Responsible AI in Health: From Innovation to Implementation

Merage Ghane, PhD

CHAI

AI THAT SERVES ALL OF US.

Learning Objectives:

- Define Responsible AI in the context of healthcare
- Identify potential legal, ethical, and safety risks
- Learn about available frameworks, tools, and resources
- Discuss governance strategies
- See examples of responsible AI principles applied to real-world use cases.

CHAI

Who is CHAI?

We are an industry-led, public-private partnership committed to advancing AI in healthcare responsibly

We bring together the broad spectrum of interdisciplinary stakeholders in the US Health Ecosystem to drive the development, evaluation and appropriate use of Responsible AI in health.

We Are The World's Largest Health AI Community

4,000

members driving consensus-driven innovation

100+

Professional organizations and patient advocacy groups, ensuring patient-centric work.

200+

Health system members, including MedStar Health, Mercy, Providence, Stanford Medicine and Mayo Clinic.

75%

Of our members define themselves as "industry", of which 24% are startups.

THE AI LIFECYCLE

The AI lifecycle is central to understanding and implementing CHAI's Responsible AI Guidance in healthcare. The six-step lifecycle outlines the essential stages and processes involved in developing, deploying, and maintaining AI systems.

By systematically addressing each phase of the lifecycle, the framework ensures that AI systems adhere to the highest standards of safety, efficacy, fairness, transparency, and security. This structured approach supports risk mitigation, managing biases, and promotes accountability and trustworthiness in AI applications.

1

- Engage stakeholders to define the problem & perform root-cause analysis
- Identify solution & plan future state
- Gather business requirements
- Assess feasibility, potential for impact, & prioritization
- Make procure/build/partner decision

2

- Select/understand model task & architecture
- Capture design & technical requirements or determine best solution to meet business requirements
- Design solution application & system workflow according to human-centered design principles
- Design deployment strategy with end users
- Design risk management, monitoring & reporting plan

3

- Access data
- Prepare data
- Develop data management plan
- Train & tune model

4

- Conduct installation qualification (when applicable)
- Validate local system performance (when applicable)
- Execute prospective, silent evaluation
- Establish risk management plan
- Train end users
- Test usefulness
- Ensure compliance with applicable healthcare regulations & standards

5

- Implement small-scale pilot to assess real-world impact
- Execute and update risk management plan
- Educate & train users on AI application reporting
- Assess usefulness and adoption

6

- Deploy at a larger scale on a general population
- Audit AI system to inform whether to maintain, refine or sunset
- Conduct ongoing risk management

CHAI

RESPONSIBLE AI PRINCIPLES

At its core, governance is to help define, document, and execute processes to ensure that these principles are being upheld at every stage of the AI lifecycle.

- Usefulness, Usability, & Efficacy
- Fairness
- Transparency
- Security & Privacy
- Safety

CHAI

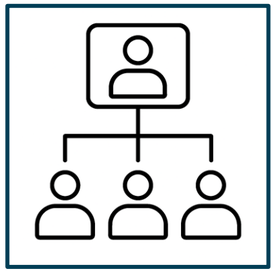
WHY GOVERNANCE MATTERS

- Aligns AI use with mission, ethics, and compliance
- Minimizes harm, bias, and system-level risks
- Builds trust, clarity, and accountability across teams
- BUT a lot goes into governance

THE CORE COMPONENTS OF EFFECTIVE AI GOVERNANCE

AI POLICIES

RESOURCES



STRUCTURE

PROCESSES

CHAI

POLICIES

What:

- AI Policy: Formal documentation outlining scope, principles, and approval processes for AI use.
- Policy Alignment: Ensures consistency with broader legal, regulatory, and internal frameworks.
- Policy Review: Periodic review and revision cycle triggered by new risks or regulations.

How:

- Maintain version-controlled, leadership-approved policies.
- Cross-walk AI policy with privacy, security, and clinical safety policies.
- Schedule annual reviews and log change rationales.

ORGANIZATIONAL STRUCTURE

What:

- Context: Defines governance scope, goals, and readiness.
- Roles & Responsibilities: Assign duties across key domains.
- Concern Reporting: Anonymous AI risk reporting protocols.

How:

- Create a governance charter and appoint cross-functional leads.
- Hold documented governance meetings regularly.
- Establish secure reporting channels and escalation paths.

CHAI

ORGANIZATION RESOURCES

What & How:

- Resource Documentation: Inventory AI tools and systems.
- Data Resources: Document data assets and uses
- Tooling/Computing: What do you have and what do you need?
- Human Resources: Staff roles and competencies.

ORGANIZATIONAL PROCESSES

What:

Define objectives and operationalize processes for:

- Responsible AI system Lifecycle Management and Use
- Risk, Benefit, & Bias Assessment
- Responsible Data Use and Management for AI Systems
- Third-Party Management
- Feedback and Training

PROCESS: RESPONSIBLE LIFECYCLE MANAGEMENT AND USE

What:

- Operationalize the objectives & processes for responsible use (e.g. how would you align with principles of responsible AI?)
- Lifecycle Stages: Define requirements for vendors and organization from development/purchasing to monitoring.
- Logging Events: Monitor for data and system drift or failure.

How:

- Maintain documentation per stage (e.g., validation reports).
- Log key model performance and use events.
- Review logs or set alerts for anomalies.

CHAI

PROCESS: RISK & IMPACT ASSESSMENT

What:

- Risk Categorization: Scores AI system's potential to cause harm.
- Risk Assessment: Evaluate operational, clinical, and ethical risks.
- Bias Profile: Track and mitigate bias risks (e.g. IEEE 7003:2024).

How:

- Assess risk before and after deployment.
- Maintain bias profiles and update throughout AI lifecycle.
- Document assessment results and use in approvals.

Title: Risk Categorization Tool
Risk Assessment Domain: Life and Patient Safety
Primary Audience: Health systems (any size) and teams responsible for pre-deployment risk review.

Risk to Patient Life & Safety

Risk Modifiers: (Low, Med, High)

- Proximity of Impact on Patients
- Human in the Loop
- Consequences of Failure
- Patient Population Vulnerability
- Level of Difficulty Monitoring Solution Output
- Data Transparency
- Solution's Clinical Level of Care
- Time to Intervention (if output is wrong)
- Breadth of Potential Harm
- Integrated Error Propagation Risk
- Population Sensitivity or Disparity Risk

CHAI

Use Case Example: AI-assisted Patient Scheduling software (e.g., scheduling chatbot) - for outpatient, primary care clinics - simplifies booking, rescheduling, and managing patient appointments. These products allow patients to select appointment times and providers to manage their calendars, reduce no-shows, and optimize clinic workflows. Typical features include automated reminders, real-time availability updates, and integration with EHRs. Human confirms appointment once scheduled

Modifier: Distance from Patient
How physically or operationally close is the AI solution to the patient?

Low: No direct impact on individual patient care, supports back-end functions such as back office, administrative tasks, population health analysis, or workflow optimization

Medium: Indirect impact on patient care, access to care, or informational use, such as scheduling, transportation, non-clinical informational chatbots

High: AI solution has a semi-direct impact on patient care, such as, used by a healthcare professional as part of a broader clinical judgment; OR AI solution is directly involved in patient care/patient interaction

CHAI

Modifier: Human in the Loop
The extent to which human oversight is involved in reviewing, verifying, or overriding the AI solution outputs before they affect patient care

Low: AI solution output is always reviewed by a relevant expert before any action is taken

Medium: AI solution output has optional human in loop review by relevant expert before any action is taken

High: AI solution output is never reviewed by provider before an action is taken

CHAI

Modifier: Data Transparency

The clarity, completeness, and accessibility of the data sources and datasets used to train, test, and validate the AI solution.

Low: health system has complete access to training data of the underlying model(s) for the AI solution; lowest level of detail for the data/datasets are shared and available (e.g. AI solution is developed internally)

Medium: partial access to training data of the underlying model(s) for the AI solution; OR some level of detail for the data/datasets are shared and available.

High: no access to training data OR no components of the data/datasets are shared or available (e.g.. Data provenance and data catalog/dictionary unavailable

??

CHAI

PROCESS: THIRD-PARTIES

What:

- Responsibility Allocation: Document **shared** duties and risks.
- Supplier Management: Enforce transparency and auditability.
- Customer Considerations: Ensure usability and trustworthiness.

How:

- Require model cards and testing data from vendors.
- Develop **shared** deployment plans and audit protocols.
- Define clear terms in contracts for incident response and feedback.

CHAI

Tool: Applied Model Card

Applied Model Card Template

Name: **Developer:**		**Inquires or to report an issue:** abc@abc.com or +1 (999) 999- 9999
Release Stage: **Global Availability:**	**Release Date:** **Regulatory Approval, If applicable:**	**Version:**
Summary: **Keywords:**		**Uses and Directions:** • **Intended use and workflow:** • **Primary intended users:** • **How to use:** • **Targeted patient population:** • **Cautioned out-of-scope settings and use cases:**

Warnings

- **Known risks and limitations:**
- **Known biases or ethical considerations:**
- **Clinical risk level:**

Trust Ingredients

AI System Facts:

- **Outcome(s) and output(s):**
- **Model type:**
- **Foundation models used in application, if applicable:**
- **Input data source:**
- **Output/Input data type:**
- **Development data characterization:**
- **Bias mitigation approaches:**
- **Ongoing Maintenance:**
- **Security and compliance environment practices or accreditations, if applicable:**
- **Transparency, Intelligibility, and Accountability mechanisms, if applicable:**

Transparency Information:

- **Funding source of the technical implementation:**
- **3rd Party Information, If Applicable:**
- **Stakeholders consulted during design of intervention (e.g. patients, providers):**

Key Metrics

Completed Example

Tool: Applied Model Card

Key Metrics					
Usefulness, Usability, and Efficacy		**Fairness and Equity**		**Safety and Reliability**	
Goal of metric(s):		**Goal of metric(s):**		**Goal of metric(s):**	
Result:	**Interpretation:**	**Result:**	**Interpretation:**	**Result:**	**Interpretation:**
Test Type:		**Test Type:**		**Test Type:**	
Testing Data Description:		**Testing Data Description:**		**Testing Data Description:**	
Validation Process and Justification:		**Validation Process and Justification:**		**Validation Process and Justification:**	

Resources

- **Evaluation References, If Available:**
- **Clinical Trial, If Available:**
- **Peer Reviewed Publication(s):**
- **Reimbursement status, if applicable:**
- **Patient consent or disclosure required or suggested:**

CHAI

PROCESS: DATA USE & MANAGEMENT

What:

- Data Use & Enhancement: Define training, validation, and deployment uses.
- Data Acquisition & Quality: Prioritize fairness, completeness, validity.
- Data Provenance & Preparation: Document transformations and sources.

How:

- Enforce data governance SOPs and source tracking.
- With third-party vendors: know where your data is going, what form it needs to be in, how it will be used, etc.
- Remember that useful AI needs quality data for training, tuning, validation, and use—data is an ASSET that can bring solutions
- Audit data for missingness and schema issues.

CHAI

HIPAA Primer

- What can you do with PHI?
  - Under HIPAA can disclose PHI for institution's own Treatment, Payment, and Operations (TPO)
    - Example: proper disclosure of PHI to develop tools to assist in patient diagnosis or improving population health likely fit within confines of HIPAA
  - Research activities are different!
    - Have to be disclosed for generalizable knowledge, which could include providing data for development of medical devices or pharmaceuticals
    - Require authorization of the individual patient to do unless IRB Waiver or limited use data set as agreed to in data use agreement
- Once data is de-identified or is considered "synthetic", HIPAA no longer applies.
  - Data de-identification needs to be done by expert determination or by removing 18 key identifiers.
  - AI can sometimes be used to re-identify individuals even when proper de-identification has taken place, so stress testing this is important

CHAI

PRINCIPLES FOR EFFECTIVE DATA MANAGEMENT & PRIVACY

Know where your data is and where it’s going

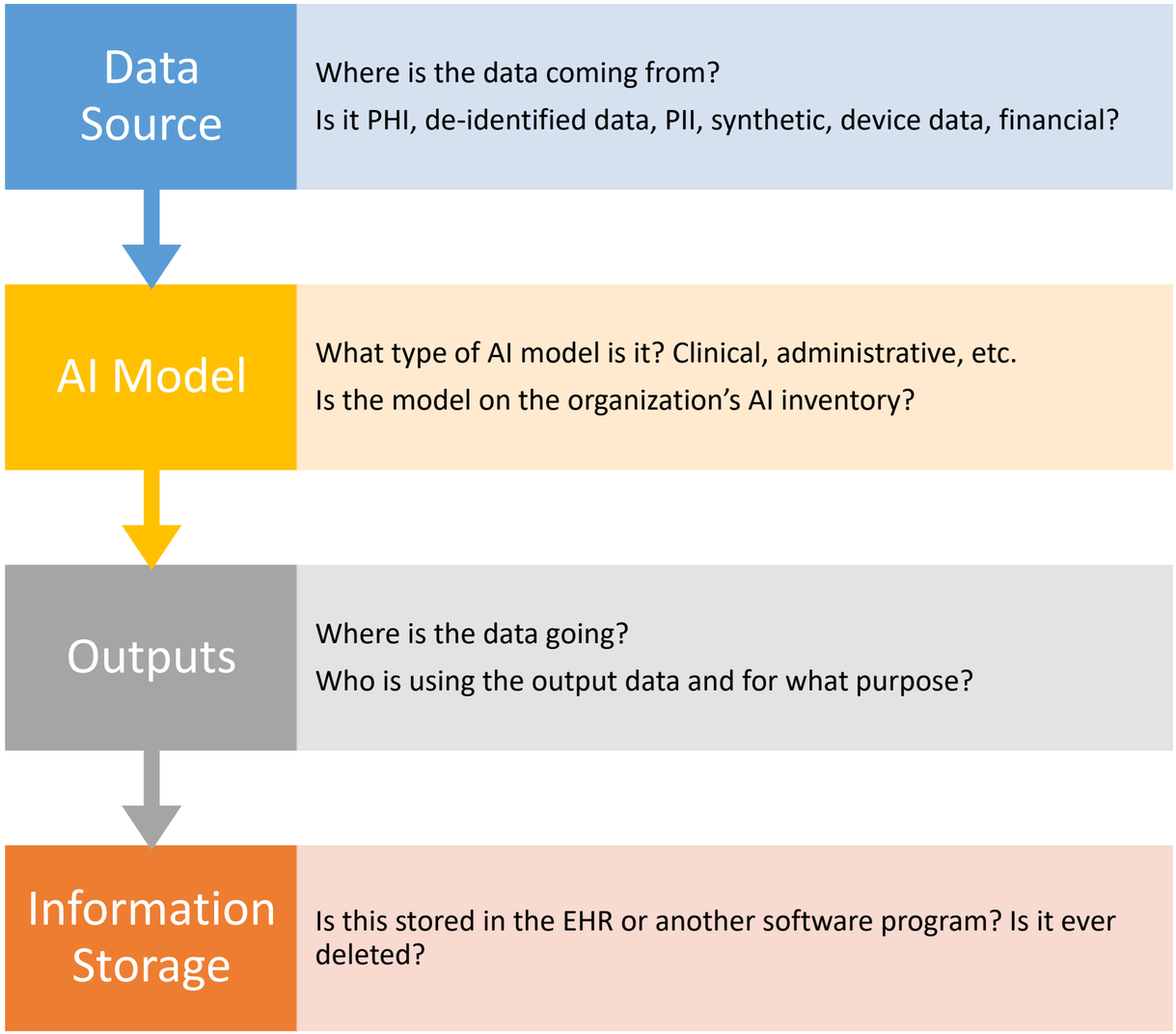
Implement proper technical safeguards

Work with the right vendors

Repeatable Processes

CHAI

MAP DATA & ASSESS RISK



HIPAA security risk analysis (45 CFR 164.308(a)(1))

- Regular risk assessments (confidentiality, integrity, etc.)
- Ensure appropriate access (who within your organization actually needs access to PHI?)
- Sanction policy (for those who inappropriately access PHI)
- IT system review and audits
- Yearly training

CHAI

Guidance for data management & use outside of HIPAA

- Encryption in transit and at rest when possible
- Access controls
- Regular security assessments
- Incident response plan
- Effective Data Use Agreements:
  - Permitted uses
  - Data minimization
  - Explicit Prohibition of reidentification
  - Third-party obligations that align with internal policies (e.g. encryption, regular security assessments, etc)
  - Audit rights

Upcoming tool: Technology & Data Risk Tool

CHAI

PROCESS: FEEDBACK & TRAINING

What:

Define and document:

- Processes for user feedback and incident reporting
- Key AI-related incident types
- Reporting pathways and responsible parties
- Processes for how users are trained

How:

- Update existing clinical and operational safety reporting processes
- Ensure the right people know the right information about the AI system
- Train relevant staff on how to use and monitor the AI solution effectively
- Consider broader AI Literacy education and change management

CHAI

GETTING STARTED

- Draft or refine an AI policy
- Form a governance group across compliance, IT, clinical, care management, patient advocate, human factors, and admin leads.
- Define a risk-based governance process for AI solutions
- Start a basic registry of existing AI solutions and readiness inventory.
- Pilot a governance process on a single low-risk AI use case.
- Don’t reinvent the wheel. Find existing processes and adapt them.

THIS IS A BIG UNDERTAKING

Currently:

1. Released guidance with Joint Commission focused on Responsible Use of AI.

2. Held governance workshops to develop playbooks, tools, templates, and resources that can help healthcare organizations of all sizes and resourcing align with AI governance processes and procedures. (First playbooks will be released EOY 2025)

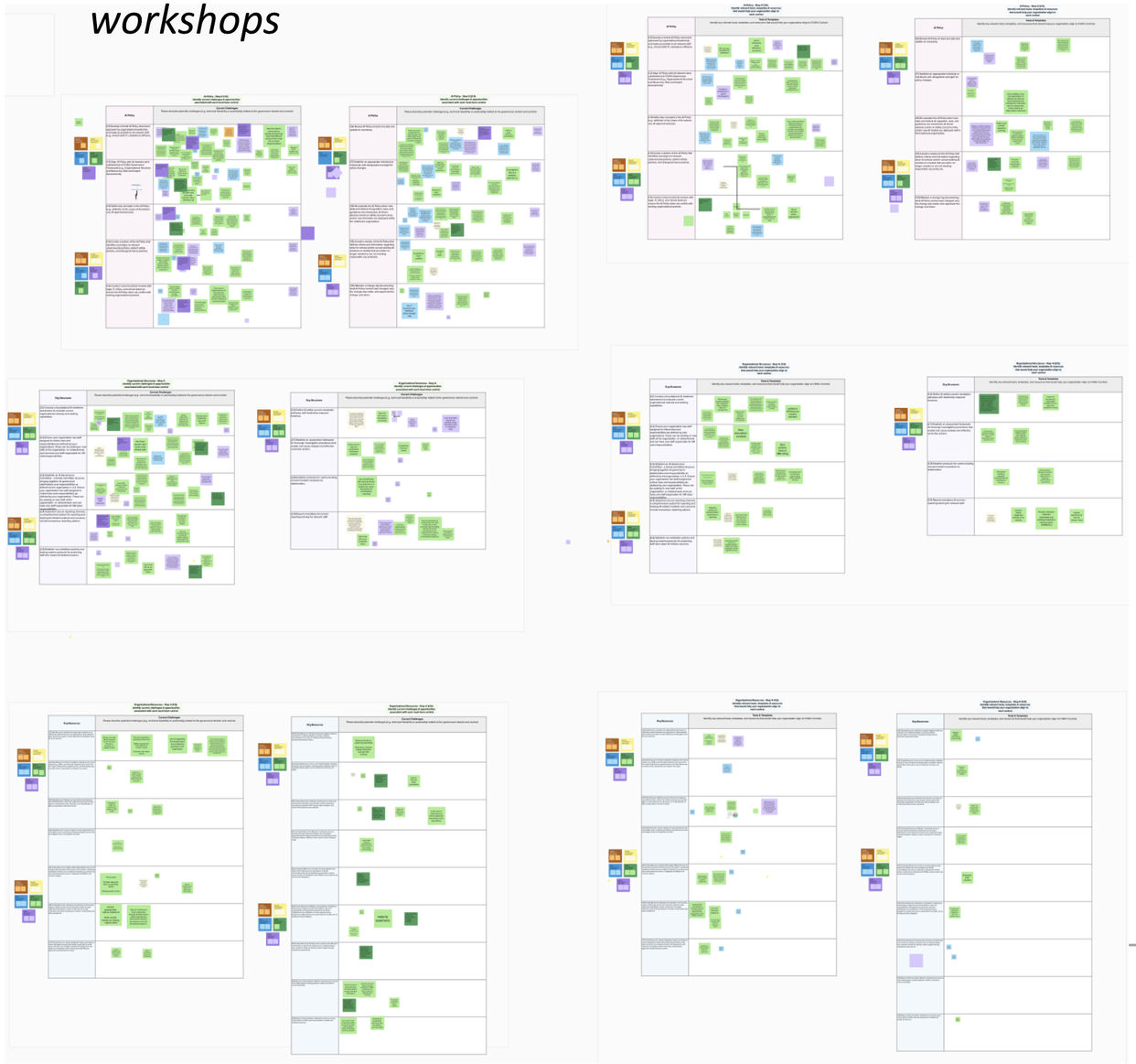
Upcoming:

- Survey to contribute to playbook content
- AI governance playbooks, with adaptations for FQHCs and their networks
- Voluntary Joint Commission Certification (not accreditation) on Responsible Use of AI based on playbooks.

CHAI

Workshop Process

MURAL Board Activity with Members: 3 x 3-hour workshops



- Which governance and responsible use controls are nice to have vs. must have?
- What challenges do you face in aligning with these controls?
- What kinds of tools, templates, or resources would help reduce some of the challenges?
- How does your organization align with some of these controls?

Insights: Workshop 1

- Policy:
  - Need for continuous tracking and updating of policies and procedures given changing landscape.
  - Need to clarify scope of what should be governed
  - Need templates, risk-based methods, continuous regulatory and policy updates.
- Organizational Structures
  - Need for upskilling and hiring resources
  - Need responsibilities to be actionable and not “just because”
  - Need trusted reporting guidance and clear escalation pathways
- Organizational Resources:
  - Need to understand the cost of infrastructure and solutions for legacy systems.
  - Need registries and model card tracking to be less resource intensive
  - Need benchmarks, checklists, and centralized tools

Governance is essential, but challenged by pace, roles, and resources. Practical tools and collaboration are key enablers.

CHAI

Insights: Workshop 2

- Lifecycle Management & Responsible Use
  - Monitoring should be risk-based, and appropriate to the solution's context to limit resource intensity (not all solutions/contexts need continuous monitoring)
  - Need clearer understanding of which metrics should be monitored and consistency. **Tool:** Testing & Evaluation Frameworks
  - Need simplified vendor oversight/logging
  - Need improved safety reporting infrastructure
  - Need dashboards, logging schemas, incident reporting hubs, and standardized templates.
- Data management and Use
  - Need clarity on what constitutes "derivative works"
  - Need more vendor accepted data use audit procedures
  - Need ways to streamline maintenance of data registers, lineage, and quality checks
  - Need more robust DUAs, minimum necessary practices, and validation pipelines

Organizations struggle with monitoring, vendor oversight, and data governance complexity. Standardized tools, common AI DUA components, and governance teams are critical enablers.

CHAI

Insights: Workshop 3

- Risk.& Impact Assessment
  - Need more consistent understanding of risk
  - Need simpler and legally protected bias profiling
  - Need templates, phased rollouts, and integration with existing review systems.
- Third-Party Management
  - Need shared liability and transparency from vendors
  - Need AI-specific clauses for contracts
  - Need improved power/resource balance in negotiations
  - Need standardized model cards, contract language, risk-based audits, and vendor-health system co-authored deployment plans
- Feedback and Training
  - Need better integration of training and reporting mechanisms with existing systems
  - Need standardized incident reporting and more vendor alignment
  - Need guidance on how and when to approach patient consent and communication
  - Need modular training, multi-access reporting, whistleblower protections, and AI-specific incident fields in existing systems.

Organizations face resource strain, vendor resistance, and reporting/training complexity. Progress depends on standardized tools, clear roles, mechanisms for shared responsibility, and patient/stakeholder trust mechanisms.

CHAI

About Coalition for Health AI (CHAI)

The CHAI (Coalition for Health AI) mission is to be the trusted source of guidelines for responsible AI in health that serves all. It aims to ensure high-quality care, foster trust among users, and meet the growing healthcare needs. CHAI membership is open and rapidly expanding with nearly 3000 organizations including health systems, patient advocacy groups, and a wide range of industry leaders and start-ups across the healthcare and technology ecosystems. CHAI is committed to convening and dialogue to achieve consensus. There is no limit to who can join and participate. Learn more about a CHAI membership here.

To learn more about CHAI or to inquire about membership:

www.chai.org

admin@chai.org

Questions about this presentation?

merage@chai.org

CHAI