

Thank You!

Constitutional AI and Medical Ethics

핵심 메시지 요약

1. Constitutional AI는 명시적 윤리 원칙으로 AI의 자기 개선을 유도
2. 의료 AI는 생명윤리 4원칙(선행, 무해, 자율성, 정의)을 준수해야 함
3. 안전 메커니즘: Red teaming, Guardrails, Output filtering
4. 편향 완화: Pre/In/Post-processing 기법 적용
5. 지속적 모니터링과 윤리 위원회의 역할이 필수적



Questions?