



Performance-Driven Risk Oversight for AI Agents in Clinical Research

Presented by

Laxmiraju Kandikatl

Maxis AI Inc, Edison, NJ

Session Objectives

1. Understand performance driven risk-based framework for oversight

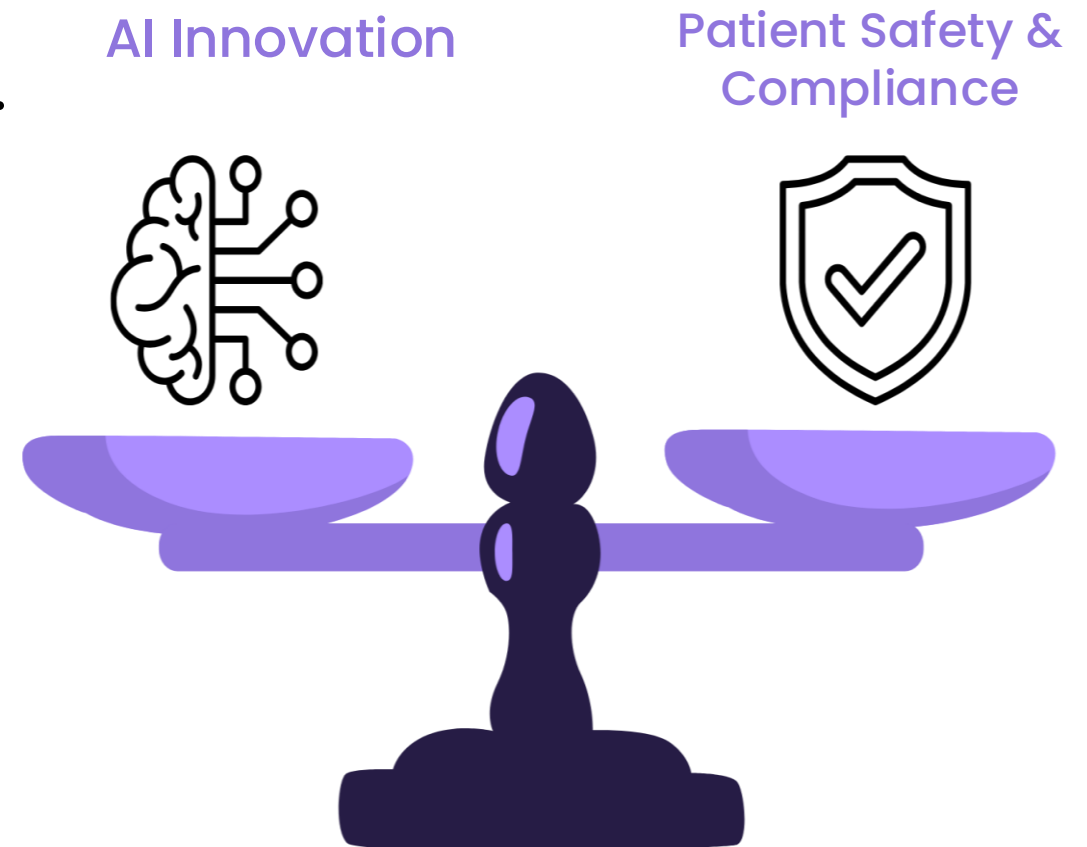
2. Explore various levels of human-oversight mechanisms

3. Apply different oversight levels according to the assessed risk

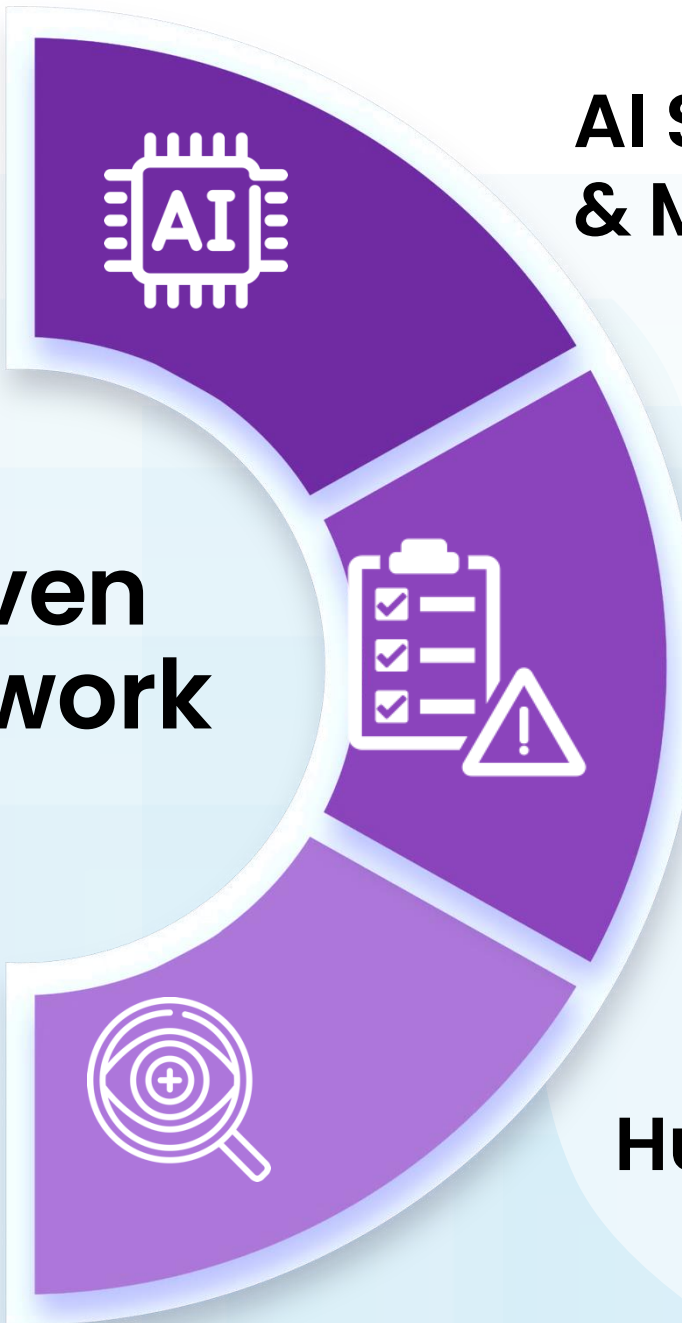
4. An example to illustrate how it works in practice

The Central Question: How Much Human Oversight is Enough?

The rapid deployment of AI in GxP environments has outpaced the development of clear governance standards.



Performance Driven Risk Based Framework



**AI System
& Metrics**

Risk Assessment

- AI System Risk
- Metrics Risk

Human Oversight
(HIC/HITL/HOTL)

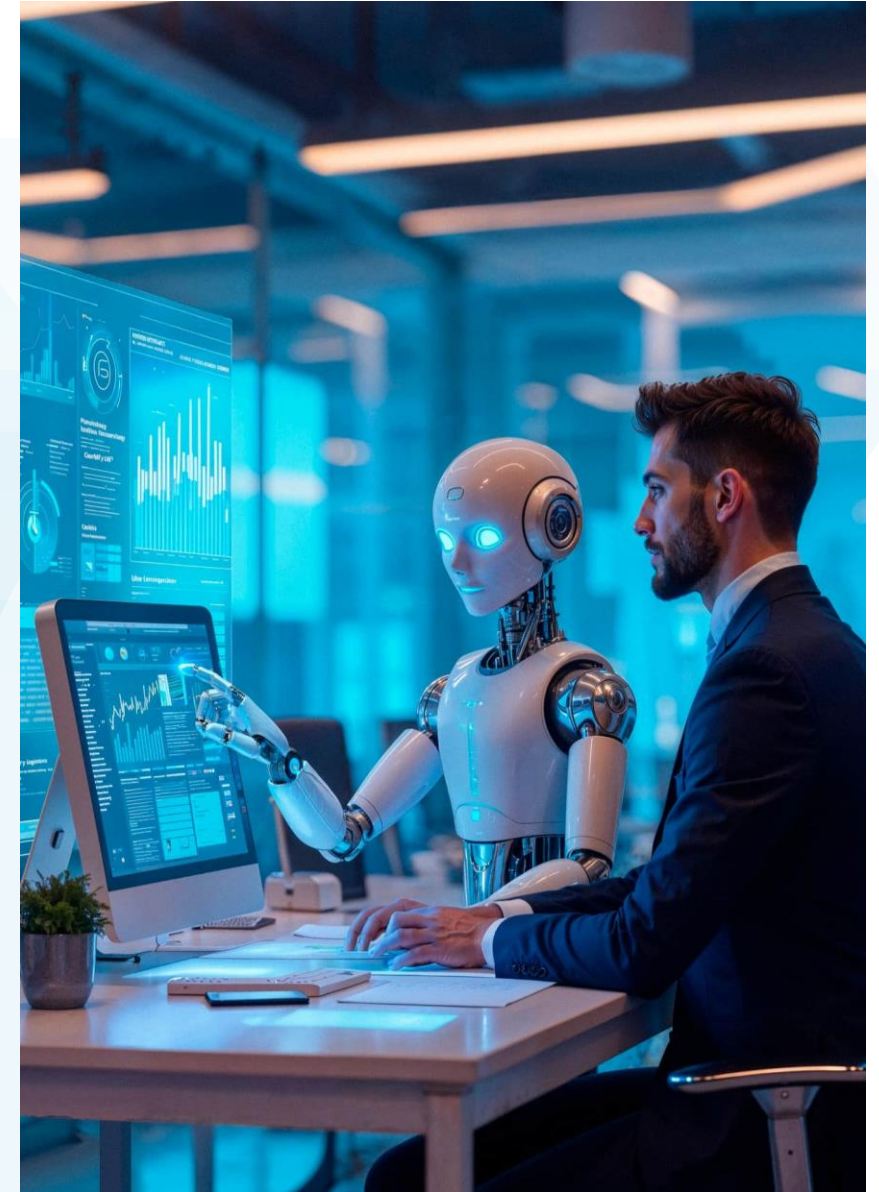
AI System

Context of Use

what question the AI is addressing.
How its output meant to be used?

Example

Clinical Data anomaly detection
Agent



Primary Metrics

Task
Completion
Rate (TCR)

Accuracy

Precision

Recall
(Sensitivity)

F1 Score

AUC-ROC



Secondary Metrics

Transparency

Bias and
Fairness

Robustness

Explainability

Consistency

Hallucinations

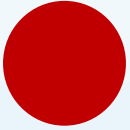
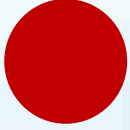
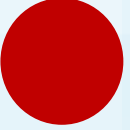


MAXIS Risk Assessment Mapping with ISO 31000



Risk Identification – AI System Risk

Model Influence

High			
Medium			
Low			
	Low	Medium	High

Decision Consequence

Model Influence:

Pertains to the relative weight attributed to the AI system's outputs compared to other sources of evidence.

Decision Consequence:

The potential severity, likelihood, and detectability of harm from an incorrect AI output

Risk Identification – Metrics Risk Classification

Metrics	High Performance (Low Risk)	Medium Performance (Medium Risk)	Low Performance (High Risk)
Task Completion Rate (TCR)	Greater than 95%	85 to 94%	80 – 85%
Accuracy	Greater than 90 %	80 to 89 %	75 % to 80 %
Precision	Greater than 90 %	80 to 89%	75 % to 80 %
Recall (Sensitivity)	Greater than 90 %	80 to 89%	75 % to 80 %
F1 Score	Greater than 0.85	0.75 to 0.84	0.70 to 0.75
AUC–ROC	Greater than 0.90	0.80 to 0.89	0.70 to 0.75

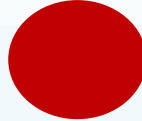
Risk Analysis

Integrated AI Risk (Metrics Risk vs AI System Risk)



Systems exhibiting high risk across both metrics risk and AI system risk warrant a halt in progression and require further investigation, remediation, and revalidation prior to continuation

Metrics Risk

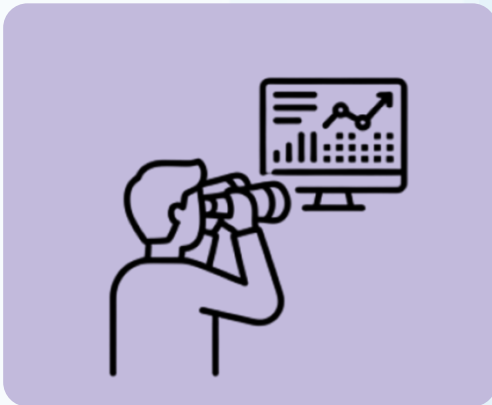
High			
Medium			
Low			
	Low	Medium	High

AI System Risk

Risk Evaluation & Treatment

The Integrated AI Risk score directly determines the required level of human involvement, ensuring oversight is always proportional to the risk.

Low Risk → Human-on-the-Loop (HOTL)



Humans monitor performance and intervene only when anomalies occur.

Medium Risk → Human-in-the-Loop (HITL)



Active human involvement in real-time decision-making processes.

High Risk → Human-in-Command (HIC)



Humans are the ultimate authority over AI systems.

Mapping Risk to Human Oversight

AI system with human oversight			
AI Risk	Low	Medium	High
Human Oversight	HOTL	HITL	HIC

Communication and Monitoring



Communication constitute a critical cross-cutting activity throughout the risk management process.



Finally, monitoring and review ensure that AI risk management remains effective over time.

Example Clinical data anomaly detection agent

Application in a clinical scenario:

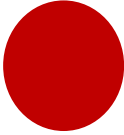
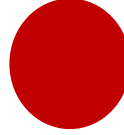
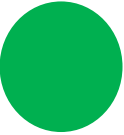
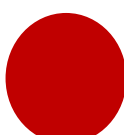
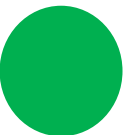
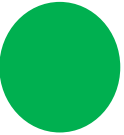
AI System Risk: High (High Influence vs. Medium Consequence)

Primary Metrics: Low Risk (e.g., TCR at 96%).

Integrated Risk: Medium – Medium AI System risk vs Low Metrics Risk

Outcome: Implementation of Human-in-the-Loop (HITL) oversight.

Metrics Risk

High			
Medium			
Low			
	Low	Medium	High

AI System Risk

Benefits and Limitations Overview

Benefits

- Improves Trust and Accountability
- Prevents Bias, hallucination and unsafe outputs
- Add Transparency
- Supports Regulatory compliance

Limitations

- Skill Gap – Need Training
- Time Consuming
- Unclear Governance
- Risk of Over- or Under-Oversight
- Continuous risk assessment

Is Human Oversight Enough?

Additionally,

- Appropriate control measures



Technical Controls

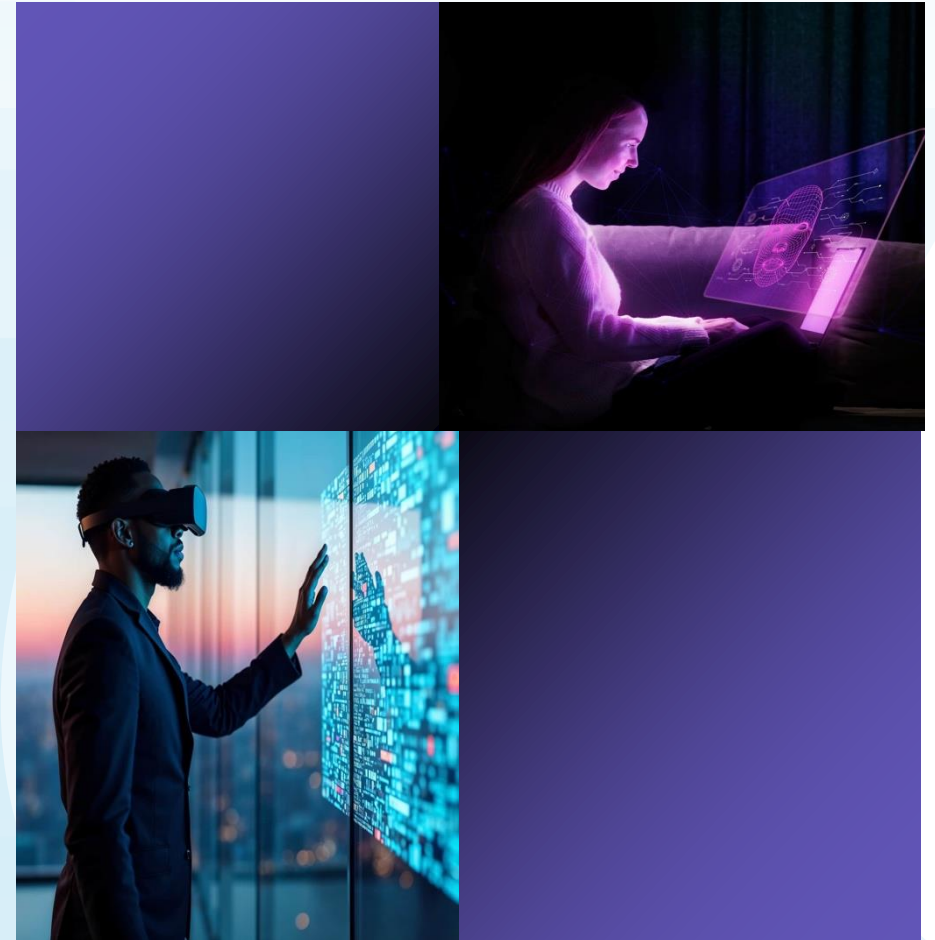


Procedural Controls , and



Organizational Controls

- are recommended to operationalize trust and ensure compliance.



Key Takeaways



Innovation vs. Accountability:

Performance Driven risk oversight balances technological advancement with clinical safety.

Proportionality:

Ensures human effort is focused where the risk is greatest.

Trustworthy AI:

Translates abstract ethical principles into concrete governance mechanisms.

Q&A

Thank you for your
time! Open to questions
& discussion.

**Laxmiraju
Kandikatla**

Maxis AI Inc



laxmiraju.k@maxisit.com