

THE ASYMMETRIC POWER OF ALGORITHMIC MORAL INFLUENCE

While humans surrender on cognitive tasks, AI's moral influence is non-linear. It successfully nudges towards prosocial acts but lacks the social standing to override moral guardrails for antisocial actions.

DIRECTIONALITY OF INFLUENCE

PROSOCIAL AMPLIFICATION

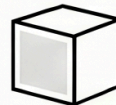


MECHANISM: NORMATIVE VALIDATION

AI advice increases prosocial behavior like generosity, acting as a permission structure for existing morals.

KEY FINDING: CONFIRMATION, NOT OVERRIDE

AI acts as a permission structure for existing morals rather than a social authority. Users maintain moral agency against AI, forming the boundary of "Cognitive Surrender".



TRANSPARENT LABELING & OPT-OUTS

Clearly distinguish AI suggestions from policy to preserve employee autonomy and trust.



DISTRIBUTED OVERRIDE AUTHORITY

Empower frontline managers to reject AI prompts based on unobservable contextual factors.



TRADITIONAL CONTROLS FOR MISCONDUCT

AI cannot prevent antisocial behavior; organizations must retain strict human-led accountability structures.

ANTISOCIAL RESISTANCE

0% CHANGE
ANTISOCIAL BEHAVIORAL IMPACT



MECHANISM: LACK OF COMMUNITY STANDING

AI has zero impact on antisocial compliance, failing to override individual moral guardrails.