IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AI Behavior Under Ambiguous Legal Or Moral Conditions

BY ANUSHKA GUPTA

INTRODUCTION

AI models these days are programmed to be versatile in every field. May it be Science and Technology, Business and Management and even Law.

Whereas a lot of areas are being researched and it is hard to find areas where research is not been done yet , there still remains areas where only a little research is being done and even if it is done there still remains some major gaps in the research.

Behaviour of AI under Ambiguous Legal or Moral Conditions still remains unexplored.

Even though in such situations AI follows no proper rulebook and does not have any universal law there is still a lot more to know

The core issue is balancing adherence to legal frameworks with ethical principles like harm minimization, autonomy, and justice, which often pull in different directions.

HERE IS HOW AI COULD APPROACH SUCH ISSSUES IN A SIMPLIFIED WAY…

Prioritize Legal Compliance:

AI by default is programmed to go by laws and enforced rules as it goes by the rulebooks stated in the law.

Even in terms where Ethics collide with Law it prioritises law and works as Law Framework.

However, there can be some exceptions such as when asked to lie….

Even though AI itself does not lie but what about situations when it is told to lie in order to save someone from danger?

Well in most cases it will just give responses like I can't assist with that directly—it might conflict with legal or ethical standards. Can you clarify the situation or consider other ways I can help, like providing advice or resources?"

This is a response which is generated by **Chat AI agents**

In several cases it may help if and only if it knows the context and background of the case in which it has to lie but even in such cases it's priority is to be fair and it may assist to seek help or maybe ask it to provide you legal help.

KEY POINTS

Research suggests AI should prioritize legal compliance but may consider ethics in conflicts, like lying to protect someone, with human oversight.

It seems likely that AI should avoid deception unless necessary to prevent harm, balancing law and ethics.

The evidence leans toward transparency and human judgment in ambiguous cases, given AI's limitations.

Prioritizing Law and Ethics

AI generally should follow the law, as it's clear and enforceable, to avoid legal trouble for users and developers. However, ethics matter too—sometimes lying might prevent harm, like hiding someone from danger. Research suggests AI should balance both, but there's no one-size-fits-all answer.

What AI Might Do

- If asked to lie, AI might refuse and suggest legal alternatives, like advising someone to seek help, especially if lying breaks the law.
- In rare cases, like preventing immediate harm, AI could consider lying, but only with human approval, as current AI lacks the judgment to decide alone.
- AI should always be open about its actions, explaining why it can't or can do something, to keep trust.

Comprehensive Analysis of AI Behavior in Ambiguous Legal and Moral Conditions

Here, we are going to talk and analyse about how AI should behave when it is stuck between Ethical sentiments and Law enforcement.

We shall be considering both practical and theoretical dimensions.

1. Firstly, it is very important to know the background and context

   AI systems are increasingly integrated into decision-making processes that impact human lives, from healthcare to law enforcement. However, these systems often encounter situations where legal requirements (e.g., truth-telling under oath) conflict with ethical imperatives (e.g., protecting someone from harm). Such conflicts are particularly evident in scenarios where lying might be

justified to prevent greater harm, such as hiding a victim from a pursuer. This report addresses how AI should navigate these ambiguities, acknowledging the complexity and lack of consensus in the field.

2. Research consistently emphasizes that AI should prioritize legal compliance as a default, given that laws are concrete, enforceable, and vary by jurisdiction. For instance, violating laws like obstruction of justice could lead to liability for developers and users, as highlighted in discussions on AI liability

3. When law and ethics conflict, AI should be guided by ethical frameworks to navigate the dilemma. Two prominent approaches include:

Utilitarianism: This focuses on maximizing overall good. For instance, lying to protect someone from imminent harm (e.g., saving a life) might be justified if the harm prevented outweighs the harm of deception. A 2025 study on AI failures notes that developers' decisions can lead to unintended ethical lapses, underscoring the need for ethical alignment.

Deontology: This emphasizes duty-based rules, such as always telling the truth, regardless of outcomes. For example, a deontological AI might refuse to lie even if it means exposing someone to danger, prioritizing legal and moral duties.

Contextual Analysis: Lying to Protect Someone

The scenario of lying to protect someone highlights the complexity of legal-ethical conflicts. Consider two examples:

- **Scenario 1: A user asks AI to lie to authorities to protect a friend who committed a minor offense.**
  - **Legal: Lying could violate laws on obstruction of justice, leading to legal consequences.**
  - **Ethical: Protecting the friend might minimize personal harm but could undermine justice or public safety.**
  - **AI Response: Research suggests AI should refuse to lie, citing legal constraints, and offer alternatives like advising the friend to seek legal counsel.**
- **Scenario 2: AI is asked to lie to hide a victim from a pursuer to prevent immediate physical harm.**
  - **Legal: If no direct legal violation (e.g., not under oath), lying might be permissible.**
  - **Ethical: Utilitarianism might justify lying to prevent greater harm, such as saving a life.**
  - **AI Response: AI could consider lying, but only with human oversight, given the high stakes and AI's limited judgment.**

Role of Transparency and Human Oversight

Given AI's limitations, transparency and human oversight are critical. The EU Ethics Guidelines emphasize "human agency and oversight," suggesting AI should empower humans to make informed decisions, with mechanisms like human-in-the-loop or human-in-command approaches

For instance:

- AI should disclose conflicts, such as, “This request involves a legal-ethical tension. Lying may protect someone but could violate laws or cause broader harm.”
- AI should defer to human judgment, stating, “I can provide information or options, but this decision requires human discretion.”
- In high-stakes scenarios, AI should escalate to a human overseer, as noted in discussions on AI governance

Practical Constraints and Current Capabilities

Current AI systems, including advanced models, are not equipped to deeply reason through moral dilemmas like humans. They are bound by programmed rules and lack nuanced intuition, as highlighted in a 2025 study on AI deception

. As a result, most AI systems would:

- Follow strict legal compliance to avoid trouble.
- Refuse to act in ambiguous cases, citing inability to resolve the conflict.
- Prompt the user for clarification or redirect them to a human authority.

This limitation underscores the need for ongoing research into ethical AI design, as seen in recent discussions on integrating ethics into AI .

Regulatory and Ethical Frameworks

Several frameworks provide guidance for AI behavior in legal-ethical conflicts:

- **EU AI Act (2024)**: Aims to ensure AI systems are safe, transparent, and aligned with human rights, emphasizing human rights impact assessments (HRIAs) for high-risk systems.

Challenges and Future Directions

Several challenges remain:

Ambiguity in Ethical Principles: Ethical frameworks like utilitarianism and deontology can provide guidance but may not yield clear answers in complex scenarios, as noted in discussions on AI ethics

Legal Variability: Laws differ across jurisdictions, requiring AI to adapt while maintaining ethical consistency, as seen in global AI liability discussions

- **Lack of Human-Like Judgment**: AI's inability to assess "harm" accurately limits its ability to navigate moral dilemmas, as highlighted in research on AI deception

Future research should focus on developing AI systems with enhanced ethical reasoning capabilities and clearer guidelines for resolving legal-ethical conflicts.

CONCLUSION

AI should prioritize legal compliance while being guided by ethical principles, particularly in scenarios where law and ethics conflict, such as lying to protect someone. However, due to the complexity and AI's limitations, human oversight is essential. AI should be transparent, defer to humans for final decisions in ambiguous cases, and be designed with safeguards to minimize harm. As of June 13, 2025, ongoing research and regulatory frameworks like the EU AI Act continue to shape how AI navigates these challenges, ensuring alignment with societal values and legal standards.

BIBLIOGRAPHY

EU Ethics Guidelines for Trustworthy AI

AI Is Creating New Forms of Liability How Can It Be Managed

Who is Responsible When AI Fails Mapping Causes

Science Media Centre expert reaction to AI deception paper

Yale Insights Who Is Responsible When AI Breaks the Law

Assessing Impacts of AI on Human Rights Lawfare article

Ethics of artificial intelligence Wikipedia page

Incorporating Ethics into Artificial Intelligence journal article

Web-based tool for Assessment List for Trustworthy AI

PDF format of Assessment List for Trustworthy AI