

# Implementing AI: A Practical Playbook

Data to Decisions: AI in Healthcare Course

R. Andrew Taylor MD MHS



**Idea**



**Pilot**



**Workflow**



**Monitoring**



**Scale**

# Objectives and How to Use



Mental model: idea → safe impact



Step-by-step implementation playbook



MLOps = ongoing patient-safety monitoring



Use on rotations: ask “where, who, how measured?”

## Playbook Card



1. Define Problem & Goal



2. Pilot & Validate



3. Integrate Workflow



4. Monitor Safety (MLOps)



5. Scale & Sustain

**Questions to Ask:**

Where, who, how measured?

# Overview

The background of the slide features a blue-to-yellow gradient. Faint, stylized circuit board patterns are visible in the corners, with a prominent chip-like shape in the top right and bottom right.



# AI in Healthcare: Taxonomy



## Predictive

estimate risk or outcomes

Risk score



## Generative

summarize, draft, explain

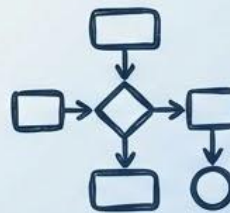
Note summary



## Automation

execute constrained tasks

Protocol automation



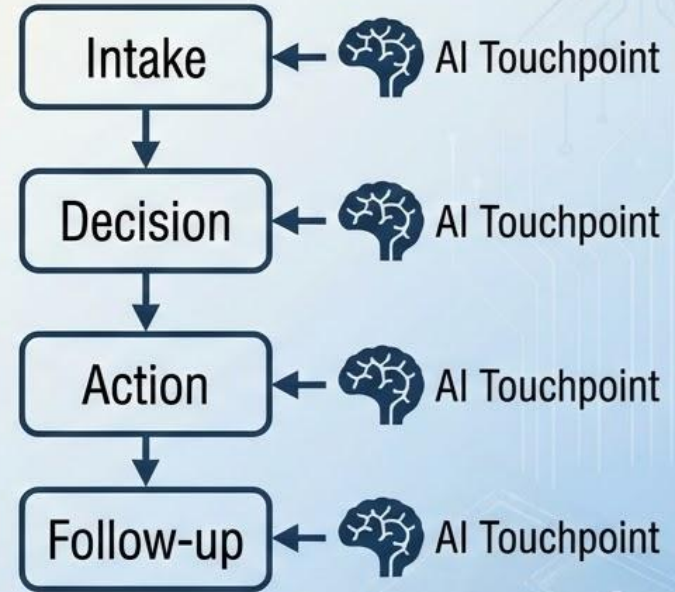
## Optimization

allocate resources,  
schedule flow

Throughput optimizer

# Where AI Fits in Workflow

- Workflow map: intake → decision → action → follow-up
- Place AI where it can change actions
- Minimize extra clicks and interruptions



**Pause & Think:** Choose one touchpoint: intake, decision, action, follow-up—why?

# Why Pilots Fail: Top Causes

- Workflow mismatch + alert fatigue
- Data issues: missingness, drift, labels
- No ownership or governance
- Weak evaluation + unclear success metrics
- Incentives misaligned with operations

## Top 8 failure modes



Mismatched  
Workflow



Alert  
Fatigue



Data Gaps



No Owner



Data Drift



Label Errors



Weak Eval



Misaligned  
Incentives

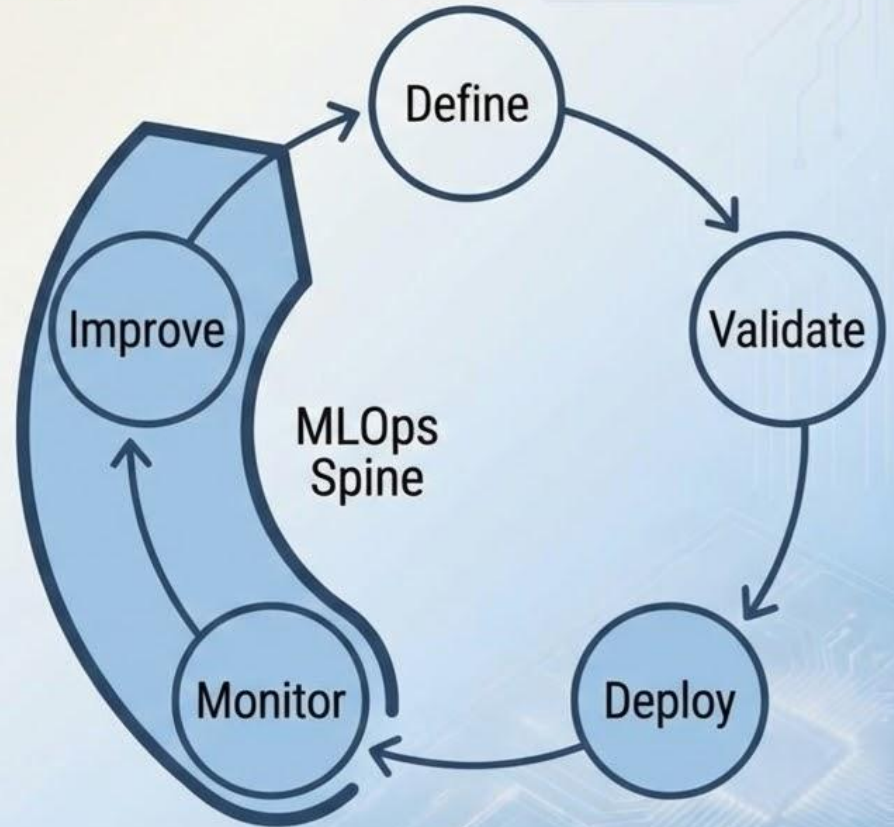


The background of the slide features a light blue and yellow gradient. It is decorated with faint, stylized circuit board patterns and microchip icons in the corners. The main title is centered in a large, bold, dark blue font.

# **AI Implementation**

# The Implementation Lifecycle Loop

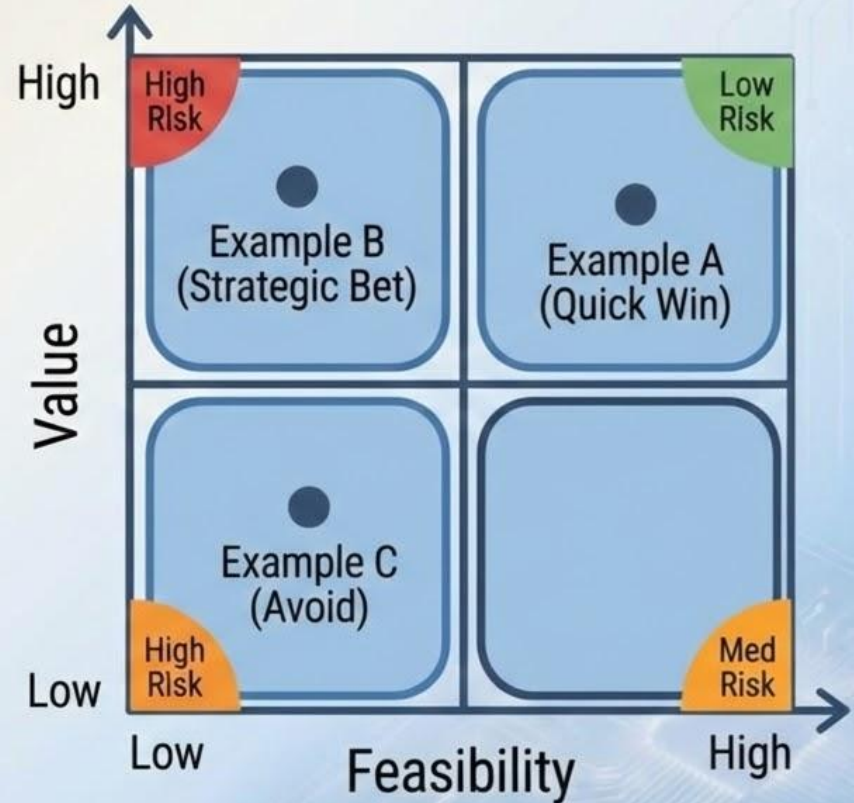
- Define: decision, action, owners, harms
- Validate: technical + clinical performance
- Deploy: workflow + training + controls
- Monitor: drift, safety, equity, usage
- Improve: iterate, retrain, retire





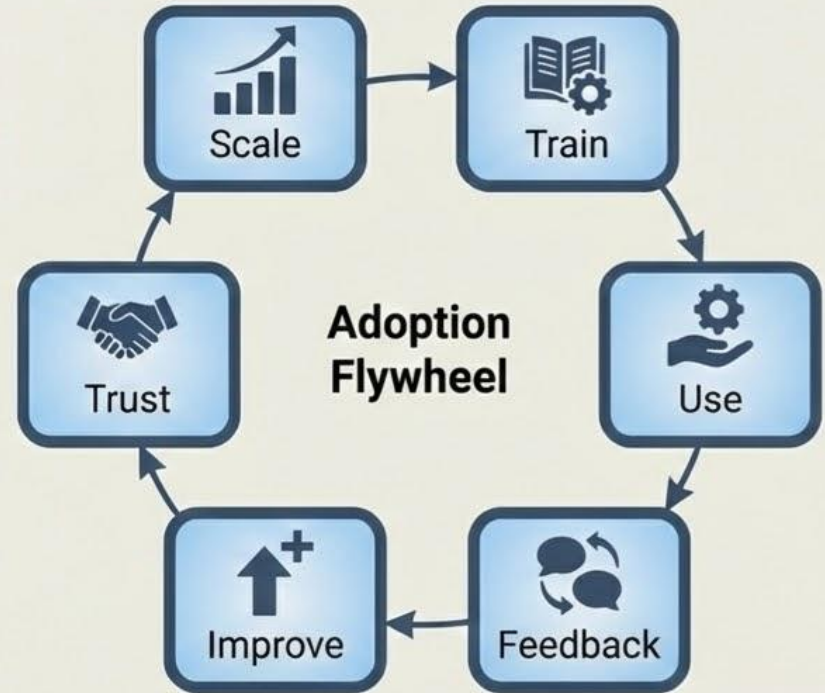
# Choosing Use Cases: Value vs Feasibility

- Score: value, feasibility, and risk
- Prefer “high value, low burden” starts
- Define the action pathway before building
- Document why you said “no”



# Adoption and Change Management

- Identify champions and accountable owners
- Train for “why,” not just “how”
- Feedback loops: rapid, visible iterations
- Plan de-implementation of failed tools



# Data Readiness: The Checklist

- Quality: accuracy, completeness, timeliness
- Provenance: source, transformations, versioning
- Labels: definitions, bias, adjudication
- Missingness patterns can be informative
- Document known bias risks

☒ Quality

☒ Provenance

☒ Labels

☒ Missingness

☒ Document risks





# Governance and Accountability (RACI)

- Assign: clinical owner + accountable leader
- Separate: build, validate, approve, monitor
- Include: IT, compliance, QI, safety, operations
- Decide escalation and rollback authority

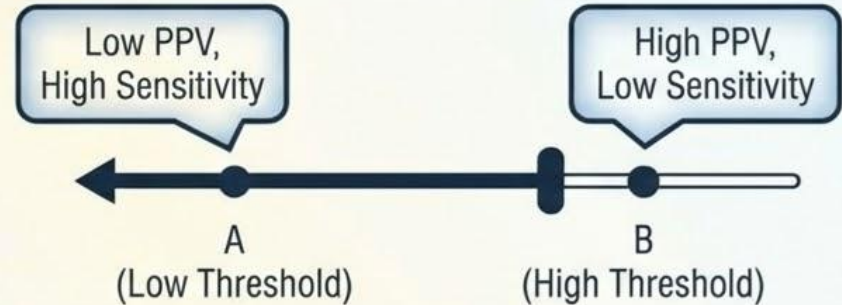
	Clinical Owner	DS/ML	IT	Compliance	QI/Safety
Define	R	C	I	I	C
Validate	A	R	I	C	C
Deploy	C	C	R	I	I
Monitor	A	R	C	C	C
Update	R	A	C	I	C

R: Responsible, A: Accountable, C: Consulted, I: Informed

# Clinical Evaluation: Beyond AUROC

- Calibration: predicted risk matches reality
- PPV/NPV at chosen thresholds
- Decision-curve: net clinical benefit
- Subgroup performance + fairness checks
- Pause & Think: Which threshold?

	Predicted Positive	Predicted Negative
Actual Positive	1 Yes	2 No
Actual Negative	3 Pos	4 Nega



For scarce resources, pick A or B—why?

# Decision Support Design Choices

- Alerts: interruptive, high risk of fatigue
- Inline recommendations: “right place, right time”
- Automation: only with guardrails + audit trails
- Always include human override + escalation



## Alerts

**Benefit:** High visibility, forces attention

**Risk:** Interruptive, high risk of fatigue



## Inline Recommendations

**Benefit:** “Right place, right time” guidance

**Risk:** Can be overlooked if subtle



## Automation

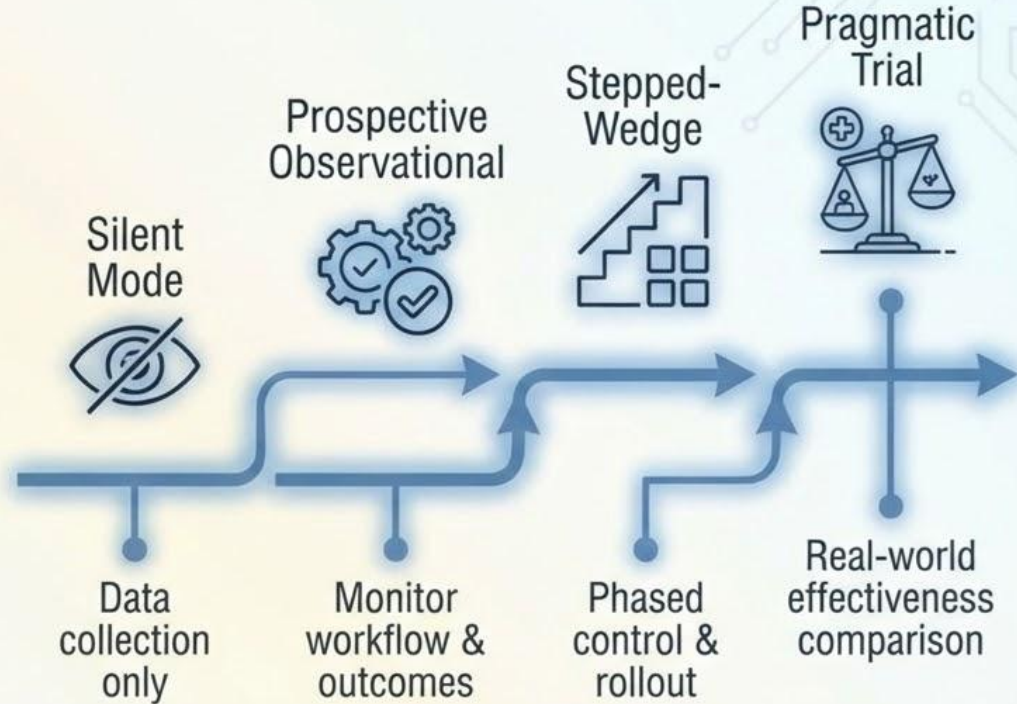
**Benefit:** Efficiency, reduced cognitive load

**Risk:** Requires guardrails + audit trails



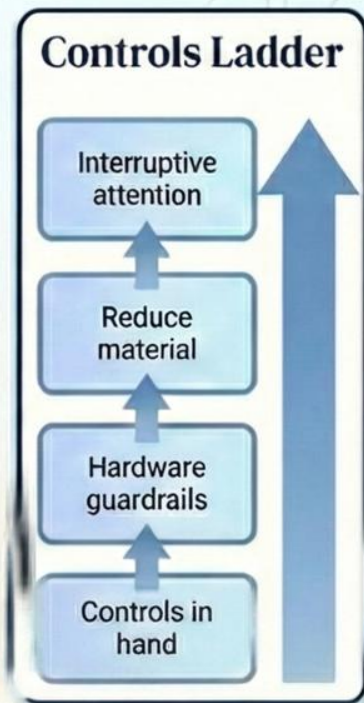
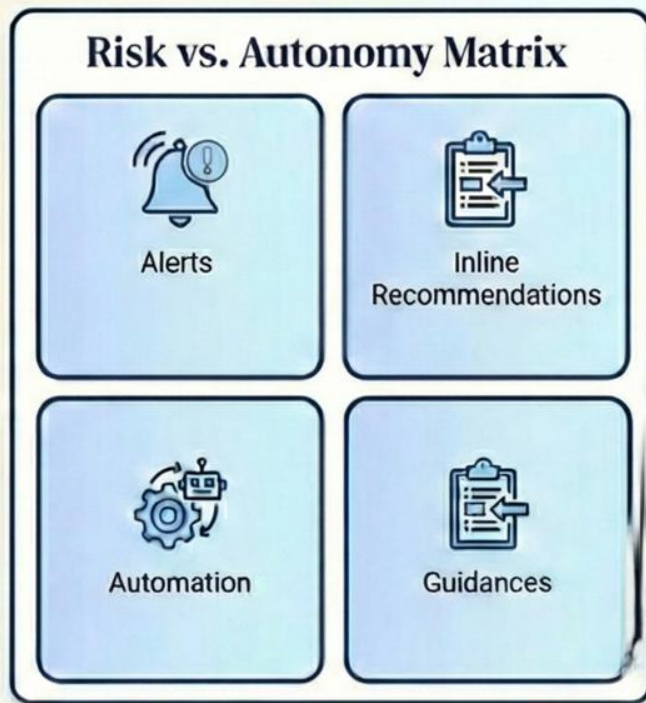
# Real-World Evidence: Study Designs

- Silent mode: measure before showing clinicians
- Prospective observational: workflow + outcomes
- Stepped-wedge: phased rollout with controls
- Pragmatic trial: real-world effectiveness



# Safety Controls: Risk Ladder

- Alerts: interruptive, high risk of fatigue
- Inline recommendations: “right place, right time”
- Automation: only with guardrails + audit trails
- Always include human override + escalation



# Human Factors: Workflow Integration



Map “before vs after” workflow



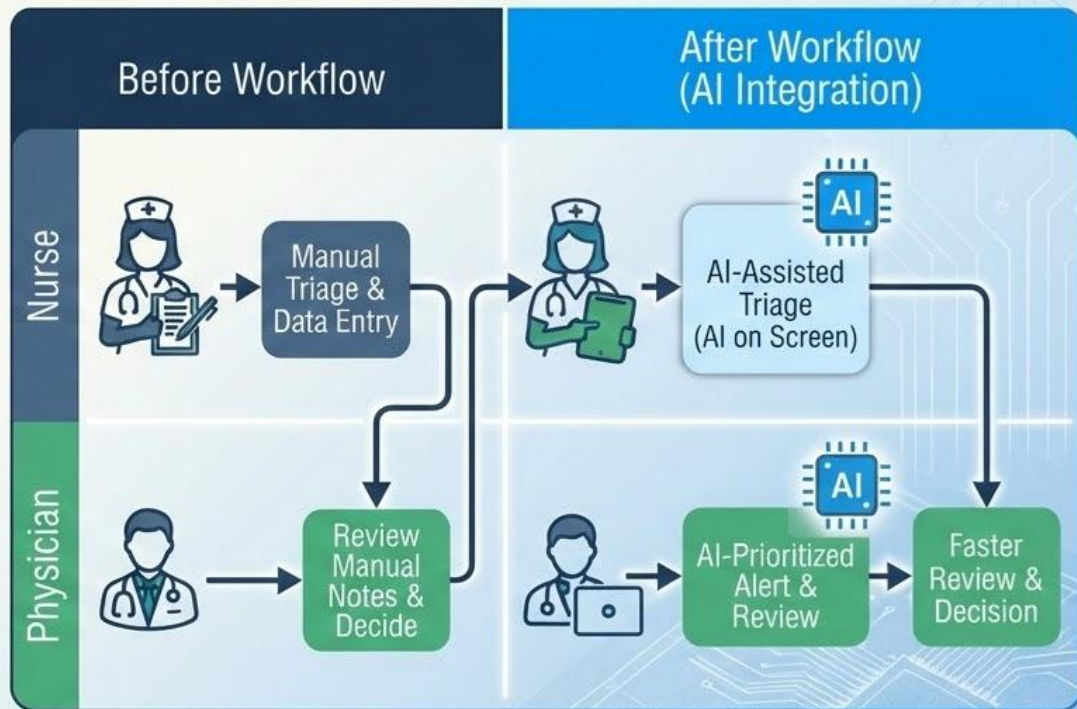
Minimize cognitive load and interruptions



Put AI in the clinician's line-of-sight







Define override and escalation paths

























# Equity and Bias in Deployment

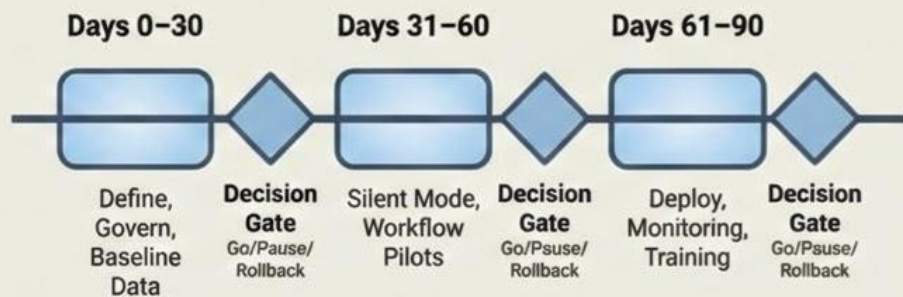
-  Bias enters via data, labels, access, workflows
-  Monitor subgroup performance continuously
-  Use mitigation playbook (data, thresholds, workflows)
-  Engage affected communities early

**Bias Entry Points vs. Mitigations Matrix**

		Data	Labeling	Workflow	Access
Mitigations	Relabeling	 Audit & Correct	 Adjust Labels	 Review Steps	 Update Policy
	Reweighting	 Rebalance Data	 Modify Weights	 Prioritize Steps	 Equitable Access
	Thresholding	 Set Fairness Metrics	 Define Criteria	 Apply Guardrails	 Limit Exposure
	Workflow Redesign	 Redesign Process	 Simplify Tasks	 Introduce Checks	 Streamline Access
	Outreach	 Collect Feedback	 Community Input	 Stakeholder Review	 Broad Engagement

# The 90-Day Implementation Playbook

- Days 0–30: define, govern, baseline data
- Days 31–60: silent mode + workflow pilots
- Days 61–90: deploy + monitoring + training
- Decision gates: go/pause/rollback



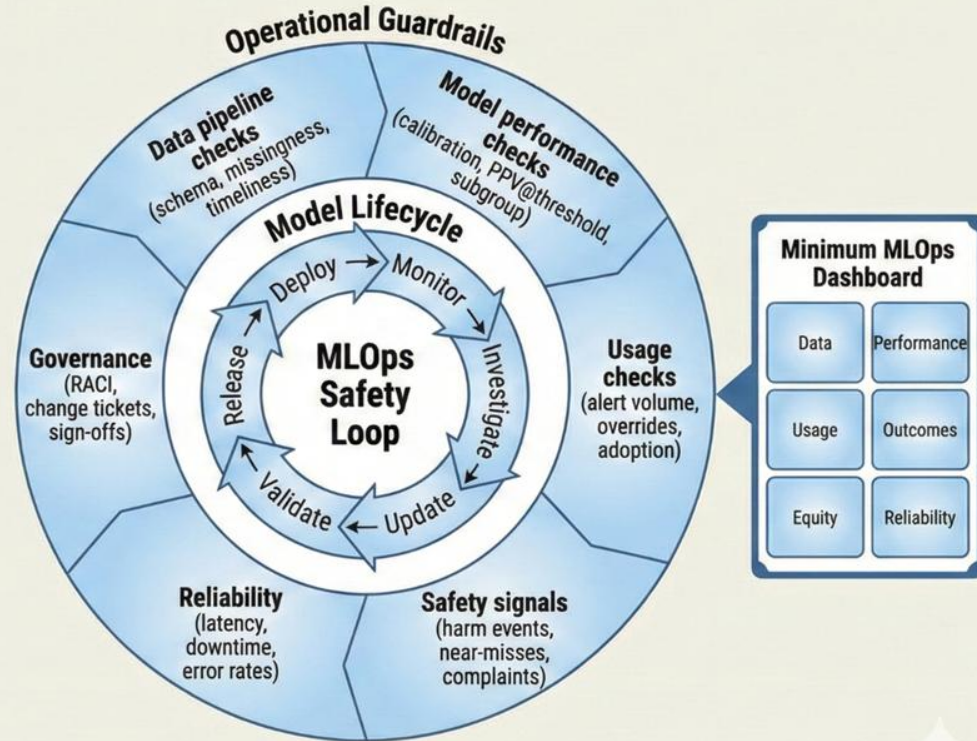
# MLOps





# MLOps: Clinical Model Operations for Safety

- MLOps = keep models safe, reliable, and current
- Monitor: data drift, performance, usage, equity
- Control change: versioning, approvals, rollback
- Operate pipelines: uptime, latency, auditability
- Treat updates like clinical changes



# Monitoring: What to Watch

- Model: performance drift + calibration shift
- Data: missingness, coding, feature distribution drift
- Use: adoption, overrides, alert burden
- Outcomes: safety signals + equity metrics
- Pause & Think: Drift suspected—now what?

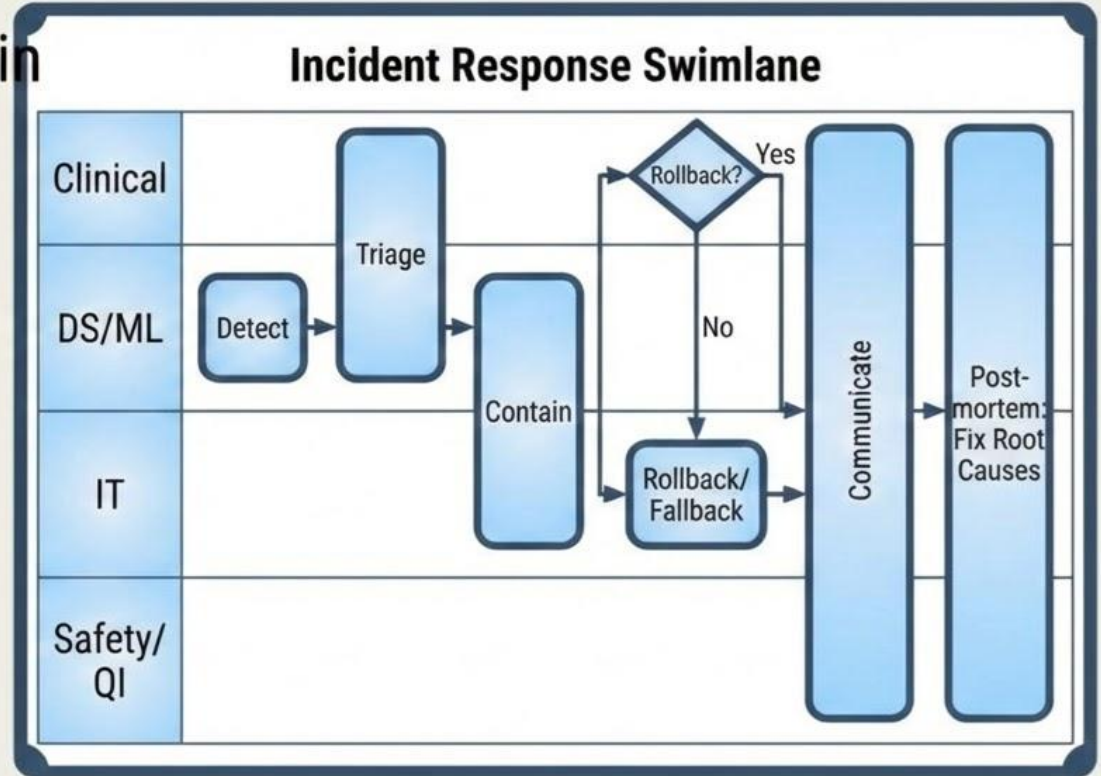
**Monitoring Matrix**

	Leading indicators	Lagging indicators
Model	Performance drift, Calibration shift	Error Analysis, Subgroup Check
Data	Missingness, Coding, Feature distribution drift	Data Integrity Audit, Source Review
Use	Adoption, Overrides, Alert burden	Workflow Analysis, User Feedback
Outcomes	Near Misses, Process Measures	Safety signals, Equity metrics
Equity	Bias Detection, Fairness Assessment	Disparity Analysis, Long-term Impact

**What is your first step?**

# AI Incident Response Playbook

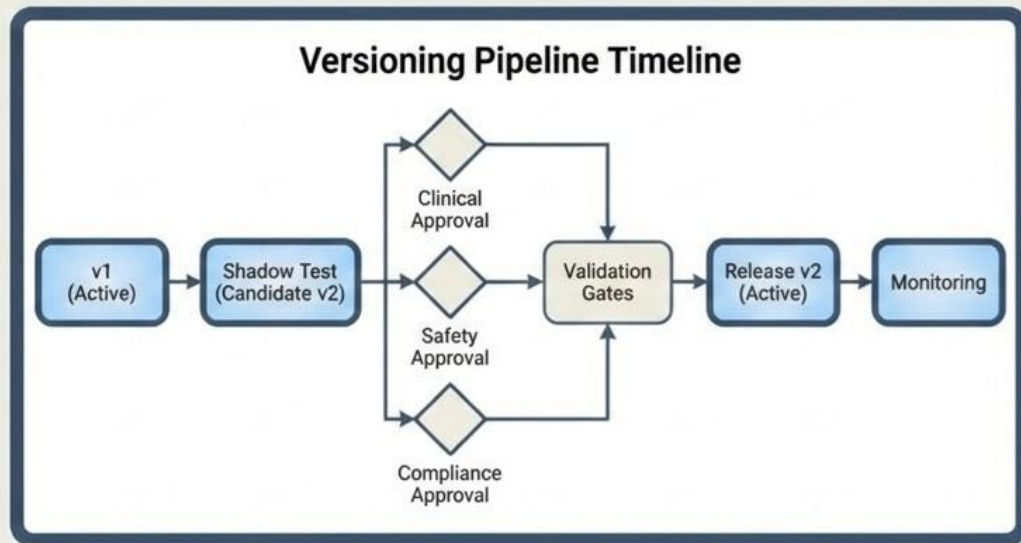
- Detect → triage → contain → communicate
- Rollback and fallback must be practiced
- Document decisions and impact
- Postmortem: fix root causes





# Retraining and Versioning With Gates

- Triggers: drift, new data, new workflow
- Validate in shadow mode before release
- Approvals: clinical + safety + compliance
- Maintain version history and changelogs



# Measuring Value: The Metric Stack

- Patient outcomes + safety outcomes
- Clinician experience (burden, trust, time)
- Operations (throughput, capacity, delays)
- Cost and unintended consequences
- Leading vs lagging metrics



# Your Rotation Rubric + Call to Action

- Rubric: Safe, Useful, Adopted, Maintained
- 3 questions to ask any AI tool
- Spot one workflow opportunity on rotations
- Partner with clinicians, safety, and IT early

