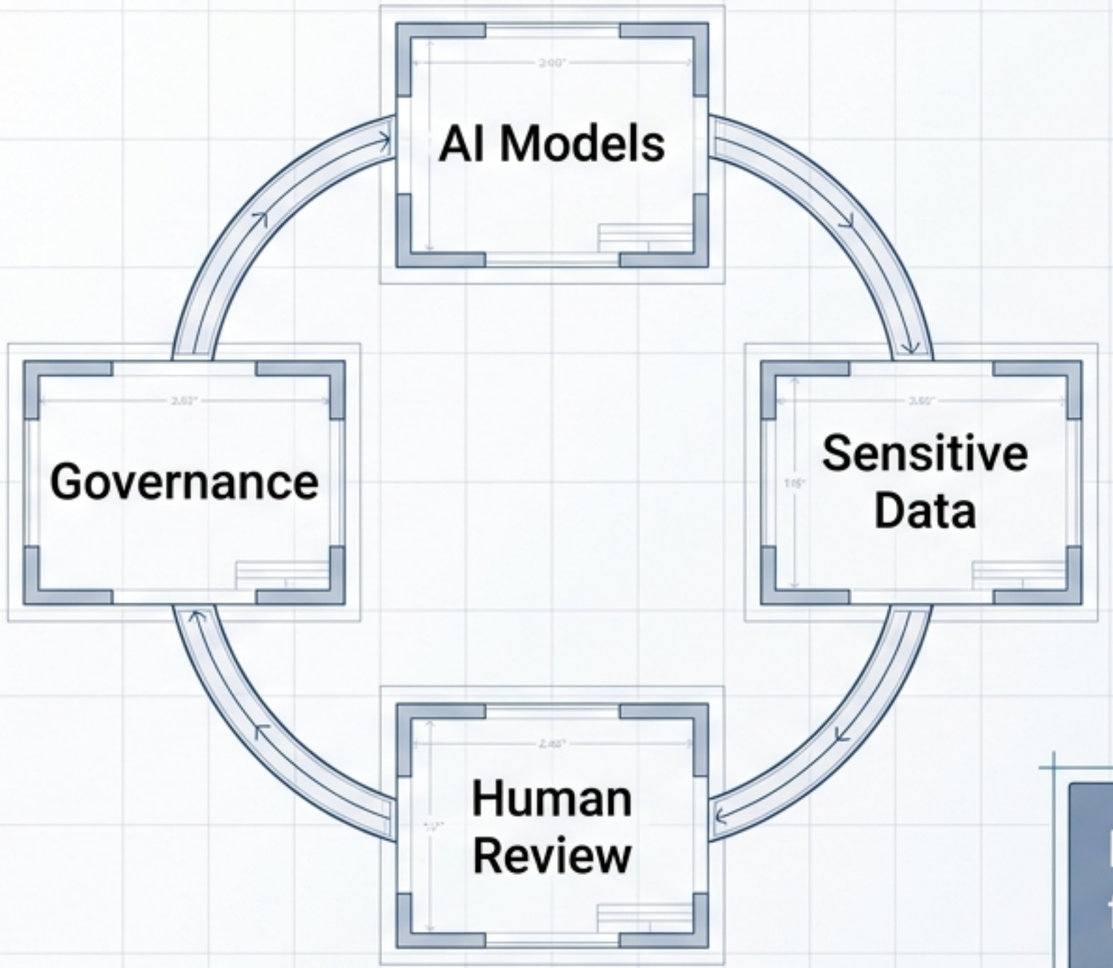
Responsible AI & Risk Review

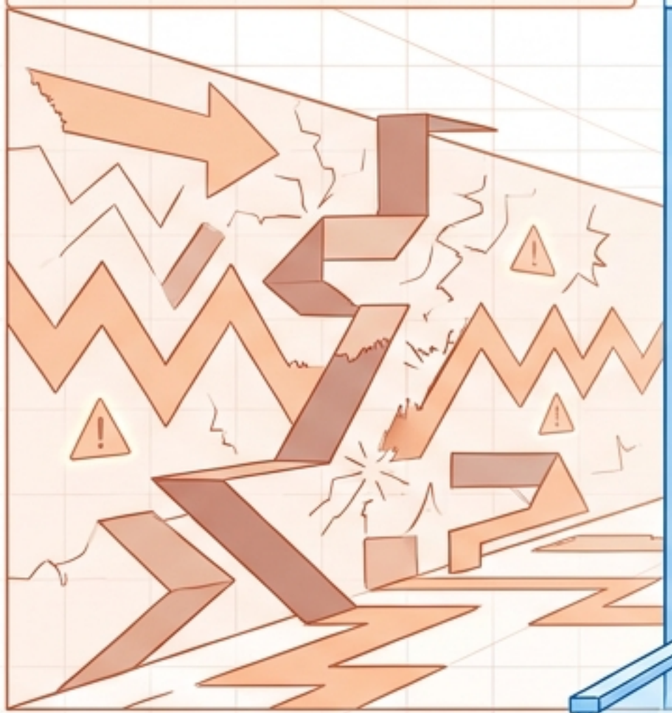
Building safe, practical design habits for healthcare AI adoption.



Human review is a design feature, not a failure.

Governance Should Enable Good Work

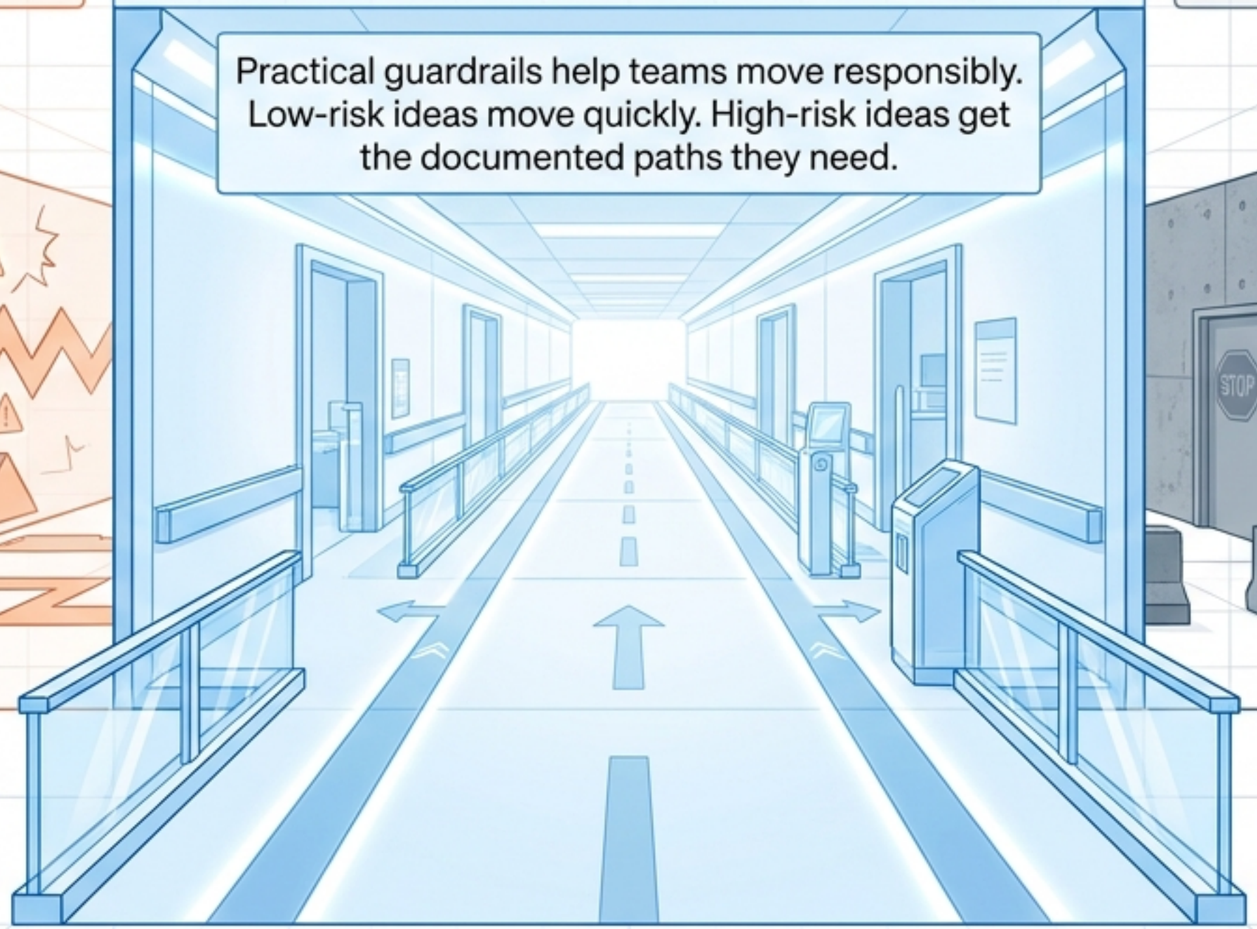
Too Loose:
Ignore Risk



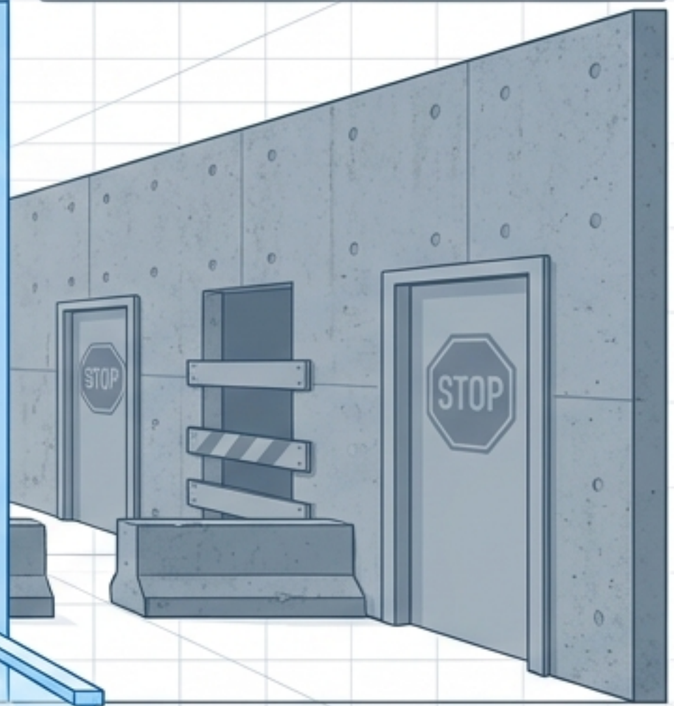
Ignoring risk exposes patients, compliance, and trust.

The Sweet Spot:
Enablement & Responsible Design

Practical guardrails help teams move responsibly. Low-risk ideas move quickly. High-risk ideas get the documented paths they need.



Too Strict:
Stop Everything



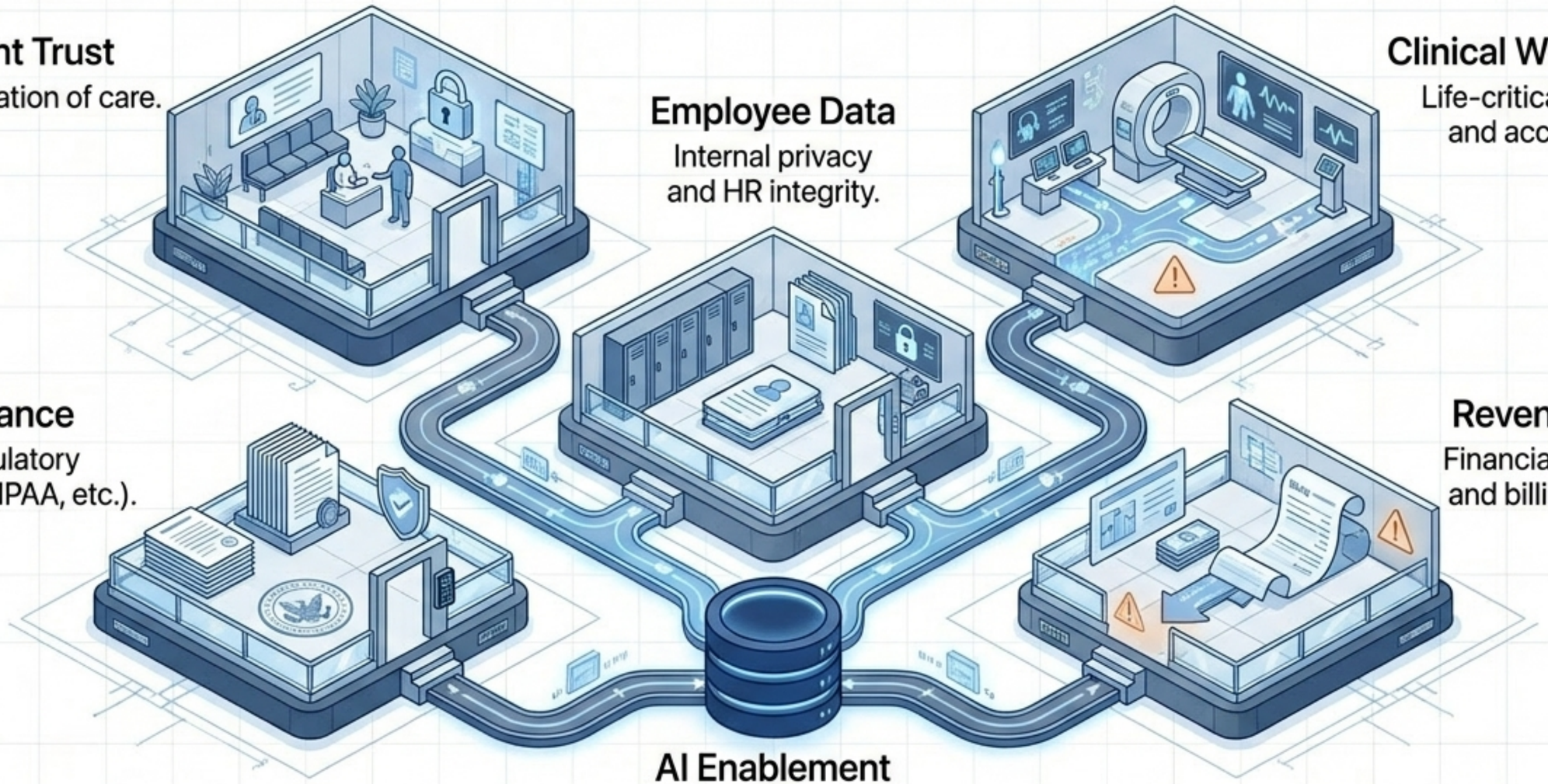
Stopping everything kills innovation and burns out teams.

Why AI Risk Matters Specifically in Healthcare

Patient Trust
The foundation of care.

Employee Data
Internal privacy and HR integrity.

Clinical Workflows
Life-critical speed and accuracy.



Compliance
Strict regulatory obligations (HIPAA, etc.).

Revenue Cycle
Financial operations and billing integrity.

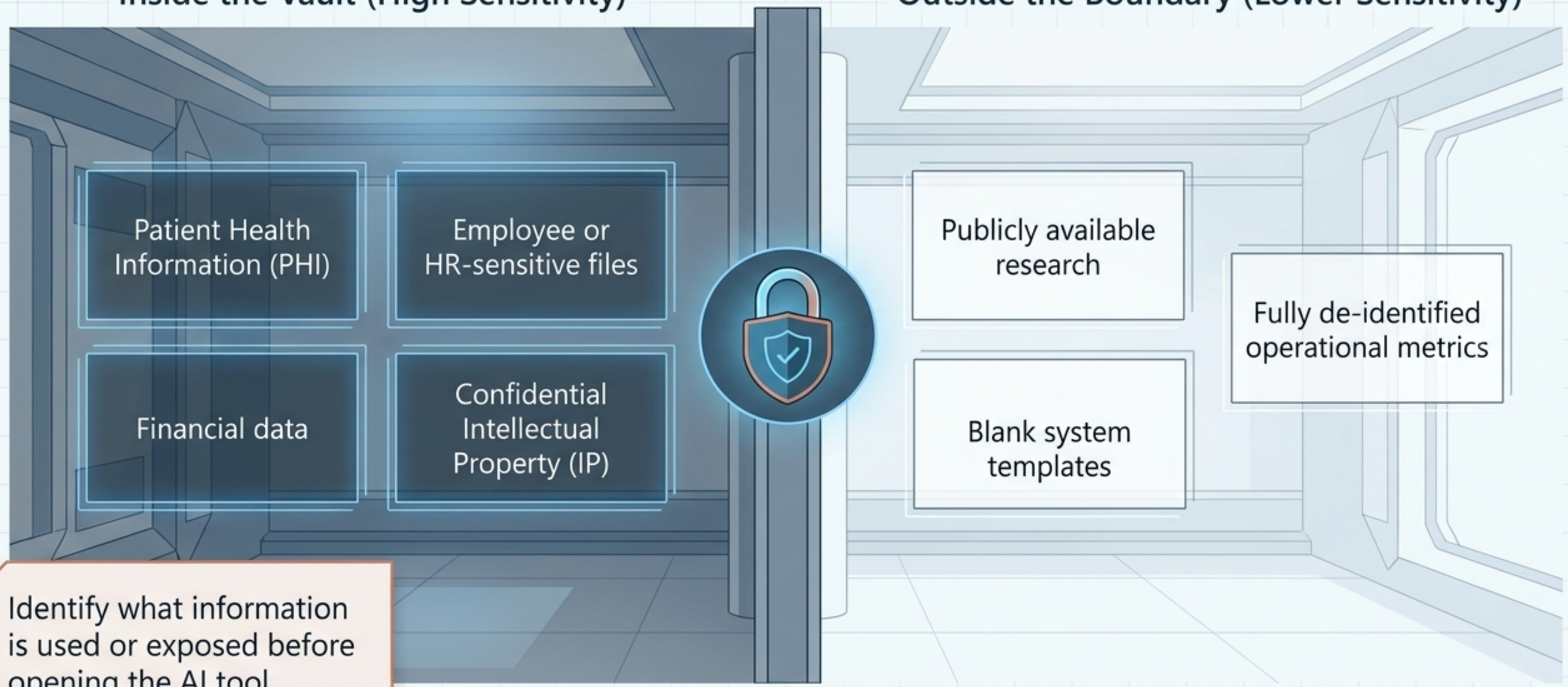
AI Enablement

AI enablement requires clear boundaries to protect our most critical ecosystem dependencies.

The Baseline Diagnostic

	Does this involve patient information?		Can a human review the output?
	Does this involve employee information?		Can the workflow be audited?
	Does this affect care, billing, legal, compliance, employment, or finance?		Can the solution be turned off or rolled back?
	Is the data source approved?		Who owns the process?

Identify Sensitive Data Before Using AI



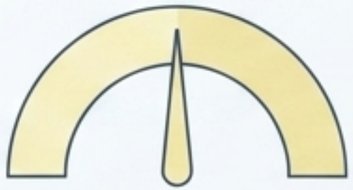
Identify what information is used or exposed before opening the AI tool.

Impact Increases Based on Where Outputs Land

AI Output

High Impact

Direct patient-specific care, financial decisions, legal/compliance reporting, automated employment actions.



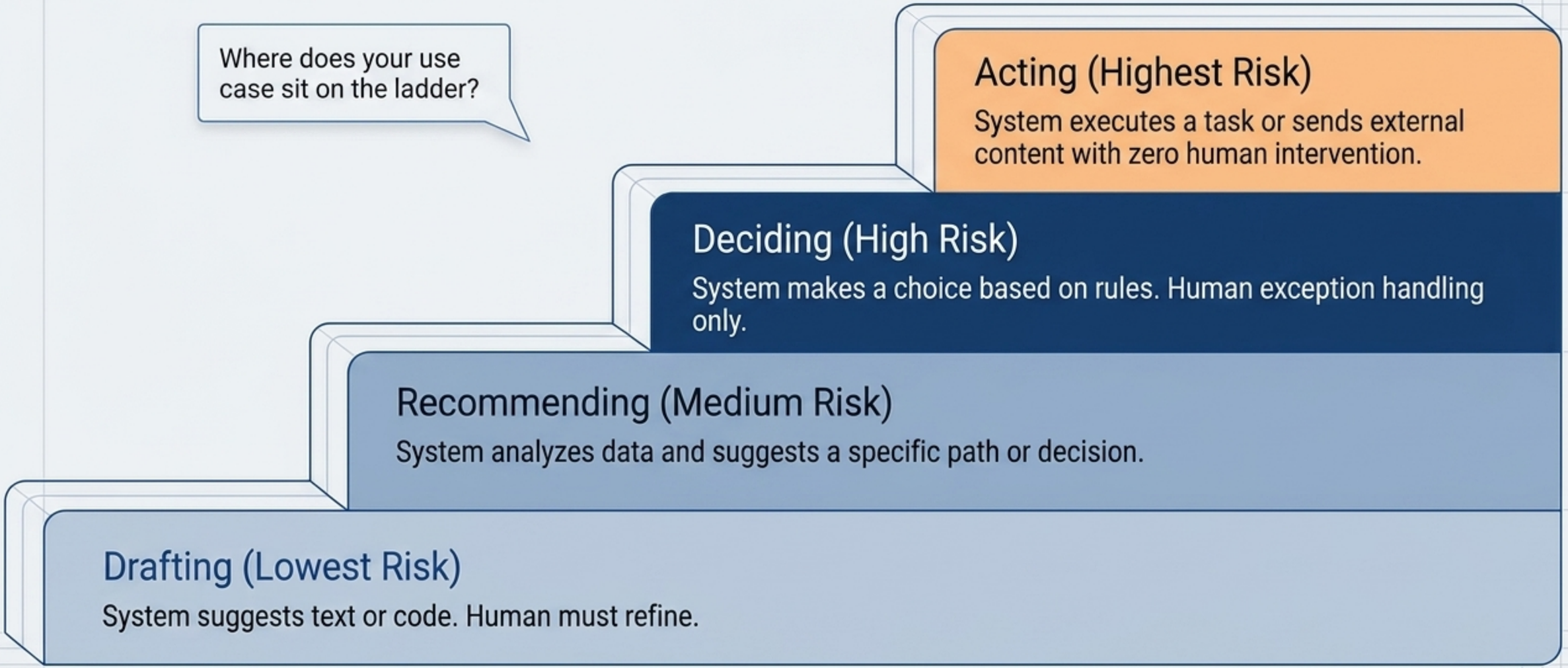
Medium Impact

Internal operational dashboards, draft communications to patients (requiring review).

Low Impact

Brainstorming, drafting routine internal-only emails, summarizing public articles.

AI Autonomy and the Gradient of Risk



Human Review Must Be Specific and Precede Impact

Mandatory Pause

Phase 1: **Generation**

AI drafts clinical note, code, or financial summary.

Phase 2: **Execution**

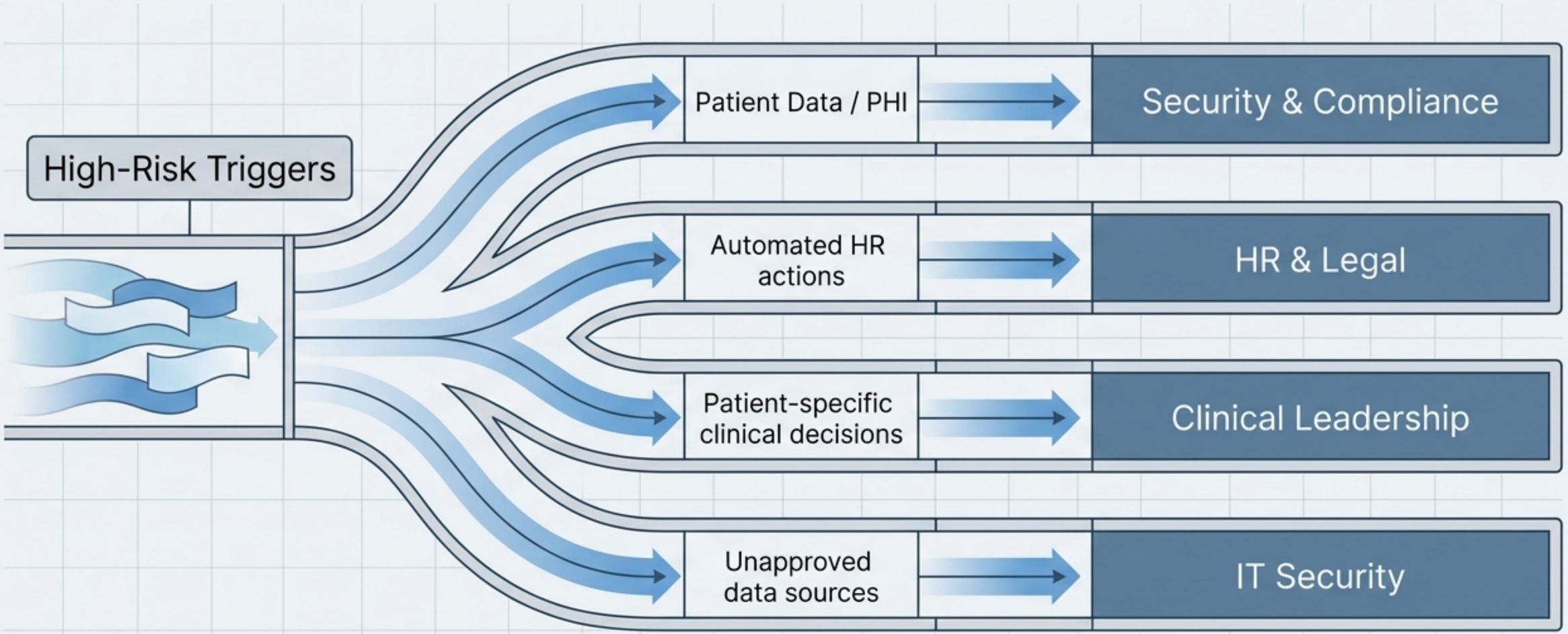
External send, clinical entry, or financial action.

Who reviews?

What exactly are they checking for?

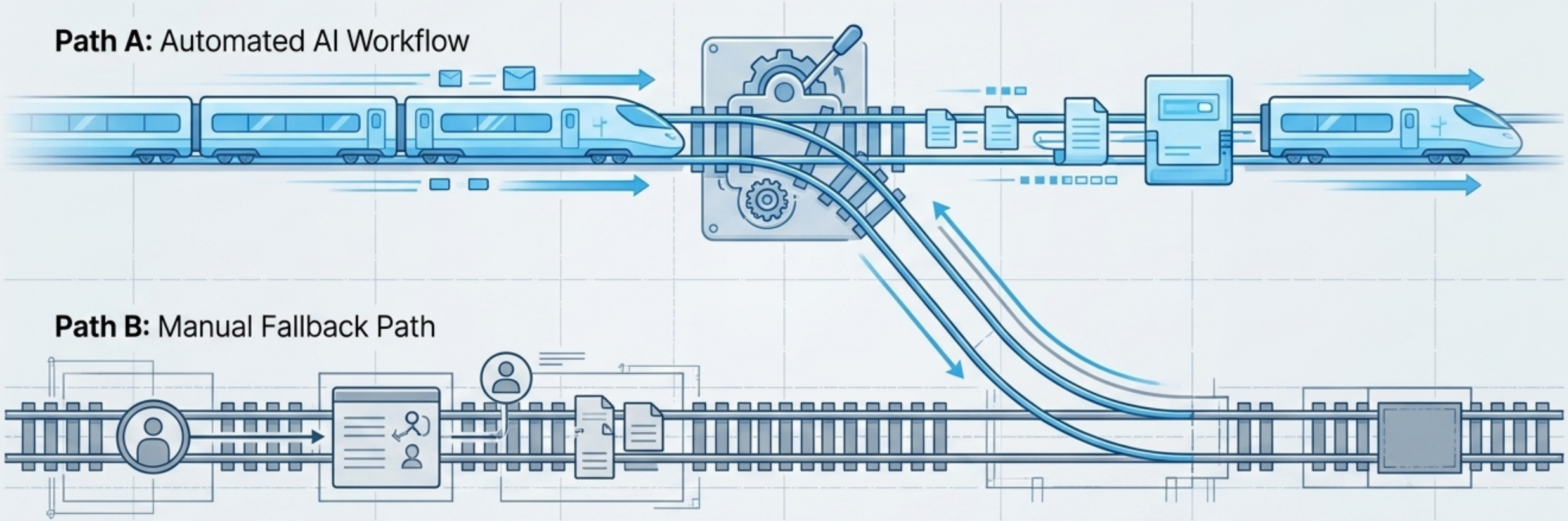
Never bypass the review gate for clinical, financial, or HR-impacting actions.

Routing Complex Cases to the Right Experts



Escalating a use case is a success metric. It means the safety system is working.

If You Cannot Roll It Back, You Cannot Roll It Out



Reversibility
Can the solution be turned off without breaking the entire department?

Fallback Plan
What happens exactly if the AI hallucinates, fails an audit, or goes offline?

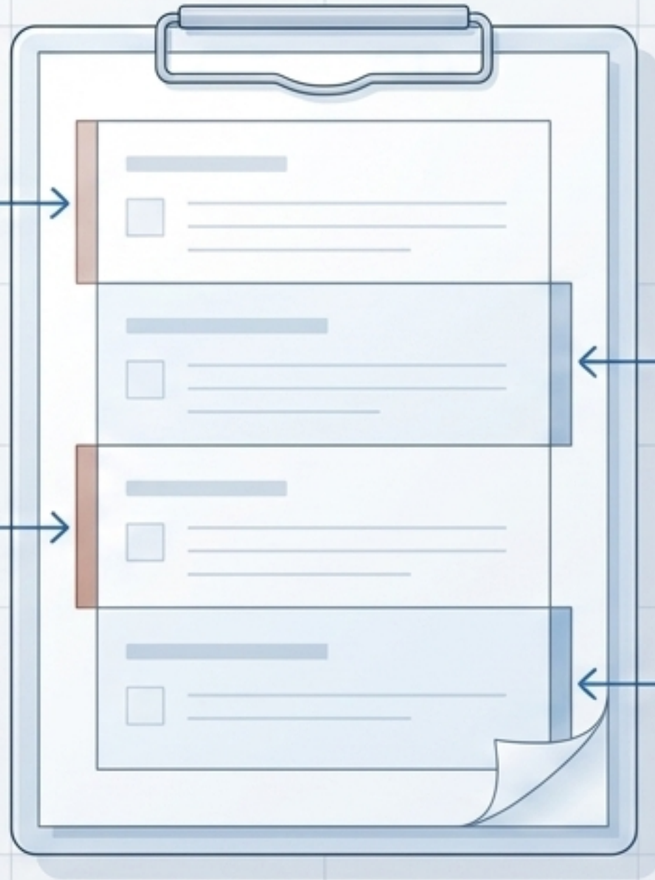
Capturing Risk Dimensions Before You Pilot

Section 1: Data & Inputs

Are sources approved?
Is data sensitive?

Section 2: Output & Impact

Who is affected? What is the automation level?



Section 3: Safeguards

Where are the human review points? Is there a fallback plan? Is it auditable?

Section 4: Assessment

Risk level determination, escalation needs, and final owner sign-off.

Classifying Real-World Scenarios

Use Case	Answers	Data	Impact	Review Point	Owner	Recommendation
Drafting patient discharge instructions.	(Requires review gate before patient sees it).	_____	_____	_____	_____	_____ _____ _____
Summarizing meeting notes (internal only).	(Low risk, ensure no sensitive HR data is included).	_____	_____	_____	_____	_____ _____ _____
Automated denial appeal letters directly to insurers.	(High risk, requires legal/finance escalation and strict review).	_____	_____	_____	_____	_____ _____ _____

Correcting Frequent Governance Assumptions

Myth

I'm using
approved
tool, so m
use case
automati
approved

Reality

Approved tools still require approved use cases. A safe tool can still execute risky workflows.

Myth

A human
will look
it.

Reality

Human review must be specifically defined, documented, and placed before the output creates impact.

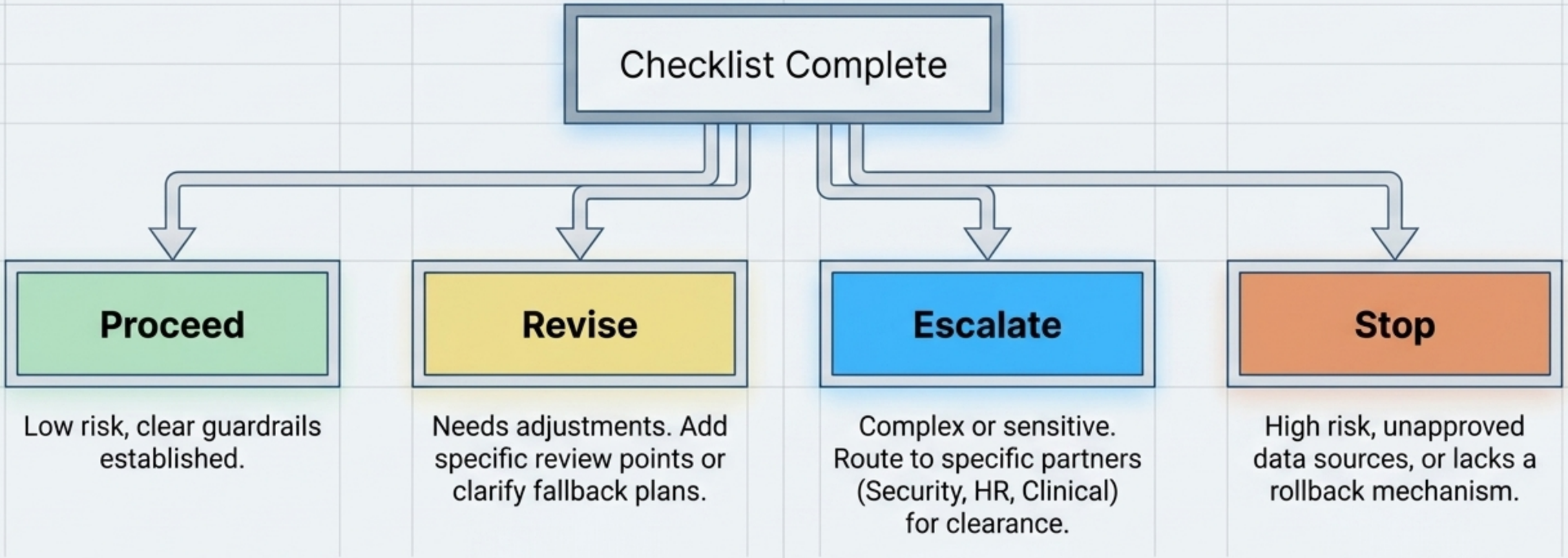
Myth

It's low-ris
so no rule
apply.

Reality

Even low-risk uses require clear data boundaries, approved sources, and assigned ownership.

The Final Decision Framework



Responsible AI protects patients, employees, the organization, and our credibility to keep innovating.