

Your Computer Can Detect Patterns— But It Can't Tell You What Matters.

- **Computational audits can flag disparities**, such as:
 - Differences in model performance across racial or ethnic groups
 - Shifts in outcome classification when endpoint definitions change
 - ...but these findings raise new questions that machines can't answer

Your Computer Can Detect Patterns— But It Can't Tell You What Matters.

- **Computational audits can flag disparities**, such as:
 - Differences in model performance across racial or ethnic groups
 - Shifts in outcome classification when endpoint definitions change
 - ...but these findings raise new questions that machines can't answer

Why We Need Humans in the Loop



Interpretation

- Is the observed bias *clinically significant*?
- Is it *ethically problematic*, or a reflection of meaningful group differences?



Moral Reasoning

- Should we change the algorithm, the model, or the clinical workflow?
- How do we weigh performance trade-offs (e.g., recall vs. fairness)?



Contextual Prioritization

- Researchers may prioritize statistical validity; patients and clinicians may prioritize *trust*, *access*, and *dignity*.
- Stakeholders bring insights about harms that aren't captured in metrics.

Your Computer Can Detect Patterns— But It Can't Tell You What Matters.

- **Computational audits can flag disparities**, such as:
 - Differences in model performance across racial or ethnic groups
 - Shifts in outcome classification when endpoint definitions change
 - ...but these findings raise new questions that machines can't answer

Why We Need Humans in the Loop



Interpretation

- Is the observed bias *clinically significant*?
- Is it *ethically problematic*, or a reflection of meaningful group differences?



Moral Reasoning

- Should we change the algorithm, the model, or the clinical workflow?
- How do we weigh performance trade-offs (e.g., recall vs. fairness)?



Contextual Prioritization

- Researchers may prioritize statistical validity; patients and clinicians may prioritize *trust*, *access*, and *dignity*.
- Stakeholders bring insights about harms that aren't captured in metrics.



Our approach is grounded in three interdependent yet tightly interlinked pillars—empirical audits, ethical reasoning, and stakeholder engagement—because only together can they move us toward truly fair and trustworthy AI.