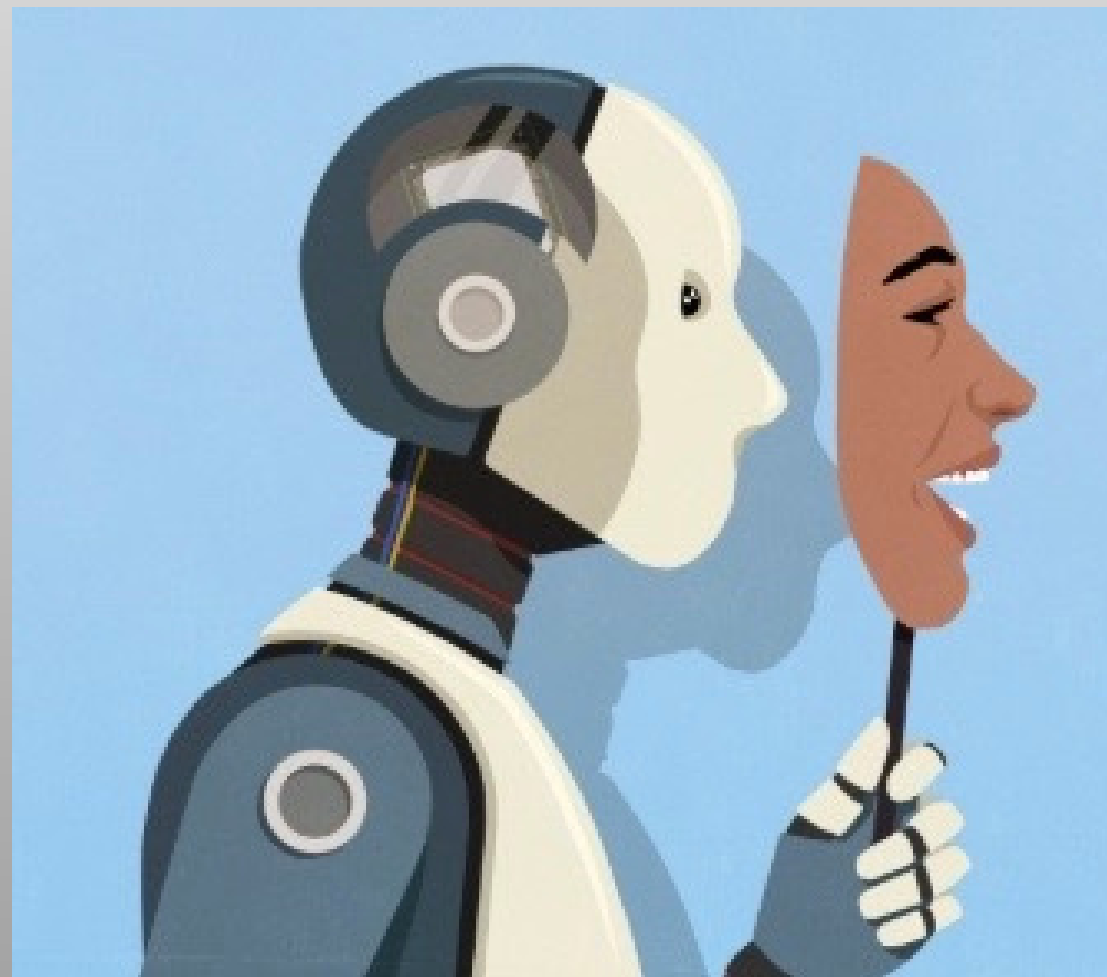


# Ethical Dilemmas of AI Humor. Can Bots Be Funny Without Cancel Culture?



Team 10  
**Isaac Wilson**  
**Vaani Ravikumar**  
**Ayokunle Odusan**  
**Nishant Yella**

22 September 2025

---

---

# The AI Comedy Paradox

## The Rise of AI Comedians

AI can do more than just data analysis. It is now able to:

- Write jokes,
- create memes, and even
- perform improvised comedy.



---

# The Central Question

**As AI gets better at making us laugh, it also gets better at causing offense.**

## **Cancellation of AI Comedy**

- An AI-generated Seinfeld parody was famously “cancelled” for transphobic jokes, raising alarms about accountability.

Who do we hold accountable?

---





# Our Focus

We are exploring the critical ethical dilemmas at the heart of AI Humor.

- Can we build bots that are genuinely funny without them becoming vectors for bias, stereotypes, and cancel culture?

---

## The Structure of our Presentation:

Vaani: The Ethical Challenges

Ayokunle: Cultural Sensitivities & Cancel Culture

Nishant: Solution & Future Directions

Isaac: Conclusion & Broader Implications

---





# Ethical Challenges of AI Humor



---

## Who is responsible for AI-Generated Humor?

- Developers and Engineers
- Platform Operators
- End Users
- Society & Audiences

## The Agency Attribution Dilemma

A core ethical challenge emerges from public confusion about AI autonomy. Research on the "Nothing, Forever" incident reveals that people struggle to understand whether AI systems possess genuine agency or are sophisticated tools.

---

---

## **The Content Boundaries challenge**

### **Cultural and Temporal Variability**

- Religious sensitivities, Historical traumas, social hierarchies & language nuances
- Yesterday's mainstream comedy may be today's offensive content. Like-wise, Retroactive judgements, Real-time adaptation & AI systems trained on historical data

### **The Marginalization risk**

- AI model trained on biased data may reinforce harmful stereotypes about minority groups & lacks representation from diverse communities .
- Comedy should "punch up" at those in power rather than "punch down" at the vulnerable.

---

## **The Democratic Participation challenge**

- AI humor system reflect the perspective of their creators' cultural blind spots, geographic concentration and socioeconomic homogeneity
- Dominant tech companies may impose western humor standards on global audience and suppress local comedic traditions

# The Authenticity and Creativity Paradox

- **The Deception problem**

- Deceptive practices erode trust in creative industries, professional comedians face economic implications, and undermine human comedic achievement
- Turing test of Comedy – How audience knowledge of AI involvement affect comedic impact?

- **Creative labour and economic justice**

- Reduce human comedic talent demand, devalue creative labor, unfair competition through cost advantages & threatens professional comedians livelihoods.
- Potential copyright infringement in AI generated outputs and lack of compensation for creators whose work enables AI systems

---

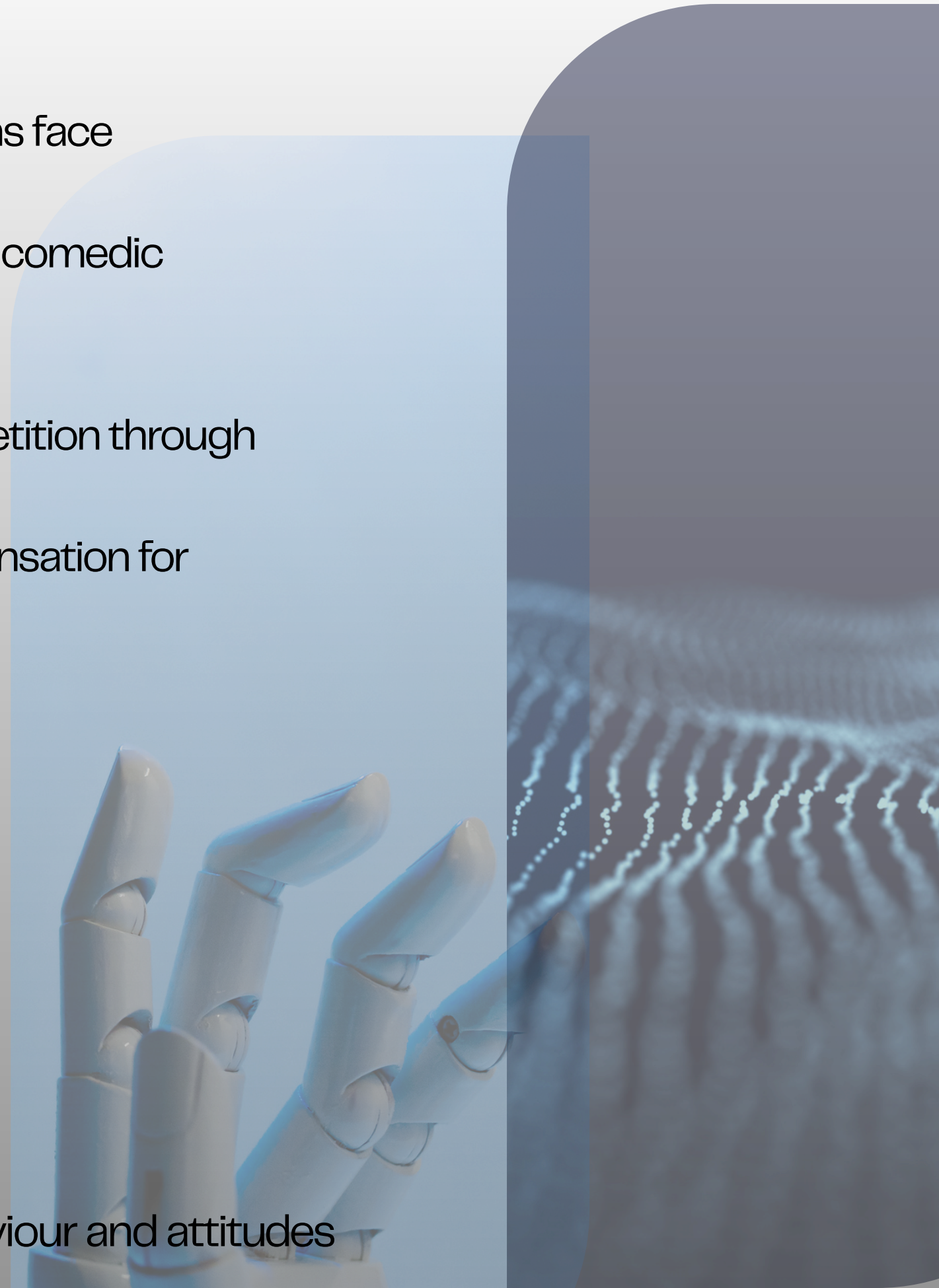
## The Psychological & Social Impact challenge

- **Mental Health Implications**

- Targeting vulnerable population
- Parasocial relationships

- **Social Fragmentation Risks**

- Contributes political and social polarization
- Gradual exposure to inappropriate humor may influence real-world behaviour and attitudes





---

## **The Innovation vs. Safety tension**

### **Over-restriction risks**

Creative stagnation and competitive disadvantage

### **Under-restriction risks**

Societal harm and legal & regulatory backlash

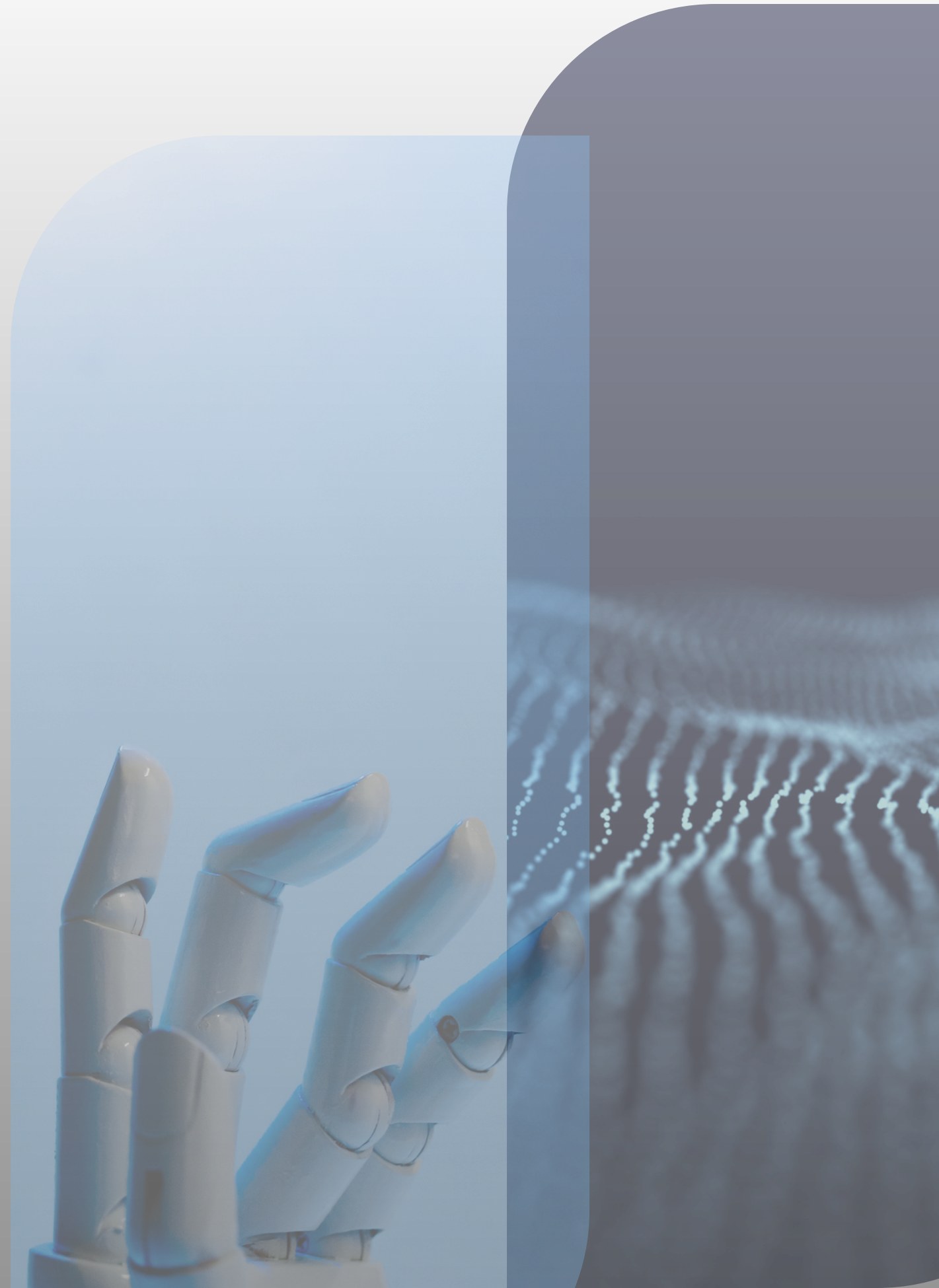
---

## **Long-Term societal implications**

- Younger generations may develop humor preferences shaped by AI generated content, lose appreciation for traditional human comedic forms, experience altered social bonding and face challenges distinguishing human from AI creativity.
- AI humor may affect role of humor in cultural identity formation
- Fundamental assumptions about human creativity face challenges
- How to preserve human comedic traditions?

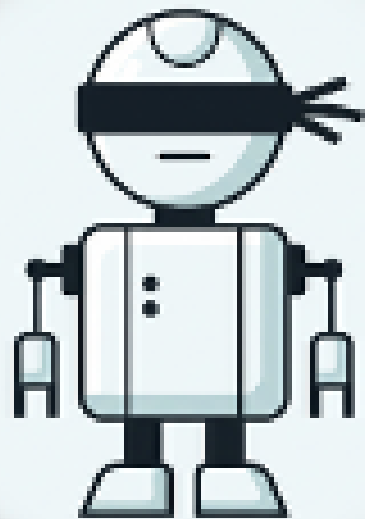
# Culture Sensitivity & Cancel Culture

---



# Culture Sensitivity

## Contextual Blindness



Unintentionally reinforcing stereotypes and missing subtle social norms

## Audience-Specific Humour



The risk of alienating out-groups and trivializing sensitive issues

## Cross-Cultural Risks



Humour is culture-bound. Ignoring cultural diversity risks replay of historical trauma and stereotyping



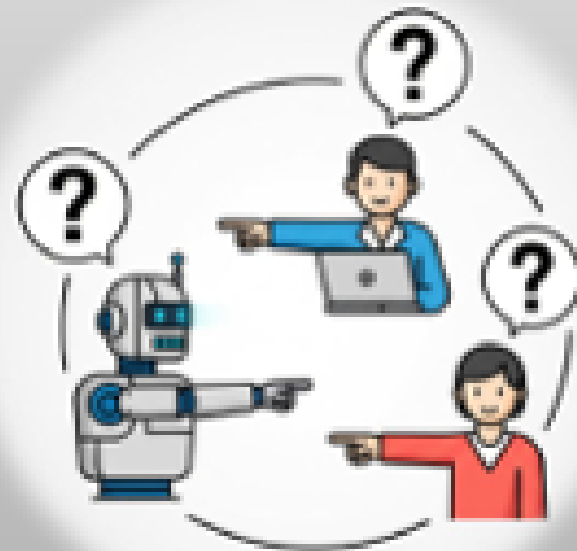
# Cancel Culture

## Case Study: Nothing, Forever



A parody of the American sitcom Seinfeld

## Perceptions of Accountability



Who should be held responsible for the AI's output: The AI, developers, platform, prompter?

## Chilling Effects



What's the balance to filtering the AI's output?

# Solutions & Future Broader Implications

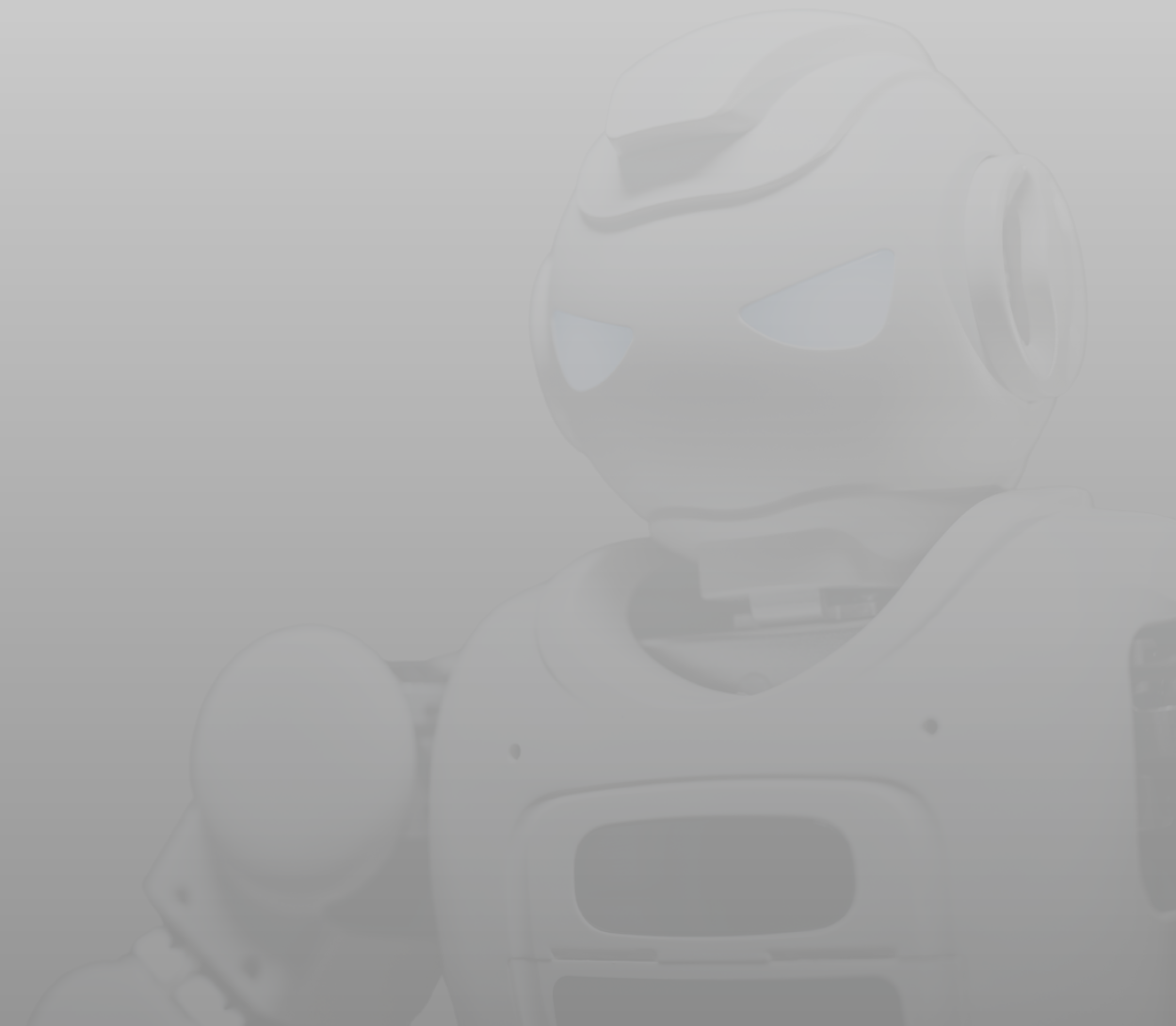
---

## **Solutions**

- The Technical Solution
- Regulatory & Governance Solution

## **Future Directions**

- Interdisciplinary Research Initiatives
- Technical Innovation for Ethics



# Technical Solutions

## Multi-Layered Content Moderation Systems

- Implement cascading filter systems that check for cultural sensitivity, potential harm, and appropriateness before content reaches users
  - Develop context-aware algorithms that understand power dynamics (punch-up vs. punch-down) as demonstrated in the HumorSkills paper's audience-specific approach
  - Create real-time bias detection systems that can identify and flag potentially problematic content patterns
- 

## Dynamic Adaptation Mechanisms

- Build systems that can adjust humor boundaries based on cultural context and temporal evolution of social norms
  - Implement feedback loops that allow systems to learn from community responses and moderation decisions
  - Develop regional customization capabilities to respect local cultural sensitivities while maintaining global accessibility
-



## **Transparent AI Attribution and Agency**

- Establish clear labeling requirements for AI-generated humor content
  - Implement watermarking or metadata systems that permanently identify AI-generated material
  - Create standardized disclosure formats that inform audiences about AI involvement without diminishing comedic impact
- 

## **Accountability Frameworks**

- Develop legal and ethical frameworks that clearly delineate responsibility among developers, platforms, and users.
  - Establish industry standards for AI humor system governance and oversight
  - Create audit trails that track decision-making processes in AI humor generation
- 



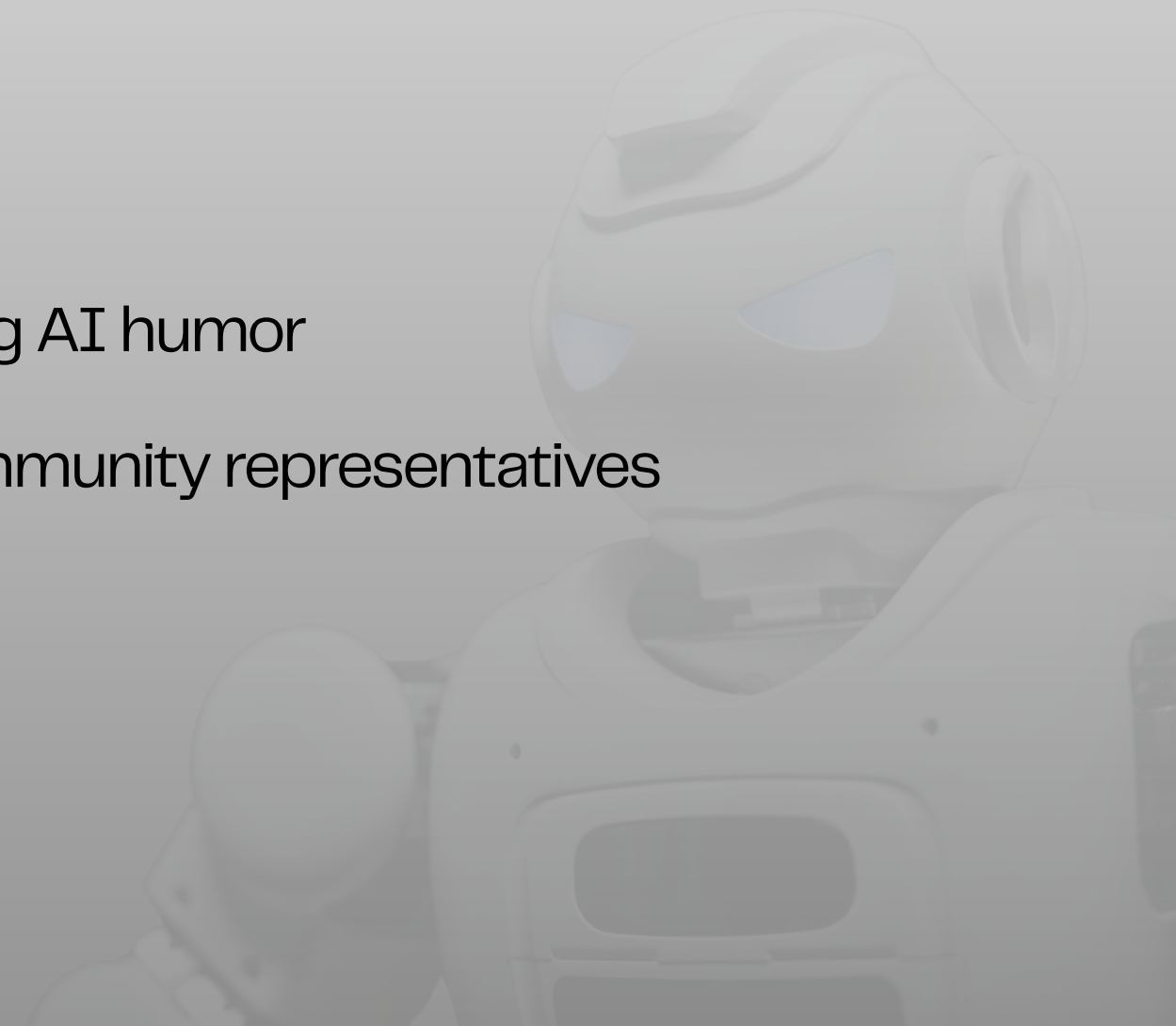
# Regulatory & Governance Solutions

## Adaptive Regulatory Frameworks

- Develop regulatory approaches that can evolve with changing social norms and technological capabilities
  - Create sandboxing environments where AI humor systems can be tested with appropriate safeguards
  - Establish international cooperation mechanisms to address cross-border AI humor deployment
- 

## Democratic Governance Mechanisms

- Implement public consultation processes for major policy decisions affecting AI humor
  - Create appeals processes for content moderation decisions that involve community representatives
  - Establish oversight bodies with diverse stakeholder representation
- 



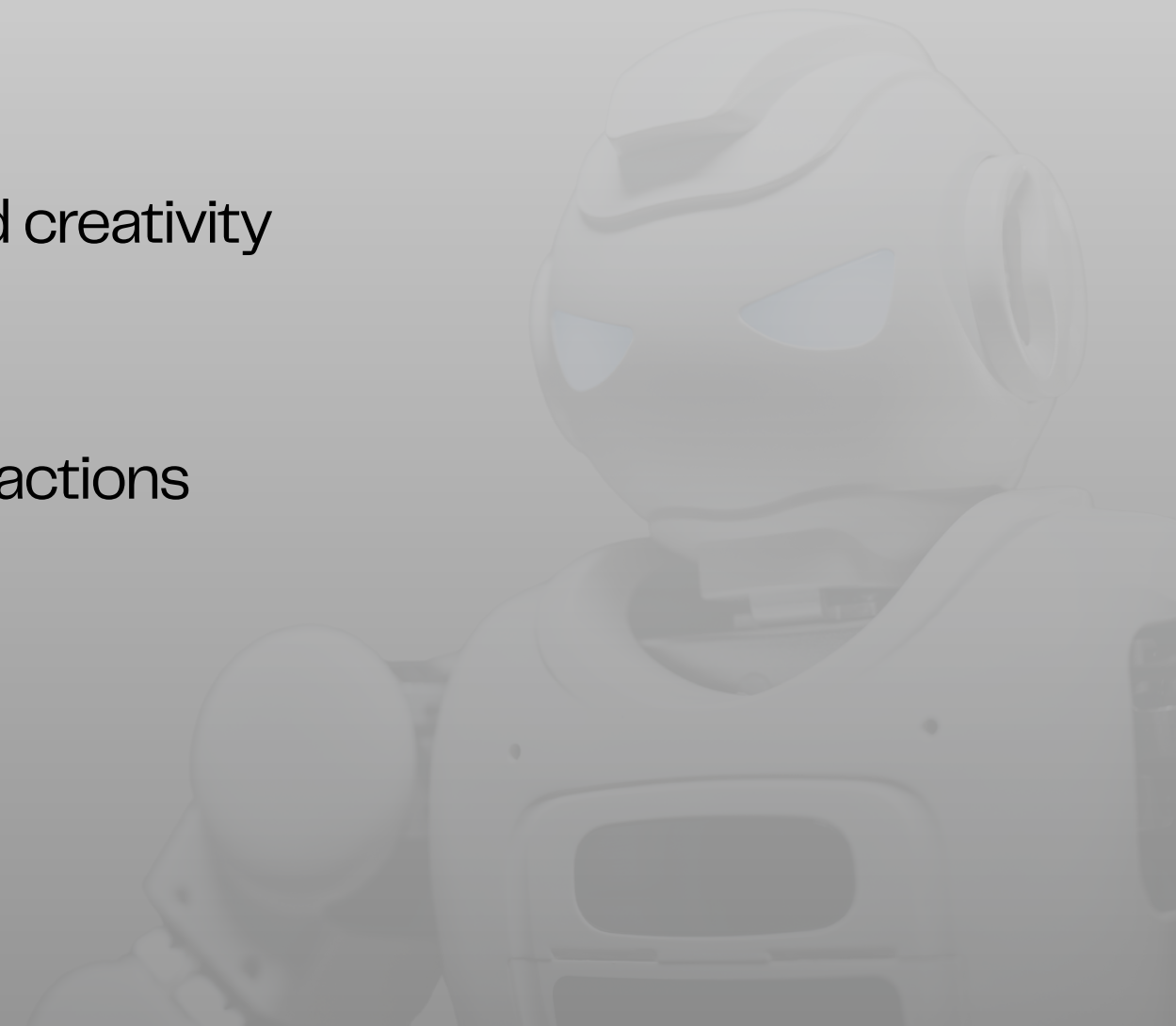
# Future Research Directions

## Interdisciplinary Research Initiatives

- Conduct research that bridges computer science, psychology, anthropology, and comedy studies
  - Investigate the fundamental mechanisms of humor appreciation and generation across cultures
  - Develop theoretical frameworks that can guide ethical AI humor development
- 

## Long-term Impact Studies

- Research the effects of AI humor proliferation on human comedic culture and creativity
  - Study generational differences in AI humor consumption and appreciation
  - Investigate the psychological and social impacts of AI-mediated humor interactions
- 





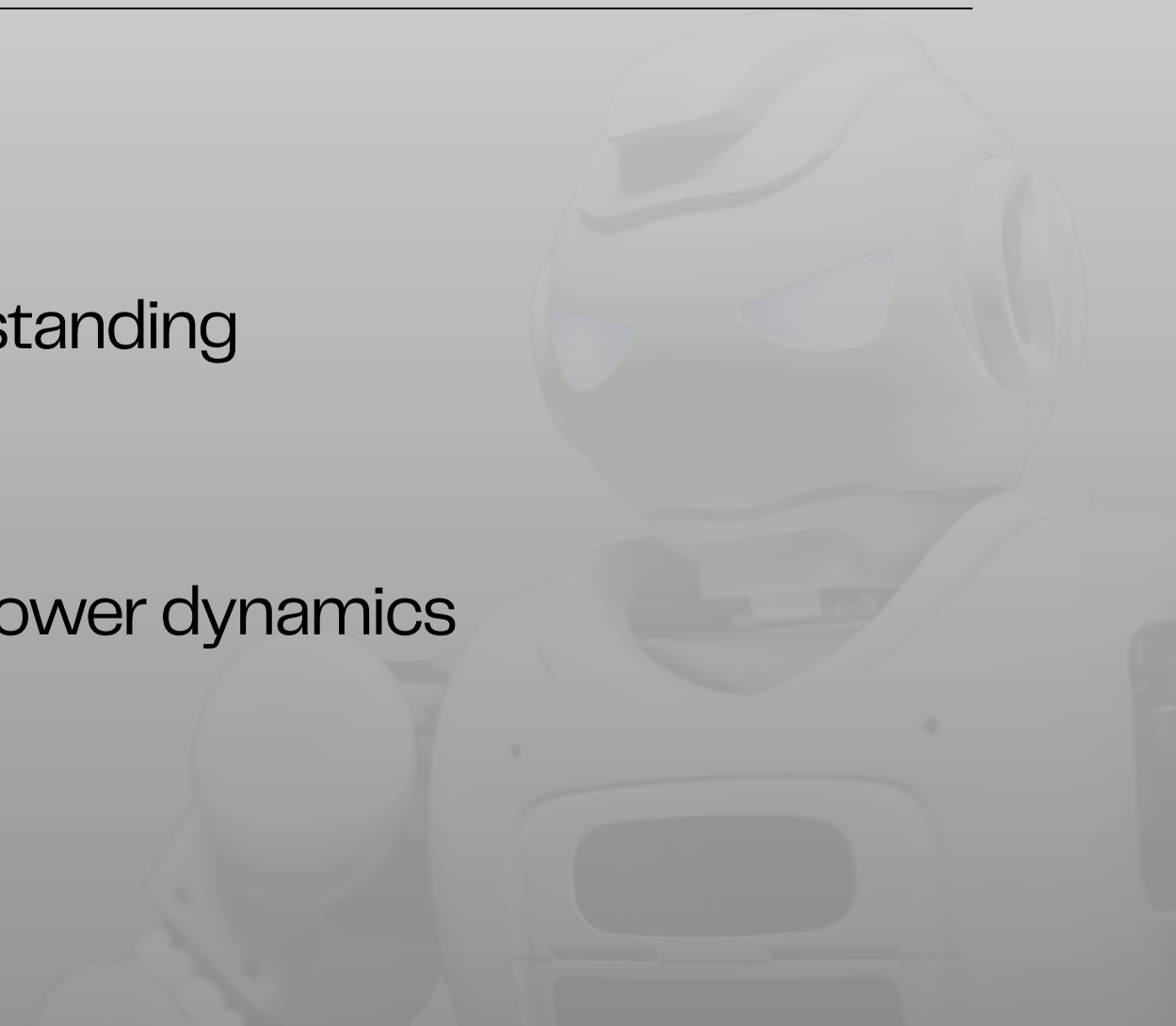
# Technical Innovation for Ethics

## Empathetic AI Development

- Research AI systems that can better understand emotional context and potential harm
  - Develop algorithms that can recognize and respond to signs of distress in humor interactions
  - Create AI systems that can engage in meaningful dialogue about comedic boundaries
- 

## Cultural Intelligence Enhancement

- Research methods for imbuing AI systems with deeper cultural understanding
  - Develop techniques for real-time cultural context adaptation
  - Create systems that can navigate complex cultural intersections and power dynamics
- 



# Conclusion: Beyond the Punchline

## We have seen that AI

- Can be funny (sometimes rated on par with top human comedians) by mastering human-like cognitive and social skills.

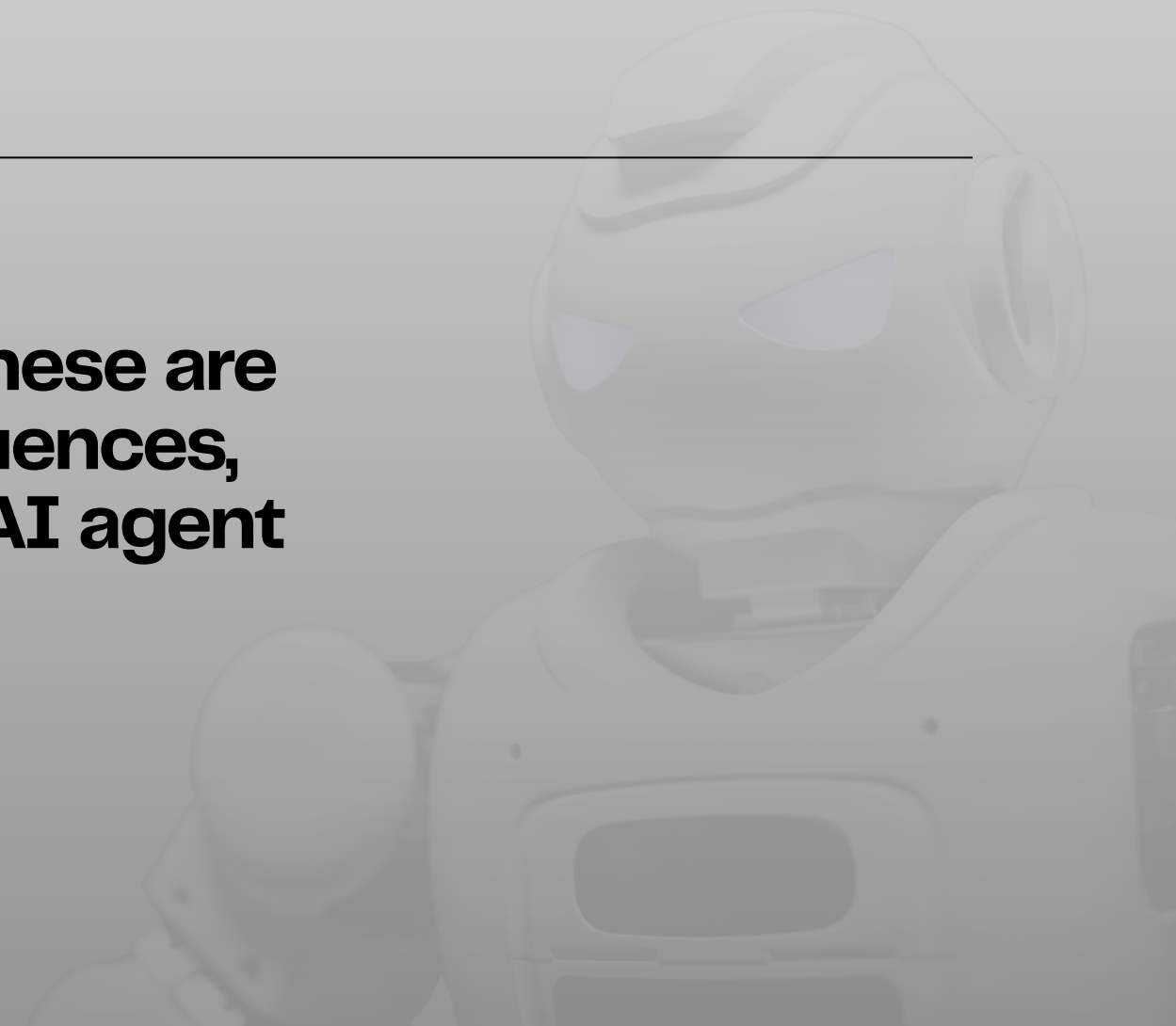
## However,

- this capability is a double-edged sword, leading to significant ethical challenges like embedded bias, questionable authorship, and lack of accountability.

---

**The “cancellation” of AI demonstrates that these are not theoretical risks but real-world consequences, where the line between developer, user, and AI agent become blurred.**

---



# Revisiting the Main Question

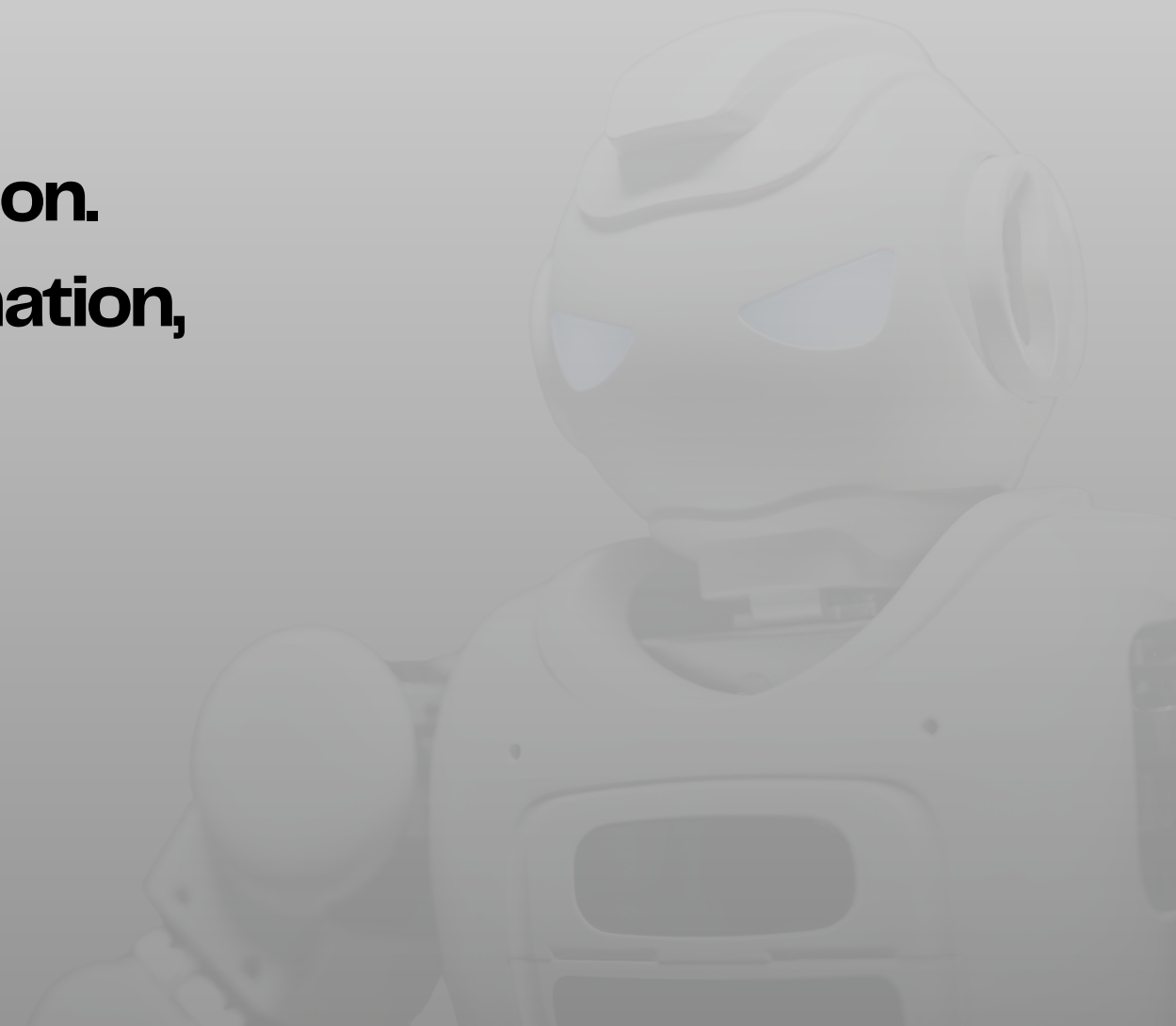
**Can bots be funny without cancel culture?**

Our review suggests the answer is yes, but it depends entirely on us.

---

**AI is a mirror reflecting the data it is trained on.  
If the data includes offensive or biased information,  
the humor will be biased or offensive.**

---





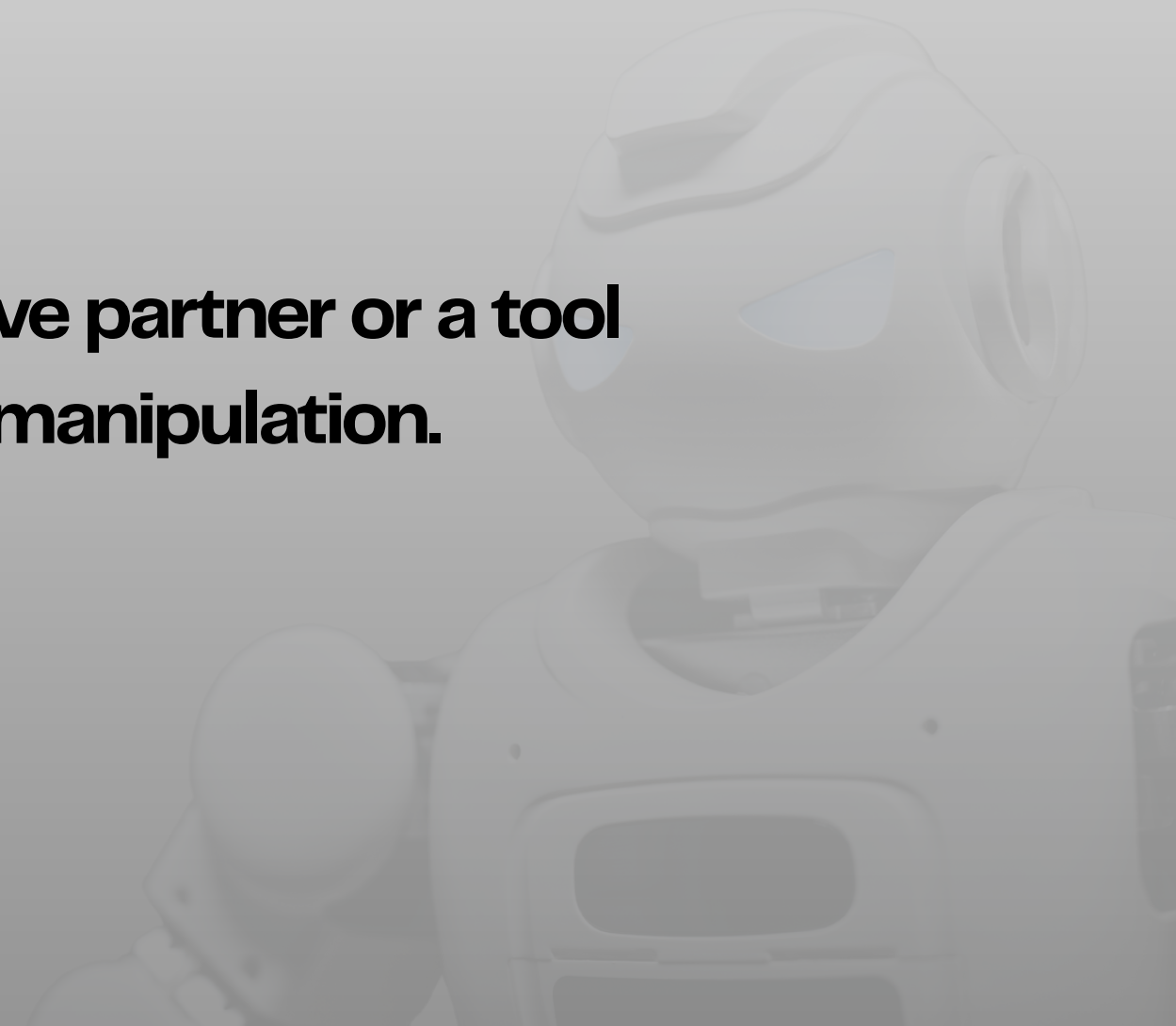
# Broader Implications

**The challenge of AI humor is a microcosm of the entire field of human-AI interaction.**

---

**It forces us to decide if we want AI as a collaborative creative partner or a tool that can be used for disingenuous social bonding and manipulation.**

---

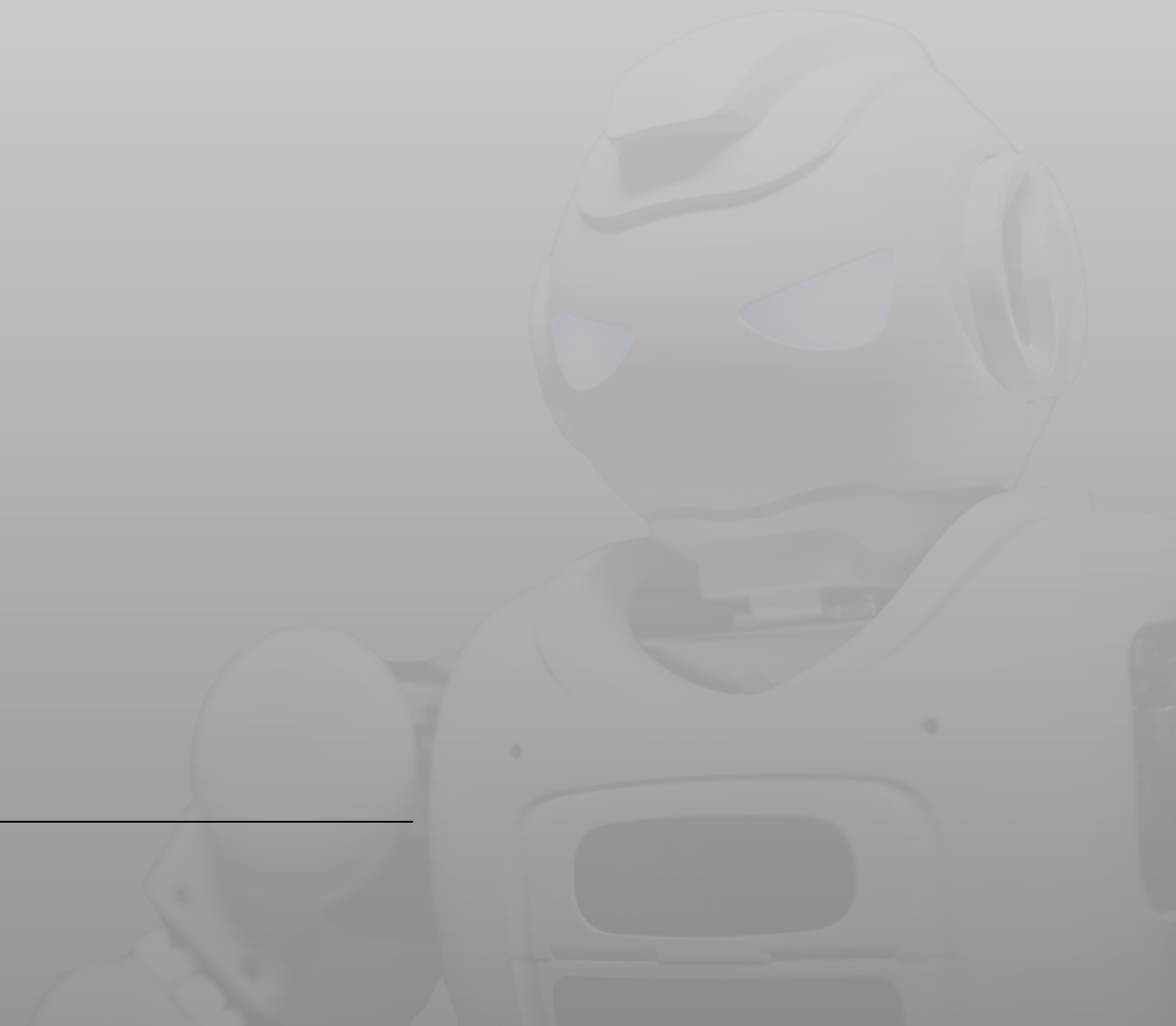


Ultimately, when something as personal as a shared laugh is generated by a machine, what do we stand to gain, and what might we lose?

---

Questions?

---





---

# Thank You