitute for advice from a qualified attorney licensed to practice in the relevant jurisdiction. Nothing in this document creates an attorney-client relationship between the reader and CHAI, its officers, directors, employees, or contributors. The contents reflect general information and observed practices as of the date of publication and may not account for changes in law, regulation, technology, or circumstance that occur thereafter.

No Warranty. While reasonable efforts have been made to ensure accuracy, no representation or warranty is made as to the completeness, timeliness, accuracy, or suitability of the information for any particular purpose. Laws, regulations, and best practices vary by jurisdiction and evolve over time, and the application of those laws and practices to specific facts can lead to different outcomes.

Illustrative Examples. Any tools, resources, examples, or technology product described or presented are for educational and illustrative purposes only and do not represent endorsement by CHAI or guarantee specific results. Readers should consult qualified legal counsel and relevant experts for specific terms, products, or services.

Independent Judgment Required. Readers should consult with qualified legal counsel, clinical experts, and other appropriate professional advisors before taking, or refraining from taking, any action based on the information in this playbook. The decision to adopt, modify, or disregard any practice described herein rests solely with the reader and their organization. CHAI expressly disclaims any and all liability for any loss, damage, or other consequence, including but not limited to claims of negligence, breach of duty, or failure to meet any standard, arising directly or indirectly from reliance on, use of, citation to, or reference to the contents of this document.

At a Glance

The **Subdomain 4.1: Responsible AI Lifecycle Management and Use** Playbook is designed to guide healthcare delivery organizations (HDOs) through the steps of implementing processes that support responsible lifecycle management and use of AI solutions while identifying and documenting broader objectives.

Background Context for Subdomain 4.1: Responsible AI Lifecycle Management and Use

Before detailing the baseline operational controls for Subdomain 4.1: Responsible AI Lifecycle Management and Use of health AI solutions, organizations should first understand how AI solutions enter the organization, referred

to as the ingestion modality. The origin and pathway through which an AI solution is introduced directly shape the appropriate lifecycle management considerations. There are four primary ingestion modalities:

1. **Research Models/Solutions**
    a. Description: AI models/solutions developed within an HDO primarily for research, experimentation, or hypothesis testing, often prior to operational deployment. These models/solutions are typically used in controlled settings or grant-funded environments and may not yet be integrated into clinical workflows. The Institutional Review Board (IRB) plays a critical role in assuring safety and ethical conduct for research AI models/solutions. Organizations where IRB representatives sit on the AI Governance Committee (AIGC) benefit from enhanced safety oversight and standardization.
    b. Challenge: Governance is highly localized and not easily generalizable across organizations.
    c. Note: Research models/solutions typically require IRB assessment and approval prior to development or deployment. HDOs should also define an explicit onramp path for transitioning a successful research model/solution into operational use.
2. **In-house Developed Models/Solutions**
    a. Description: AI models/solutions designed, built, and maintained by an HDO's internal teams (e.g., data science, informatics, IT) for use within its own operations. The organization retains primary control over development, validation, and lifecycle management.
    b. Challenge: Processes are highly variable and resource-dependent, limiting cross-institutional standardization.
    c. Note: IRB approval or waiver may be required for in-house developed models/solutions if there is an intent to publish findings.
3. **Vendor or Co-developed Models/Solutions**
    a. Description: AI models/solutions that are commercially licensed from external vendors or jointly developed with third-party partners. These include both fit-for-purpose or general-use solutions. Development responsibilities are shared or externalized, while the HDO remains accountable for safe implementation and oversight within its environment. Partner universities serve as a key external collaborator for many healthcare organizations that are not university-operated Academic Medical Centers; the role and responsibilities of academic partners should be explicitly defined in co-development agreements.
    b. Challenge: Ambiguous accountability, contractual constraints that restrict access to performance metrics or model change logs needed for ongoing management, limited visibility into model/solution development and training data, and/or dependency on vendor release cycles for remediation or retraining, intellectual property ownership ambiguities (e.g., who owns improvements to the vendor model/solution), and insufficient contractual obligations for vendors to provide ongoing performance metrics and data.
4. **AI-Enhanced Features in Existing Health IT Systems**
    a. Description: These are AI functionalities integrated into pre-existing health IT systems, such as Electronic Health Records (EHRs), imaging platforms, laboratory systems, or revenue cycle solutions, as modular add-ons or embedded features. The core system remains the same, but new AI capabilities are layered on to enhance specific tasks like documentation, decision support, image analysis, or claims processing.
    b. Challenge: Automatic activation of AI features through system upgrades or configuration changes without explicit organizational review or approval, lack of required vendor disclosure of model documentation, absence of a formal review or approval process when AI capabilities are enabled, modified, or expanded during software updates, low user awareness that AI is influencing outputs or recommendations, increasing the risk of over-reliance or automation bias, and/or diffuse ownership between IT, clinical operations, and compliance teams, and economic incentives that may encourage enabling pre-purchased AI features without independent governance review (e.g., bundled EHR contracts where AI solutions are already included in existing licensing fees). Lastly, some vendors will not voluntarily commit to advance notification of model/solution updates or feature changes: in those cases,

HDOs should consider adding periodic feature-inventory audits and/or contractual notification requirements as compensating controls.

c. Note: HDOs should strive for vendors to provide advance notification of new AI functionality or upgrades. This allows new models/solutions to be re-assessed through the organizational intake and governance process. Vendors may also expose new functionality within existing terms and conditions, or rely on API calls to third-party models whose terms differ from the platform contract; understanding how product architecture maps to applicable contract terms is critical.

Note: Organizational v. Personal Use of AI

AI solutions may enter an organization for *organizational use* (deployed through formal intake, contracting, and governance) and for *personal use* (employees or clinicians accessing consumer-grade solutions such as ChatGPT, Microsoft Copilot, or mobile/"shadow" AI applications on personal or work devices). Shadow AI refers to AI used by staff outside the HDO's formal intake and approval processes. Some organizations permit limited personal use under acceptable-use policies; others restrict use to a curated set of organization-approved LLM-backed solutions. Either way, organizations should explicitly state which solutions are permitted, the conditions of use (including protected health information (PHI) handling), and the channels for requesting new solutions. Mobile and shadow AI should be addressed in policy, training, and ongoing monitoring.

HDOs should refer to their AI Policy document (see Domain 1: AI Policy Playbook) regarding which types of ingestion modalities are in-scope and out-of-scope. It is recommended that each ingestion modality be paired with a formally designated accountable owner(s) (e.g., an executive sponsor and an operational owner) responsible for lifecycle compliance. Ambiguity in accountability is a leading governance failure point, particularly for vendor AI and embedded EHR features.

The ingestion modalities in which AI solutions enter the organization need to be governed and managed across the organization's defined AI lifecycle. The components of an AI lifecycle to consider at your organization are described below.

Responsible AI (RAI) Lifecycle Management Steps

The Responsible AI Lifecycle steps described below represent components of widely adopted lifecycle frameworks (e.g., CHAI's Responsible AI Guide Executive Summary). Organizations may adapt step definitions to reflect their unique context, while staying in alignment with emerging best practices. Ensure a risk-proportionate approach across each lifecycle step (e.g., lightweight path for low-risk use cases).

Define the Problem & Plan

Before HDOs can work through subsequent lifecycle management steps, it is critical to establish a strategy to help drive use case prioritization, business case development, and governing body pre-approval. For specific use cases, approaching selection by starting with the problem-statement (what problem is AI intended to solve), will drive the intended purpose of the solution. The intended purpose a) defines the scope of verification & validation activities and b) allows reasonable delineation between user responsibility and vendor responsibility. HDOs should identify a clear problem, its setting, and stakeholders involved (which together comprise a *use case*), thereby quantifying potential return on investment, return on health outcomes, and improvement of healthcare operations. Implementers use this information to decide whether to build an in-house AI solution, procure one from a third party, or partner with a third party to develop a solution. The following lifecycle management steps are important independent of ingestion modality.

Intake/Procurement

Not all AI solutions enter an organization through a formal procurement process; however, the initial step in managing any AI solution is a structured intake process. Intake consists of gathering and documenting essential information about the AI solution and its underlying model(s) to establish a baseline understanding of intended use, risk, and governance needs. Intake should explicitly capture intended *users* in addition to the intended *use*. CHAI recommends collecting this information through a model card (see CHAI's Applied Model Card, hereafter "Applied Model Card AMC)") or comparable documentation and, where applicable, a standardized intake questionnaire (see CHAI's AI Intake Questionnaire, under the "Tools & Resources" section in Subdomain 4.4: Third Party Management Playbook).

A model card describes an AI solution in the context of a specific health use case and reflects the full operational system, including the model, technical infrastructure, and human workflows. Key components include intended use and workflow, primary users, targeted patient population, out-of-scope use cases, known risks and limitations, known biases or ethical considerations, AI system facts (e.g., model type, data sources, inputs/outputs, foundation models used, development data characterization), and key evaluation metrics related to usefulness, usability, efficacy, fairness, safety, reliability, transparency, privacy, and security (see CHAI's AMC example).

Intake questionnaires should supplement model documentation by capturing additional organizational, operational, and contextual information not otherwise documented. These questionnaires are often organized into domains such as data management, cybersecurity, ethical considerations, deployment and monitoring planning, legal and regulatory compliance, stakeholder engagement, and training and awareness. Sample intake questions include:

- Data: e.g., Are there plans to obtain, store, engineer and protect data required for the AI?
- Cybersecurity: e.g., Does the AI solution introduce new attack surfaces or require security review? Has an IT security assessment been completed?
- Ethical Considerations: e.g., Are there any ethical concerns or potential societal impacts associated with the AI project?
- Bias Considerations: e.g., Are there any bias concerns or potential population-specific (e.g., race, ethnicity, payer, socioeconomic status) impacts associated with the AI solution? How will these be identified and addressed?
- Deployment and Monitoring: e.g., How will features be mapped to production data from the EHR?; How do you identify instances where immediate actions or taking the AI solution offline are needed after deployment?
- Training and awareness: e.g., What training or educational resources are available to stakeholders involved in the AI project?

Note

HDOs typically have a two-part intake process with external vendors.

Part 1 – initial, light-touch set of intake questions to categorize risk (low risk, medium risk, high risk) for triage purposes. Initial light touch questions to understand:

- Does this solution solve an important problem?
- How does the solution solve the problem/what impact does it aim to have?
- What type of AI is this using (LLM, regression model)?
- Does/how the AI impact patients (decision support or autonomous)?
- Does it touch PHI?
- What are the high level data dependencies underlying the fidelity of the model output?
- In what settings has the model been developed and applied? And how extensively?

Part 2 – additional, more robust set of intake questions for medium risk and high risk solutions. Minimum vendor governance requirements should be defined by the organization and added to the AI Intake questionnaire.

- Part 2 is a detailed intake and governance approval, and it is typically completed in collaboration with both the vendor and business/clinical stakeholders from the HDO.

It is recommended that all AI systems, regardless of ingestion modality, be logged in a centralized AI inventory that captures: intended use, risk tier, monitoring owner, version tracking, and other relevant components. AI inventories are becoming a baseline governance expectation aligned with U.S. Food and Drug Administration (FDA), Office of the National Coordinator for Health Information Technology (ONC) HTI-1, and ISO/IEC 42001:2023 requirements. Note that existing Information Technology (IT) or Governance, Risk, and Compliance (GRC) systems can serve as the AI inventory if they capture important fields.

Risk

Once intake information has been gathered, organizations should categorize the AI solution's risk prior to deployment. Risk categorization synthesizes information from the model card and intake questionnaires to assess risk domains. The resulting risk categorization informs which mitigation controls are required, the depth of evaluation and monitoring, and expectations across subsequent lifecycle steps. Additional guidance on risk categorization is detailed in Subdomain 4.2: Risk and Impact Assessments Playbook and made practical through CHAI's Risk Categorization Tool (v2).

Pre-deployment

Pre-deployment occurs after design, development, and assessment activities and before initiating a pilot. The goal of this stage is to assess readiness for real-world use at limited scale and to proactively address risks that may emerge once the AI solution interacts with production data, users, and workflows. Updates to clinical, population, and organizational risk should be incorporated based on insights from development, validation, and silent evaluation. Evidence standards for validation should be risk-tiered, including: (1) what constitutes sufficient validation evidence at each tier; (2) when local validation is mandatory versus accepted by vendor attestation; and (3) the role of the governance body in defining acceptable and unacceptable uses. This lifecycle step applies to both vendors and HDOs. It includes readiness assessments for the purchasing or deploying organization, clarifies roles and responsibilities between vendors and HDOs (where applicable), and evaluates broader AI system design considerations such as safety, privacy, security, usability, and monitoring plans. For acquired or vendor-provided solutions, organizations may use this stage as part of procurement by requiring evidence of best

practices and engaging jointly with developers to confirm appropriate planning, risk mitigation, and usability for the full AI system.

Note

In practice at most HDOs, third party vendor contracts are signed before pre-deployment assessment is complete, often because procurement timelines and governance timelines don't align, or because executives approve a vendor relationship before the governance team has completed review.

Where organizational timelines require contract execution before pre-deployment assessment is complete, contracts should include a conditional right-to-terminate clause: the organization retains the right to exit the agreement without penalty if pre-deployment assessment reveals unacceptable risk, if the vendor fails to provide required documentation within a defined period, or if governance committee review results in a no-proceed decision.

Piloting

Piloting is a controlled, limited-scale deployment designed to validate real-world performance, usability, workflow integration, and risk assumptions prior to full rollout. During this stage, organizations should confirm that the AI solution is used as intended, evaluate early safety and performance signals, and assess human-AI interaction, including user reliance, overrides, and escalation behaviors. Pilot results should be reviewed against predefined success criteria, risk thresholds, and Responsible AI principles, and findings should inform decisions to proceed, iterate, expand, pause, or terminate deployment. HDOs should ensure that pilot data environments are segregated to prevent pilot data from contaminating production training pipelines.

Additionally, pilot exit criteria should be predefined and include: measurable Responsible AI thresholds (e.g., fairness, safety, accuracy benchmarks), explicit go/no-go authority designations, and formal sign-off from the designated AI oversight body before proceeding to full deployment. Piloting is strongly encouraged in most cases and should be considered a requirement for high-risk AI solutions prior to full deployment.

Deployment

Deployment represents the transition from pilot to broader operational use. At this stage, the AI solution is integrated into standard workflows, supported by finalized policies, training, and operational ownership. Deployment should be accompanied by activation of monitoring mechanisms, clear accountability for oversight and incident response, and communication to affected stakeholders regarding the AI's role, intended use, and limitations. Any material changes to scope, configuration, or intended use identified during deployment should trigger reassessment through intake or risk categorization processes as appropriate. Incident response recommendations are included in Subdomain 4.5 Education, Training, and Feedback Playbook.

Monitoring

Monitoring is an ongoing lifecycle activity focused on evaluating safety performance, effectiveness, and responsible use over time. Organizations should regularly review defined metrics through dashboards or comparable reporting mechanisms, track usage patterns, and flag unexpected behavior or deviations from intended use for manual review by IT, compliance, and other oversight functions. Monitoring should account for model drift, workflow changes, user behavior, and downstream impacts, particularly where direct model telemetry is limited. Additionally, monitoring activities may include tracking performance over time, identifying potential bias or unintended consequences, and assessing continued alignment with clinical or operational workflows.

Organizations may also define conditions that trigger re-review, modification, temporary suspension, or termination of AI solutions.

Human factors monitoring (e.g., override behavior tracking and over-reliance/automation bias detection) should be incorporated into monitoring programs. Training updates should be triggered by monitoring findings that reveal problematic usage patterns.

Monitoring should vary in frequency and intensity based on the risk of the AI solution. Higher risk solutions require more frequent review cycles and enhanced escalation pathways. Risk-proportionate monitoring ensures resources are directed where patient safety exposure is greatest.

Note: Monitoring Continuously Adaptive or Self-learning AI Systems

Continuously adaptive or self-learning AI systems are becoming increasingly common. To support these types of deployments, HDOs may consider a few monitoring behaviors:

(1) Periodic batch retraining (e.g., discrete event, clearly triggers re-assessment notification)

(2) Continuous online calibration (e.g., ongoing, requires pre-defined performance threshold triggers rather than event triggers — e.g., "if performance on any key metric changes by more than X% from baseline on a rolling 90-day window, the change constitutes a governance notification event")

(3) Static model / Dynamic retrieval corpus (e.g., retrieval-augmented generation architectures — the model is frozen but the data it reasons over changes, which requires corpus governance policies rather than model retraining governance).

Termination *(if applicable)*

Termination involves the planned or unplanned retirement, removal, or disabling of an AI solution. Organizations should define thresholds and criteria for termination in advance, informed by safety concerns, performance degradation, changes in intended use, regulatory developments, or organizational priorities. Prior to sunset, organizations should assess the operational, clinical, and workflow impacts of removal and ensure appropriate transition planning, communication, and data handling. Documented termination criteria, a stakeholder notification plan, archival of monitoring data, and a lessons-learned review are important components of any AI sunsetting process. Formal sunsetting processes and audit retention are industry best practice; AI removal can create clinical disruption and liability exposure if unmanaged.

Lastly, HDOs should explicitly assess whether the AI solution *can* be terminated. For AI deeply embedded into a product or solution that is mandatory for a workflow, termination may be technically infeasible without removing the entire system, in which case compensating controls should be documented. Termination should also require a data-disposition sign-off covering model weights, inference logs, embeddings, and cached predictions, each mapped to a retention or destruction decision consistent with the HDO's Business Associate Agreement (BAA) terms.

****

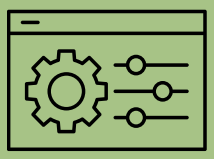
Across all lifecycle steps, organizations are encouraged to introduce formal Decision Gates, requiring documented approval status, evidence reviewed, and defined next review trigger dates. Decision Gates mitigate informal progression through lifecycle steps and create auditable governance checkpoints.

Keeping all the above lifecycle and ingestion modality information in mind, the below implementation guidance represents how HDOs may consider operationalizing each of the baseline operational controls. Each control includes a table with (1) columns = ingestion modality, (2) rows = AI lifecycle steps, and (3) cells = populated with **example** questions to help guide organizations to operationalize the control (if/where the control meaningfully applies).

Key Tools & Resources

CHAI Responsible AI Guide (RAIG) and Responsible AI Guide Executive Summary

- *A playbook for the development and deployment of AI in healthcare, providing actionable guidance on ethics and quality assurance. It stems from the consensus-based approach of the CHAI community, drawing upon the collaborative work of patient advocates, technology developers, clinicians, data scientists, civil servants, bioethicists, and others.*
- *Note: CHAI's RAIG highlights core, industry-aligned responsible AI (RAI) principles (usefulness/usability/efficacy, fairness/bias management, safety/reliability, transparency, privacy, and security) that should be considered as HDOs develop their AI lifecycle management processes.*



Baseline Operational Controls for Subdomain 4.1: Responsible AI Lifecycle Management and Use

Develop and maintain a process that appropriately assesses and documents intended use of each AI solution.
Define responsible AI principles, as aligned with the organization's mission and goal (e.g., usefulness, fairness, safety, transparency, security, privacy); document the objectives and processes that the organization will use to align with these principles across relevant lifecycle steps.
Operationalize a process to evaluate defined responsible AI principles (e.g., usefulness, fairness, safety, transparency, security, privacy) and reference industry recommended evaluation frameworks where available.
Define a process to monitor the safety, performance, and responsible use of AI solutions.
Establish a monitoring mechanism or process for ongoing performance reporting; monitoring mechanisms or processes should flag unexpected patterns for review by responsible parties.
Establish a process to inform relevant stakeholders about responsible AI principles and processes (e.g., definitions, evaluation, monitoring).

Assess and Document Intended Use of Each AI Solution

Control 4.1.1: Develop and maintain a process that appropriately assesses and documents intended use of each AI solution.

Goal: Prior to managing the lifecycle of an AI solution, HDOs should first understand how the solution enters the organization, as the ingestion modality fundamentally shapes lifecycle management expectations and downstream controls. AI solutions may be introduced through research activities, internal development, vendor or co-development arrangements, or as AI-enabled features embedded within existing health IT systems - each presenting distinct challenges related to accountability and scope control. This control requires organizations to establish a standardized process to explicitly assess and document the intended use of each AI solution at intake and revisit it across key lifecycle steps. Clear documentation of intended use helps prevent scope creep, supports appropriate risk categorization and mitigation, and ensures that AI systems are deployed, monitored, and retired in alignment with organizational objectives.

Implementation Guidance

Included below is a template for how HDOs may think through operationalizing **control #1**. Within the template are decision gate considerations (e.g., lifecycle step owner(s), helpful artifacts, decision criteria, and guiding questions) across each lifecycle step that HDOs can use as step-by-step guidance during the implementation and operationalization of AI lifecycle management. Guiding questions depend on (1) ingestion modality and (2) AI lifecycle step. All content included in the below table are examples, not prescriptive. The decision gate considerations and guiding questions are designed to be adapted to your organizational context, industry best practices, and maturity.

	Decision Gate Considerations *(examples, not prescriptive)*
Intake/ Procurement	**Lifecycle step owner(s):** Executive sponsor, operational/business owner, AI governance committee chair, intake review lead(s). **Helpful artifacts:** Model card, intake questionnaire, vendor terms & conditions, architecture diagrams (including any third-party LLM API dependencies). **Decision criteria:** • Stated problem and proposed AI solution align with organizational goals • Intended user population is explicitly named • Permissible and non-permissible uses are defined by the governance body • Risk tolerances and contract architecture are understood • Accountable ownership (e.g., executive sponsor + operational owner) is in place **Guiding questions:** • Is the defined problem consistent with goals, user needs, and risk tolerances? • Who specifically are the intended users (e.g., physicians, registered nurses, medical students, call-center agents, patients), and do the users have appropriate licensure, expertise, and availability? • Has a designated accountable owner(s) been identified? • In its intended use, will the AI solution directly target the stated problem? • Have permissible and non-permissible uses been defined by the governance body? • Does the product architecture (including any third-party LLM API calls) map cleanly to applicable contract terms, BAA, and acceptable-use policy?
Risk	**Lifecycle step owner(s):** Risk committee or risk officer, clinical safety lead, AI governance committee **Helpful artifacts:** Risk categorization tool, risk and impact assessments **Decision criteria:** • Risk categorization is assigned based on organizational risk modifiers • Downstream validation, monitoring, and/or training rigor will scale to the category **Guiding questions:** • Is there a framework for categorizing and evaluating risks and impacts? • Has the intended use been categorized by risk accounting for use case, user authority, and reversibility of harm? • Are validation, monitoring, and/or training expectations tied to the risk category?
Pre-deployment	**Lifecycle step owner(s):** Operational owner, validation lead, IT / security. **Helpful artifacts:** Validation evidence, integration plan, permissible / non-permissible use list.

	Decision criteria: • Intended use is finalized and matches what will be communicated to users • Selected features align with the overarching design and architecture • Evidence is sufficient for the assigned risk category **Guiding questions:** • Do selected features align with the overarching design and architecture? • Has the intended use been validated for the specified user role and context? • Are permissible / non-permissible uses reflected in user-facing disclosures and training plans?
Piloting	**Lifecycle step owner(s):** Pilot lead, clinical / operational champion, intended user representatives. **Helpful artifacts:** Pilot plan, override / disagreement log, end user feedback mechanism, data-environment segregation plan. **Decision criteria:** • Pilot is bounded to documented intended use • Deviations are detected and reviewed • Pilot data is segregated from production training pipelines **Guiding questions:** • Is there a plan to manage end user disagreements with AI outputs? • Is there a method to measure whether user actions are consistent with intended use? • Are actions taken by end users different than originally anticipated? • Are pilot data environments segregated to prevent contamination of production training pipelines?
Deployment	**Lifecycle step owner(s):** Deployment lead, operational owner, governance committee. **Helpful artifacts:** Approved intended use statement, training plan, audit-log configuration, patient opt-out documentation. **Decision criteria:** • Intended use is communicated and trained on • Mechanisms exist to monitor adherence and to surface opt-out paths • A re-evaluation cadence is defined **Guiding questions:** • Is data available to evaluate if the system is being used as intended? • Has a defined cadence been established to re-evaluate intended use post-deployment (e.g., annual or upon material change)? • Have opt-out paths for users and patients been documented and surfaced?
Monitoring	**Lifecycle step owner(s):** Named monitoring owner, performance review committee.

	Helpful artifacts: Usage audits, drift dashboards (internally-developed or via vendor), off-label / unintended-use logs, subgroup performance reports. **Decision criteria:** • Off-label use, drift, and unintended-use patterns trigger review at predefined thresholds • Subgroup performance is examined **Guiding questions:** • Are there audits to identify "off-label" or unintended uses, and is there a mechanism to detect unintended use? • Are specific criteria defined for shifts in model performance within subgroups? • Is there a mechanism for reporting unintended uses to stakeholders? • Are there technically defined thresholds for significant data drift?
Termination	**Lifecycle step owner(s):** Operational owner, data governance, governance committee. **Helpful artifacts:** Termination decision record, transition plan for users / patients, data disposition plan, dependency map. **Decision criteria:** • A termination path exists, or compensating controls are documented if the AI is embedded such that it cannot be cleanly removed • User / patient transitions are planned • Data disposition is signed off **Guiding questions:** • Is there guidance on transitioning patients / users to a new solution? • Are plans in place for handling patient data (migration / deletion) during sunsetting? • Can the AI solution be reasonably or technically terminated, or is it so embedded that compensating controls are required? • Has data-disposition sign-off covered model weights, inference logs, embeddings, and cached predictions, mapped to retention / destruction decisions consistent with BAA terms?

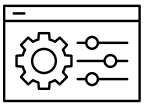
Note: Differences across Ingestion Modalities

When reviewing/implementing decision gates across the lifecycle steps, HDOs should keep the below differences in mind across ingestion modalities.

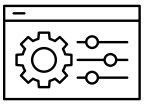
- Research: Intended use is typically narrow and IRB-scoped. Define an explicit onramp into production if a research project succeeds.
- In-house developed: The organization has the greatest control over intended-use definition; the dominant risk is typically biased toward technical feasibility.
- Vendor / Co-developed: Intended use may be shaped by vendor marketing claims; the governance committee should define acceptable use.
- AI-enhanced features: Intended use should be re-documented for the specific organizational deployment. Limited transparency may force reliance on contractual disclosures.

A note on agentic solutions: Intended use should extend beyond the model/solution to bound the agent's autonomy, including the allowed tool inventory, the decision-delegation rules, the potential human-on-the-loop checkpoints, and the reversibility of actions taken. The permissible / non-permissible use list should explicitly cover autonomous actions.

Define and Operationalize Responsible AI Principles



Control 4.1.2: Define responsible AI principles, as aligned with the organization's mission and goal (e.g., usefulness, fairness, safety, transparency, security, privacy); document the objectives and processes that the organization will use to align with these principles across relevant lifecycle steps.



Control 4.1.3: Operationalize a process to evaluate defined responsible AI principles (e.g., usefulness, fairness, safety, transparency, security, privacy) and reference industry recommended evaluation frameworks where available.

Goal: HDOs should define, operationalize, and consistently apply their Responsible AI (RAI) principles (e.g., usefulness, fairness, safety, transparency, security, privacy) across all AI systems, with expectations appropriately scaled based on ingestion modality, intended use, and risk. This control requires organizations to document clear objectives and processes for how each principle will be evaluated at relevant lifecycle steps, from intake and risk categorization through deployment, monitoring, and termination. Objective evaluation should leverage industry-recommended frameworks (e.g., CHAI Testing & Evaluation Frameworks) and combine first-party testing, vendor attestations, and outcome-based monitoring as appropriate. Where full evaluation is constrained (e.g., due to vendor opacity or embedded AI features), organizations should explicitly document limitations, compensating controls, and residual risk to ensure principled and consistent lifecycle management decisions.

Controls #2 and #3 are complementary and should be operationalized together. HDOs should operationalize and apply RAI principles consistent with emerging professional consensus and best practice in relevant literature, and pair principles with consistently operationalizable criteria. Control #2 focuses on defining and documenting how RAI principles will be applied. Control #3 focuses on objective measurement and evaluation of those principles.

Key Tools & Resources

CHAI Testing & Evaluation (T&E) Frameworks

- *A Testing and Evaluation (T&E) Framework includes a consensus-defined set of methods, metrics, measures, and/or benchmarks for developers and implementers to more concretely evaluate the responsible use of health AI solutions for a given use case.*

Implementation Guidance

Included below is a template for how HDOs may think through operationalizing **controls #2 and #3**. Within the template are decision gate considerations (e.g., lifecycle step owner(s), helpful artifacts, decision criteria, and guiding questions) across each lifecycle step that HDOs can use as step-by-step guidance during the implementation and operationalization of AI lifecycle management. Guiding questions depend on (1) ingestion modality and (2) AI lifecycle step. All content included in the below table are examples, not prescriptive. The decision gate considerations and guiding questions are designed to be adapted to your organizational context, industry best practices, and maturity.

	Decision Gate Considerations *(examples, not prescriptive)*
Intake/ Procurement	**Lifecycle step owner(s):** AI governance committee, RAI principles owner, ethics lead. **Helpful artifacts:** Organizational RAI principles statement (anchored in industry consensus), evaluation framework references, vendor RAI commitments and supporting evidence. **Decision criteria:** • RAI principles are operationalized consistent with emerging industry consensus • Criteria are consistently operationalizable • Evaluation evidence aligns to recognized frameworks • Evaluation limitations are documented up front **Guiding questions:** • Have RAI principles been operationalized and applied consistent with emerging industry consensus, and paired with consistently operationalizable criteria? • Has the developer / vendor provided objective evidence aligned with recognized evaluation frameworks? • For embedded features, have evaluation limitations been documented before activation?
Risk	**Lifecycle step owner(s):** RAI and/or governance committee, named risk mitigation owner(s). **Helpful artifacts:** Risk and impact assessments, risk mitigation library, residual-risk register. **Decision criteria:** • Risks are documented with a named owner for every AI solution regardless of ingestion modality • Opacity is addressed with documented mitigation controls • Risk assessment reflects the specific context and population **Guiding questions:** • Have risks been documented with a named owner for mitigation, regardless of ingestion modality? • Where transparency is limited (vendor opacity or embedded logic), are operational mitigation controls documented? • Have risk domains been assessed against the relevant population and clinical / operational context?
Pre-deployment	**Lifecycle step owner(s):** Validation lead, data science / quality lead. **Helpful artifacts:** Local validation evidence, vendor claims documentation, disclosure plan.

	Decision criteria: • Evidence standards are risk-tiered • Transparency and outcome proxies are defined before go-live • Vendor claims are either locally validated or constraints documented. **Guiding questions:** • Has objective validation confirmed alignment with RAI principles using representative production data? • Has independent testing verified usefulness, fairness, safety, transparency, privacy, security, etc. at a rigor proportionate to the risk tier? • For vendor or embedded AI, has the organization validated claims locally or documented constraints? • Have transparency, disclosure, and outcome-proxy measures been defined before enabling AI functionality?
Piloting	**Lifecycle step owner(s):** Pilot lead, RAI principles owner, safety officer. **Helpful artifacts:** Pilot safety-monitoring plan, override / deviation logs, vendor performance-feedback loop, pilot performance dashboard. **Decision criteria:** • Safety signals and override rates are observed during pilot • Vendor commitments are validated through pilot performance data. **Guiding questions:** • Are safety signals monitored prospectively during pilot? • Are override rates and unexpected workflow deviations documented? • Is pilot performance data shared with and from any vendor and used to validate RAI commitments? • Are unintended workflow or risk impacts measured during pilot?
Deployment	**Lifecycle step owner(s):** Governance committee, deployment approver. **Helpful artifacts:** Approval / sign-off documentation, threshold definitions, communication plan. **Decision criteria:** • Deployment is contingent on meeting predefined RAI thresholds. **Guiding questions:** • Are deployment approvals contingent on meeting predefined RAI thresholds? • Has the governance committee approved the permissible-use criteria?
Monitoring	**Lifecycle step owner(s):** Named monitoring owner, RAI and/or governance committee.

	Helpful artifacts: Drift dashboards or related monitoring mechanism, subgroup reports, vendor reporting / analytics outputs (as able). **Decision criteria:** ● RAI metrics are tracked longitudinally with named accountability and surfaced regardless of where the AI sits in the stack. **Guiding questions:** ● Are RAI metrics tracked longitudinally, including drift, subgroup bias, and real-world safety outcomes? ● Is a named monitoring owner accountable for regular review of RAI metrics? ● For vendor / embedded AI: How will the vendor enable data access and visibility to AI system performance? Have vendors been asked to provide reporting / analytics
Termination	**Lifecycle step owner(s):** RAI committee, governance committee, data governance owner(s) **Helpful artifacts:** Sunsetting decision record, lessons-learned documentation, vendor-exit summary. **Decision criteria:** ● RAI performance informs sunsetting ● Performance failures are explicit grounds for termination ● Lessons are captured. **Guiding questions:** ● Are RAI metrics reviewed as part of the sunsetting decision? ● Are vendor performance failures against RAI commitments grounds for termination? ● Are lessons learned on RAI performance documented for future model or solution development?

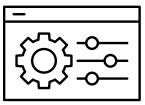
Note: Differences across Ingestion Modalities

When reviewing/implementing decision gates across the lifecycle steps, HDOs should keep the below differences in mind across ingestion modalities.

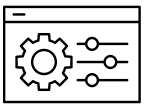
- Research: RAI principle alignment is typically evaluated against research objectives and the IRB-defined population; future production use will require re-evaluation.
- In-house developed: The organization owns the evidence end-to-end and is responsible for both defining and operationalizing RAI principle adherence.
- Vendor / Co-developed: Evidence is shared / supplied; opacity should be treated as a residual risk with internal, operational controls.
- AI-enhanced features: Evidence is typically least visible; RAI principle adherence often depends on disclosure attestations and contractually-secured reporting access.

A note on agentic solutions: RAI principles should extend to the agent's autonomous decision space (e.g., transparency of the agent's reasoning trace, the scope of tool access, and the reversibility of actions). Fairness includes equitable treatment across actions an agent might take, not only the output it produces.

Monitor AI Safety, Performance, and Responsible Use



Control 4.1.4: Define a process to monitor the safety, performance, and responsible use of AI solutions.



Control 4.1.5: Establish a monitoring mechanism or process for ongoing performance reporting; monitoring mechanisms or processes should flag unexpected patterns for review by responsible parties.

Goal: HDOs should implement a risk-proportionate monitoring process to continuously evaluate the safety, performance, and responsible use of AI systems across relevant lifecycle steps, with monitoring expectations informed by ingestion modality, intended use, and risk categorization. This control requires establishing a centralized monitoring mechanism - such as a dashboard or equivalent reporting interface - that enables frequent review of key performance, utilization, and bias indicators, and flags unexpected or out-of-scope usage patterns for manual review by IT, compliance, and other designated oversight functions. Where direct model telemetry is unavailable (e.g., vendor-managed or embedded AI features), organizations should monitor downstream workflows (e.g., outcome proxies and user behavior) and document escalation pathways, remediation actions, and termination thresholds to ensure ongoing safety and accountability.

Note: Monitoring Maturity Pathway

- **Foundational:** Organizations should have a named monitoring owner(s), a defined review cadence (e.g., quarterly), and documented triggers that warrant re-review or escalation. Even organizations without formal dashboards should have these elements.
- **Structured:** This may include standardized monitoring checklist, tracking log documenting review outcomes, and formal process for escalation to compliance or AI governance committee.
- **Advanced:** This may include automated dashboards with alert-based escalation and triage; integration with patient safety reporting systems; vendor-enabled data access for centralized performance analytics.

Implementation Guidance

Included below is a template for how HDOs may think through operationalizing **controls #4 and #5**. Within the template are decision gate considerations (e.g., lifecycle step owner(s), helpful artifacts, decision criteria, and guiding questions) across each lifecycle step that healthcare delivery organizations can use as step-by-step guidance during the implementation and operationalization of AI lifecycle management. Guiding questions depend on (1) ingestion modality and (2) AI lifecycle step. All content included in the below table are examples, not prescriptive. The decision gate considerations and guiding questions are designed to be adapted to your organizational context, industry best practices, and maturity.

	Decision Gate Considerations *(examples, not prescriptive)*
Intake/ Procurement	**Lifecycle step owner(s):** Monitoring owner, IT operations, contract / vendor manager. **Helpful artifacts:** Monitoring plan, key performance indicator (KPI) definitions, data-access / contract terms, log-availability assessment. **Decision criteria:** • Monitoring ownership, infrastructure, and KPIs are defined before go-live • Data access (especially from vendors) is contractually secured • Log availability is confirmed **Guiding questions:** • Has ownership, infrastructure, and KPIs been established for ongoing monitoring? • For vendor / embedded AI, does the contract define shared safety-monitoring responsibilities and data access for dashboard reporting (or similar mechanism)? • Has IT confirmed availability of usage logs and defined how AI safety signals will be tracked?
Risk	**Lifecycle step owner(s):** Risk committee, governance committee, monitoring owner(s). **Helpful artifacts:** Risk categorization definitions, monitoring-frequency matrix, escalation pathway documentation. **Decision criteria:** • Monitoring intensity scales with risk tier • High-risk solutions have enhanced thresholds and escalation • Vendor dependency risk is reflected in monitoring frequency **Guiding questions:** • Is monitoring intensity and reporting frequency aligned with the solution's risk category? • Are high-risk AI solutions assigned enhanced monitoring thresholds and escalation pathways? • Has vendor dependency risk been reflected in monitoring frequency and dashboard oversight?
Pre-deployment	**Lifecycle step owner(s):** IT, DevOps, monitoring owner, safety officer. **Helpful artifacts:** Dashboard test results, alert-logic specifications, escalation runbook (e.g., a tactical guide that outlines exactly who to notify, when to do so, and how to hand off an issue when an incident cannot be resolved). **Decision criteria:** • Monitoring infrastructure is functional and validated before go-live.

	Guiding questions: • Are predefined safety thresholds, alert triggers, and escalation workflows configured before pilot launch? • Has dashboard functionality and alert logic been tested prior to go-live?
Piloting	**Lifecycle step owner(s):** Pilot lead, monitoring owner. **Helpful artifacts:** Adverse-event log, override-tracking dashboard, automation bias indicators, vendor performance feed. **Decision criteria:** • Monitoring during pilot captures adverse events, near misses, override behavior, and reliance patterns • Vendor performance is integrated. **Guiding questions:** • Are adverse events, near misses, and unexpected usage patterns systematically logged and reviewed through a dashboard or similar mechanism? • Are override rates and workflow deviations actively flagged for manual review? • Are automation bias indicators or unusual reliance patterns detected during pilot? • For vendor AI, is pilot performance data shared with the vendor and incorporated into dashboard reporting or similar mechanism?
Deployment	**Lifecycle step owner(s):** Monitoring owner, governance committee. **Helpful artifacts:** Enterprise monitoring report, incident-integration logs, spot-check schedule. **Decision criteria:** • Monitoring is integrated into enterprise reporting • Vendor incident reports flow into organizational monitoring cycles • Embedded AI is periodically spot-checked **Guiding questions:** • Is vendor incident reporting integrated into organizational monitoring cycles? • Are AI outputs (including embedded features) periodically spot-checked and reflected in enterprise reporting?
Monitoring	**Lifecycle step owner(s):** Monitoring owner, compliance / safety lead. **Helpful artifacts:** Monitoring review records, change-management records, anomaly logs. **Decision criteria:** • Certain model and solution updates (vendor- or internally-driven) automatically trigger renewed monitoring review • Anomalies and usage spikes are reviewed

	Guiding questions: • Do model and solution updates (vendor- or internally-driven) automatically trigger renewed safety-monitoring review? • Are workflow anomalies and usage spikes flagged for compliance review?
Termination	**Lifecycle step owner(s):** Monitoring owner, governance committee, data governance. **Helpful artifacts:** Performance trend reports, archived dashboard data, retirement decision record. **Decision criteria:** • Sustained performance degradation or safety-threshold breaches trigger retirement review • Dashboard (or similar mechanism) data is archived per retention policy before shutdown. **Guiding questions:** • Have predefined safety thresholds or sustained performance degradation triggered retirement review? • Is dashboard (or similar mechanism) data archived and reviewed prior to shutdown decisions?

Note: Differences across Ingestion Modalities

When reviewing/implementing decision gates across the lifecycle steps, HDOs should keep the below differences in mind across ingestion modalities.

- Research: Monitoring is typically time-bounded and tied to study endpoints rather than continuous performance.
- In-house developed: Organizations have the most direct telemetry access but must invest in the monitoring infrastructure themselves.
- Vendor / Co-developed: Monitoring depends on contractual data access and vendor cooperation.
- AI-enhanced features: Monitoring is often constrained to what the host platform exposes. Where the vendor does not provide adequate telemetry, organizations should work with vendors to align on contractual controls or processes. HDOs should keep in mind that vendor model updates may occur silently as part of system upgrades.

A note on agentic solutions: Monitoring must extend beyond model output drift to *action drift* (e.g., the actions the agent takes, tool-call success / failure rates, hallucinated tool invocations, and unexpected action sequences). Run traces and replay logs become important monitoring artifacts. Automation bias and over-reliance indicators are especially material when an agent can act autonomously.

Stakeholder Awareness on Responsible AI

Control 4.1.6: Establish a process to inform relevant stakeholders about responsible AI principles and processes (e.g., definitions, evaluation, monitoring).

Goal: HDOs should embed Responsible AI Use policy training into new hire orientation and recurring compliance education to ensure that all relevant stakeholders – including clinicians, administrators, IT staff, operational leaders, and patients – understand how AI systems are used within the organization, their intended use, limitations, and associated risks. Training content should be role-based and aligned to ingestion modality, emphasizing differences in accountability, transparency, and appropriate reliance for research models/solutions, internally developed solutions, vendor-developed solutions, and AI-enabled features embedded in existing health IT systems. Effective training is a primary mechanism for reducing AI risk. Organizations should verify that planned training is being implemented and refreshed on the cadence required by the risk category.

Implementation Guidance

Included below is a template for how HDOs may think through operationalizing **control #6**. Within the template are decision gate considerations (e.g., lifecycle step owner(s), helpful artifacts, decision criteria, and guiding questions) across each lifecycle step that healthcare delivery organizations can use as step-by-step guidance during the implementation and operationalization of AI lifecycle management. Guiding questions depend on (1) ingestion modality and (2) AI lifecycle step. All content included in the below table are examples, not prescriptive. The decision gate considerations and guiding questions are designed to be adapted to your organizational context, industry best practices, and maturity.

	Decision Gate Considerations *(examples, not prescriptive)*
Intake/ Procurement	**Lifecycle step owner(s):** Education / training lead, RAI program owner, intended user representatives. **Helpful artifacts:** Stakeholder map, training-needs assessment, draft role-based curriculum, AI-disclosure standards. **Decision criteria:** • Training needs are identified by intended user role at intake • Required curriculum scope is included in intake documentation. **Guiding questions:** • Have the specific intended user roles been identified, and have training needs been mapped to each role? • Have training requirements been captured as part of the intake decision package?
Risk	**Lifecycle step owner(s):** RAI training lead, risk committee **Helpful artifacts:** Risk categorized training matrix **Decision criteria:** • Training depth and refresh cadence align with the AI's risk category **Guiding questions:** • Does training depth and frequency scale with the risk category of the AI solution? • Does training reflect the licensure and expertise of the intended user(s)?
Pre-deployment	**Lifecycle step owner(s):** Training lead, vendor liaison (if applicable) **Helpful artifacts:** Training completion records, role-based curricula, AI disclosure standards. **Decision criteria:** • Role-based training is completed before access is granted • AI disclosures and embedded-feature labels are part of workflow education **Guiding questions:** • Is role-based training completed prior to granting system access? • For vendor AI, is vendor co-facilitated training delivered before go-live? • Are AI disclosures (including for embedded features) incorporated into workflow education?
Piloting	**Lifecycle step owner(s):** Pilot lead, training lead

	Helpful artifacts: Pilot misuse-pattern reports, refresher training materials, vendor feedback **Decision criteria:** • Misuse, misunderstanding, and early safety findings from pilot feed back into targeted retraining **Guiding questions:** • Are pilot misuse patterns and early safety findings addressed through targeted retraining? • Are user misunderstandings corrected through structured education sessions? • For vendor AI, is vendor feedback integrated into updated training materials?
Deployment	**Lifecycle step owner(s):** Training lead, compliance / onboarding lead. **Helpful artifacts:** Any additional onboarding curricula **Decision criteria:** • RAI training is embedded in onboarding and relevant compliance programs • Embedded features are clearly labeled • Administrators trained on monitoring / escalation **Guiding questions:** • Is RAI training embedded in onboarding and relevant compliance programs, and is it being implemented as planned? • Are administrators trained on monitoring and escalation responsibilities? • Are vendor updates reflected in training refresh cycles? • Are embedded AI capabilities clearly labeled in training materials?
Monitoring	**Lifecycle step owner(s):** Training lead, monitoring owner. **Helpful artifacts:** Monitoring findings reports, training update logs, case-study library **Decision criteria:** • Monitoring findings translate into recurring training updates and case studies. **Guiding questions:** • Are monitoring findings (including safety incidents and identified misuse trends) incorporated into recurring training updates and used as case studies? • Are workflow changes communicated through updated training?
Termination	**Lifecycle step owner(s):** Training lead, communications lead. **Helpful artifacts:** Sunset communication plan, post-termination lessons-learned document.

Decision criteria:

- Staff and intended users are informed of changes
- Lessons are captured into future training and policy education

Guiding questions:

- Are staff educated on workflow changes following AI retirement and informed when embedded AI features are disabled?
- Are post-termination lessons (including vendor exit lessons) integrated into future training modules and policy education?

Note: Differences across Ingestion Modalities

When reviewing/implementing decision gates across the lifecycle steps, HDOs should keep the below differences in mind across ingestion modalities.

- Research: Training is typically protocol-specific and IRB-aligned.
- In-house developed: Training scope is more visible because deployment is intentional and internally developed.
- Vendor / Co-developed: Vendor-supplied materials should be reviewed and adapted to local workflows.
- AI-enhanced features: Training should explicitly label what is AI-driven, what user actions imply trust delegation, and what disclosures are required to patients or downstream recipients of the output.

A note on agentic solutions: Training should cover how to interpret and verify an agent's proposed action plan, where humans can/should intervene, and how to safely abort or roll back agent actions. Intended users need explicit education on the difference between recommendation-only AI and action-taking AI, and on the boundaries of the agent's tool access.

Real-World Examples

Explore these real-world examples of how HDOs are successfully approaching responsible AI lifecycle management. These examples offer valuable insights and practical strategies that you can adapt to your own organization. CHAI recognizes that organizations needing the most governance support (including single systems, rural health systems, and non-AMC organizations) may have different structures, resources, and maturity levels than some of the large academic medical centers featured below.

Example #1: Memorial Sloan Kettering

The iLEAP Operating Model

The cornerstone of the lifecycle management is the iLEAP model, which stands for Legal, Ethics, Adoption, and Performance. This framework provides three distinct pathways for AI practitioners based on the model's origin:

- Research Path: Governed primarily by the Institutional Review Board (IRB) to preserve scientific freedom.
- Home-Grown Build Path: For models developed internally by hospital departments.
- Acquired/Purchased Path: For third-party vendor tools and collaborations.

AI Lifecycle Management Gates

The management process is structured around five specific decision gates (G1-G5) that a model must pass before and after deployment:

- G1 (Model Proposal): Proposals are reviewed for strategic alignment and informatics feasibility.
- G2 (Model Registration): The model is officially entered into a centralized Model Registry.
- G3 (Model Information Sheet): Practitioners must complete a "nutrition card" detailing data sources, training methods, and a risk assessment. This stage includes screening for FDA regulatory status.
- G4 (Model Monitoring Plan): Before full launch, models undergo silent evaluation in a testing environment to validate performance.
- G5 (Model Monitoring Report): Post-launch, the AIGC reviews performance metrics, clinician trust (using the TrAAIT tool), and any AI-induced Adverse Events (aiAEs). This may also include model retirement.

The "Express Pass" Methodology

To prevent governance from stifling innovation, the organization utilizes an "Express Pass" for rapid review. Models can qualify for an expedited path—sometimes receiving approval in as little as two weeks—if they are low-risk, have an engaged sponsor, use existing core platforms, and meet all predefined "gate criteria".

Safety and "Algorithmovigilance"

A unique feature of this lifecycle management approach is the proactive monitoring of safety through algorithmovigilance. The system is designed to intercept and remediate aiAEs by leveraging existing safety-event reporting tools. This ensures the AIGC has the authority to pause or terminate models if they demonstrate degraded performance or safety issues.

Source: Responsible Artificial Intelligence governance in oncology

Example #2: Stanford Health Care

At Stanford Health Care, AI is governed through an enterprise Responsible AI Lifecycle (RAIL) led by a multidisciplinary group spanning academic and clinical experts (physicians, nurses), legal and compliance partners, security and privacy leaders, computer scientists and data scientists, ethicists, and executive sponsors—with additional perspective from our Patient and Family Advisory Council. Every AI request enters through a standardized intake process, is risk-tiered, and is routed to the appropriate level of governance and required assessments to ensure Fair, Useful, Reliable Models (FURM). Higher-risk use cases—particularly those that could influence clinical decisions or create potential patient harm—undergo a rigorous end-to-end review and validation pathway along with a monitoring plan, while lower-risk changes (e.g., productivity-focused updates or routine vendor releases) follow a lighter-weight process proportional to risk but always requiring a monitoring plan. Governance is operationalized by codifying intake, decisions, and required artifacts in our project management and CMDB (Configuration Management Database) tool for consistency and traceability, and by aligning governance checkpoints with product development and DevOps release cycles so testing, release controls, and post-go-live monitoring are integral parts of delivery. The process undergoes periodic review and updates led by the Project Management office in the IT department at SHC.

Source: Stanford Health Care Responsible AI Lab (RAIL)

Example #3: Five Pillar Lifecycle Management at a Regional Health System

Background and Objectives

To ensure the safe and effective adoption of artificial intelligence, a prominent regional health system established a comprehensive five-pillar lifecycle management framework. This framework governs the entire lifecycle of AI tools - from initial proposal to long-term clinical integration - ensuring they serve clear clinical needs rather than acting as "technology for its own sake".

AI Lifecycle Management Pillars

1. **Utility, Validity, and Validation**: The lifecycle begins by identifying AI tools that solve specific clinical problems effectively. Once a tool is proposed, it undergoes rigorous testing using real-world data that reflects the institution's unique patient demographics and conditions. This ensures the model performs reliably and consistently across diverse populations before it is deployed.
2. **Ethics and Bias Mitigation**: A critical stage in the management process is the assessment of fairness and transparency. The framework requires identifying and mitigating biases within training data to prevent unequal treatment and ensure consistent outcomes across all patient groups. Additionally, the process is designed to ensure AI enhances, rather than replaces, human decision-making and autonomy in clinical settings.
3. **Compliance, Risk, and Cybersecurity**: The management lifecycle includes a strict review of regulatory adherence, specifically ensuring tools comply with Centers for Medicare & Medicaid Services (CMS) and FDA standards. To manage organizational risk, the system documents the impact of all AI tools to protect against legal issues
4. **Cybersecurity**: Simultaneously, the Cybersecurity pillar enforces strong access controls and secure model development to protect sensitive data and maintain operational resilience against breaches.
5. **Learner Impact and Adaptation**: Unique to this lifecycle approach is an emphasis on educational support. The framework ensures that AI tools act as mentors to reinforce foundational clinical skills and are adaptable to various learning environments, specifically benefiting those in academic settings or early in their careers,.

Governance and Decision-Making Process

The lifecycle is managed through a structured, multi-tiered review process:

- Independent Review: A core team first conducts an independent assessment of the AI tool.
- Multidisciplinary Committee: A broad team of experts—including clinical, operations, IT, research, legal, compliance, and cybersecurity professionals—meets monthly to evaluate the risks and benefits.
- Approval and Guardrails: This committee is responsible for voting on the approval of new tools and implementing the necessary guardrails to ensure ongoing safety and efficacy.

Example #4: Operationalizing AI Lifecycle Management in a Large Health System

AI Governance Lifecycle Structure

To manage the AI lifecycle, the institution created the AI & Automation Advisory (AAA) group, which functions as a strategic subgroup of the broader Information Services Oversight Committee (ISOC),. This group oversees the strategic alignment and adoption of AI across two primary levels:

- Strategic: Involvement from transformation executives and entity leadership.
- Tactical: Six multidisciplinary workgroups—including Clinical, Research & Analytics, and Business & Revenue Cycle—identify and vet operational use cases.

AI Lifecycle Management Process

The institution utilizes a structured Decision to Partner Framework and a high-level RAI process to move an AI model/solution from inception to deployment:

1. Categorization and Risk Assessment: Use cases are identified and mapped on a Risk/Impact matrix. High-impact, low-risk projects are categorized for "Rapid Innovation," while high-risk projects may be placed in "Watchful Waiting" or require "Aggressive Pursuit" only with trusted partners.
2. RAI Survey and Evaluation: For every "build" (internal) or "buy" (vendor) solution, an RAI Survey is completed. A core team then reviews and scores these responses using a Responsible AI Evaluation Rubric.
3. The Four Pillars of Review: The management lifecycle focuses on four critical pillars:
    a. **Fairness**: Ensuring models are representative of the system's patient population and include processes for recalibration if the population changes.
    b. **Transparency**: Documenting intended users, out-of-scope use cases, and the model or solution's building/maintenance process.
    c. **Accountability**: Defining clear roles for developers and subject matter experts (SMEs), and establishing processes to track performance over time.
    d. **Trustworthiness**: Educating clinicians and operational users about how the model works and providing independent validation of findings.
4. Recommendation and Voting: Based on the rubric scores, the core team makes an RAI recommendation to the AAA group, which then votes on whether the solution is approved to proceed into the Enterprise Architecture process.

Operational Principles: FAVES

Throughout the lifecycle, the system adheres to the FAVES AI Principles to ensure all models are:

- **Fair**: Avoiding prejudice toward specific groups.
- **Appropriate**: Matched to specific contexts and populations.
- **Valid**: Accurately estimating targeted values in both internal and external data.
- **Effective**: Demonstrating real-world benefit.
- **Safe**: Ensuring probable benefits outweigh known risks.

Tools & Resources

Included below are tools and resources that may be helpful for developing and operationalizing your own Responsible AI Lifecycle.

CHAI Testing & Evaluation (T&E) Frameworks

- *A Testing and Evaluation (T&E) Framework includes a consensus-defined set of methods, metrics, measures, and/or benchmarks for developers and implementers to more concretely evaluate the responsible use of health AI solutions for a given use case.*

CHAI Partner Ecosystem

- *Collaborate with trusted partners delivering excellence in AI governance, testing, data management, and advisory services - all aligned with the CHAI AI Governance Framework.*

CHAI Responsible AI Guide (RAIG) and Responsible AI Guide Executive Summary

- *A playbook for the development and deployment of AI in healthcare, providing actionable guidance on ethics and quality assurance. It stems from the consensus-based approach of the CHAI community, drawing upon the collaborative work of patient advocates, technology developers, clinicians, data scientists, civil servants, bioethicists, and others.*
- *Note: CHAI's RAIG highlights core, industry-aligned responsible AI (RAI) principles (usefulness/usability/efficacy, fairness/bias management, safety/reliability, transparency, privacy, and security) that should be considered as HDOs develop their AI lifecycle management processes.*

CHAI Responsible AI Lifecycle Management and Use Worksheet

- *This worksheet supports healthcare delivery organizations in operationalizing Subdomain 4.1: Responsible AI Lifecycle Management and Use of CHAI's AI Governance Framework. This worksheet provides structured guidance for assessing intended AI system use, applying Responsible AI principles, and establishing monitoring processes across the full AI lifecycle.*

Next Steps

Links below will connect you to other AI governance playbooks in this series.

Domain 1: AI Policy

Domain 2: Organizational Structures

Domain 3: Organizational Resources

Domain 4: Organizational Processes

- Subdomain 4.1: Responsible AI Lifecycle Management and Use
- Subdomain 4.2: Risk & Impact Assessments
- Subdomain 4.3: Responsible Data Management and Use
- Subdomain 4.4: Third Party Management
- Subdomain 4.5: Education, Training, and Feedback

Attribution

CHAI AI Governance Playbooks were developed through the contributions and collaboration of more than 150 organizations across the healthcare ecosystem. CHAI extends its sincere appreciation to all participating organizations and contributors for their time, expertise, and thoughtful feedback throughout the development process. We would also like to provide special recognition to the individuals below who requested attribution by name for their contributions to these efforts:

- Josh Hjelmstad, MHA, AVIA Health
- Sandra Gregg, DNP, MHA, RN, PMP, LSSBB, CCMP, HCD, Booz Allen Hamilton, Inc.
- Vidhi Vinay Kothari, Carnegie Mellon University
- Stephen McLaughlin, PhD, Children's National Hospital
- Ryan Marshall Felder, PhD HEC-C, Cleveland Clinic
- Po-Hao Chen, MD, MBA, Cleveland Clinic Foundation
- Katherine Mulready, JD MPS, cliexa
- Holden Groves, MD, Columbia University/NewYork-Presbyterian
- Shakira J. Grant, MBBS, MSCR, CROSS Global Research & Strategy
- Namrata Rastogi, MBBS MRCGP MS, Enigma Ventures
- Aleocidio Sette Balzanelo, MD, MBA, FOLKS
- Rocco Orlando, MD, Hartford HealthCare
- Michael Blumental, HealthLeap
- Romanus M. Joseph, Innovative Health Accountable Care Organization (ACO)
- Rebecca G. Mishuris, MD, MS, MPH, Mass General Brigham
- Nasibeh Zanjirani Farahani, PhD, CSM, Mayo Clinic
- Jason Gillman, MD, FIDSA, MCG Health
- Douglas Graham, Individual Contributor
- Nicole Petroski, Mercy Health Services
- Haris Shuaib, Newton's Tree
- Rémy Ibarcq, Individual Contributor
- Srikar Nallan, MS, Stanford Healthcare
- Jaimie Weber, MD, Tampa General Hospital
- R. Brandon Hunter, MD, FAAP, Texas Children's Hospital
- Angela L. Lipscomb, JD, LL.M, The University of Texas MD Anderson Cancer Center
- Christopher H. Lee, MBA, BSN, RN-BC, UCLA Health
- Abby Steele, Individual Contributor
- Shiba Kuanar, University of Minnesota
- Joshua Miller, JD, CHC, University of Rochester Medicine
- Anne Travis, MD, MSc, Wolters Kluwer
- Michael L. Shaw, JD, ZS

License