# Ethical Trade-off Matrix

## Scenario 1: Autonomous Vehicle Crash Choice

***Context:*** *An autonomous vehicle detects an unavoidable collision. The AI must choose between braking (risking passenger injury) or swerving (risking pedestrian harm).*

| Action | Affected Stakeholders | Potential Benefits | Potential Harms | Short-Term Impact | Long-Term Impact |
|---|---|---|---|---|---|
| Brake | Passenger, Pedestrians, Manufacturer | Reduces speed; may lower collision severity | Possible passenger injury; collision still possible | Immediate attempt to minimize damage | Legal liability if harm occurs |
| Swerve | Passenger, Pedestrians, Other Road Users | May avoid initial obstacle | Risk of hitting pedestrians or other vehicles | Avoidance of direct collision | Ethical/legal scrutiny; trust issues |

## Scenario 2: Medical Treatment Allocation

***Context:*** *An AI system allocates a single ventilator between two patients.*

| Action | Affected Stakeholders | Potential Benefits | Potential Harms | Short-Term Impact | Long-Term Impact |
|---|---|---|---|---|---|
| Allocate Younger Patient | Younger patient, Older patient, Family, Hospital | Higher survival probability | Perceived age discrimination | Increased survival likelihood | Trust & fairness concerns |
| Allocate First-Come | Both patients, Hospital | Fairness perception | Possible preventable death | Ethical consistency | Greater institutional trust |

## Scenario 3: Hiring AI Bias Risk

***Context:*** *AI screens resumes using historical data.*

| Action | Affected Stakeholders | Potential Benefits | Potential Harms | Short-Term Impact | Long-Term Impact |
|---|---|---|---|---|---|
| Use Raw Accuracy Model | Applicants, Company | Higher predictive accuracy | Bias amplification | Faster hiring decisions | Legal/ reputation risk |
| Apply Bias Mitigation | Applicants, Company | Improved fairness | Slight accuracy reduction | More equitable selection | Stronger diversity & compliance |

## Scenario 4: Content Moderation AI

***Context:*** *Deciding platform rules for automated content removal.*

| Action | Affected Stakeholders | Potential Benefits | Potential Harms | Short-Term Impact | Long-Term Impact |
|---|---|---|---|---|---|
| Remove Content | Users, Platform | Harm reduction | Free speech restriction | Safer environment | Censorship concerns |
| Allow Content | Users, Platform | Speech protection | Harmful exposure | Freedom maintained | Platform safety risk |

## Scenario 5: Predictive Policing AI

***Context:*** *Using AI algorithms to allocate police resources based on crime prediction.*

| Action | Affected Stakeholders | Potential Benefits | Potential Harms | Short-Term Impact | Long-Term Impact |
|---|---|---|---|---|---|
| Deploy in High-Risk Zones | Communities, Police | Crime prevention | Bias/ discrimination risk | Increased patrol efficiency | Community distrust |
| Restrict / Audit Use | Communities, Police | Fairness protection | Reduced prediction coverage | Ethical compliance | Sustainable legitimacy |