

# WHY WE ARE ALWAYS RIGHT

ECHO CHAMBERS AND POLARIZATION:  
A DILEMMA FOR DEVELOPERS



# CONTENT

**Introduction**

**Case Analysis**

**Moral Problem**

**Possible Solutions**

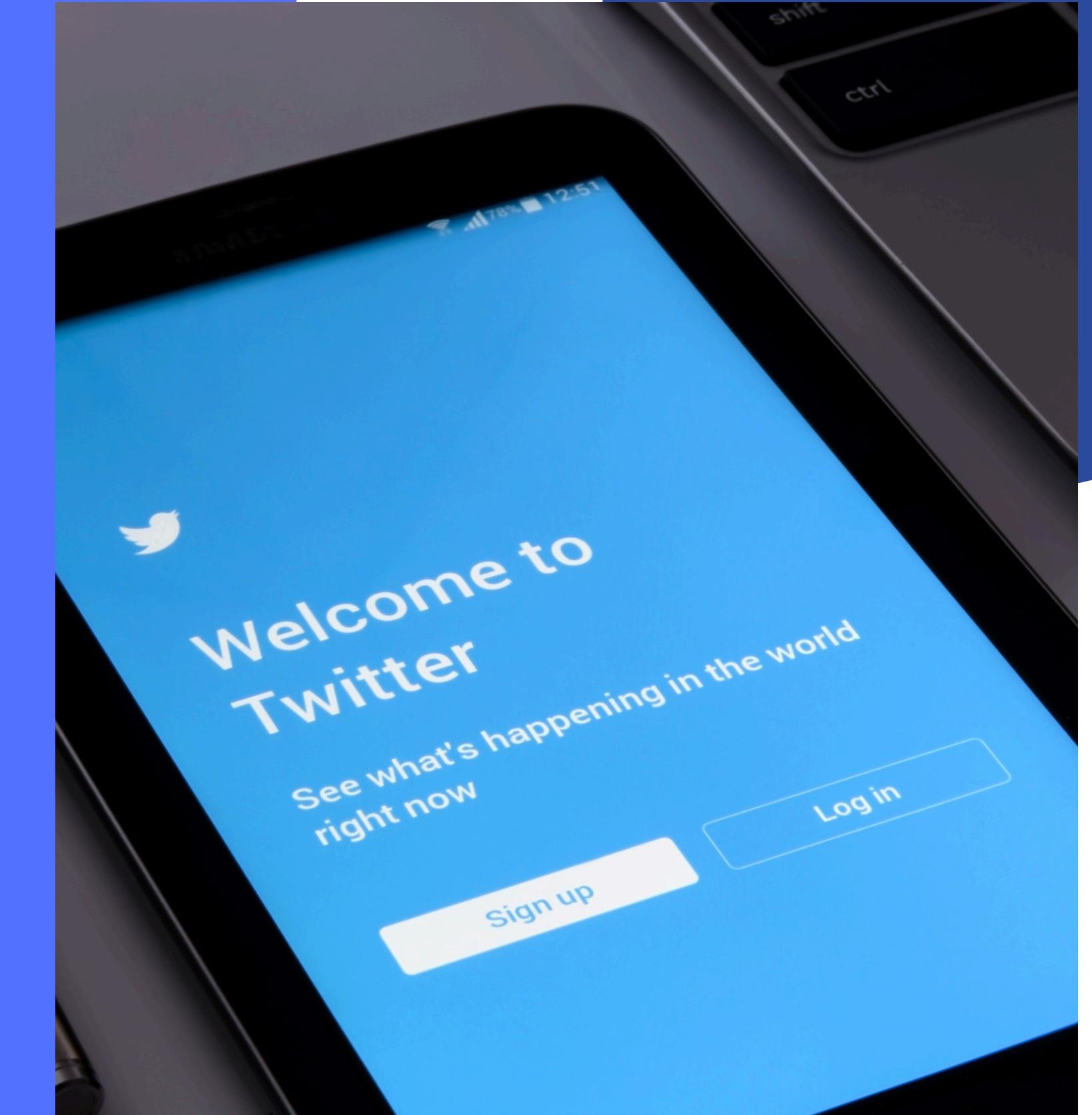
**Ethical Intuition**

**Conclusion**

**Objections and Responses**

# INTRODUCTION

## TWITTER AND LARGE LANGUAGE MODELS



# LANGUAGE PROCESSING

## Natural Language Processing

- WHAT TWITTER USES
- CONTENT CLASSIFICATION
- CONTENT MODERATION
- CONTENT RECOMMENDATION
- IMMEDIATE WORD CONTEXT

## Large Language Models

- ADVANCED AI ALGORITHM
- TRAINED ON EXTENSIVE TEXTUAL DATA
- TEXT GENERATION, PREDICTION, AND SUMMARIZATION
- LARGER TEXTUAL CONTEXT
- SUPERIOR CONTENT RECOMMENDATIONS

# CASE ANALYSIS

CONTENT RECOMMENDATIONS, ECHO  
CHAMBERS, AND POLARIZATION

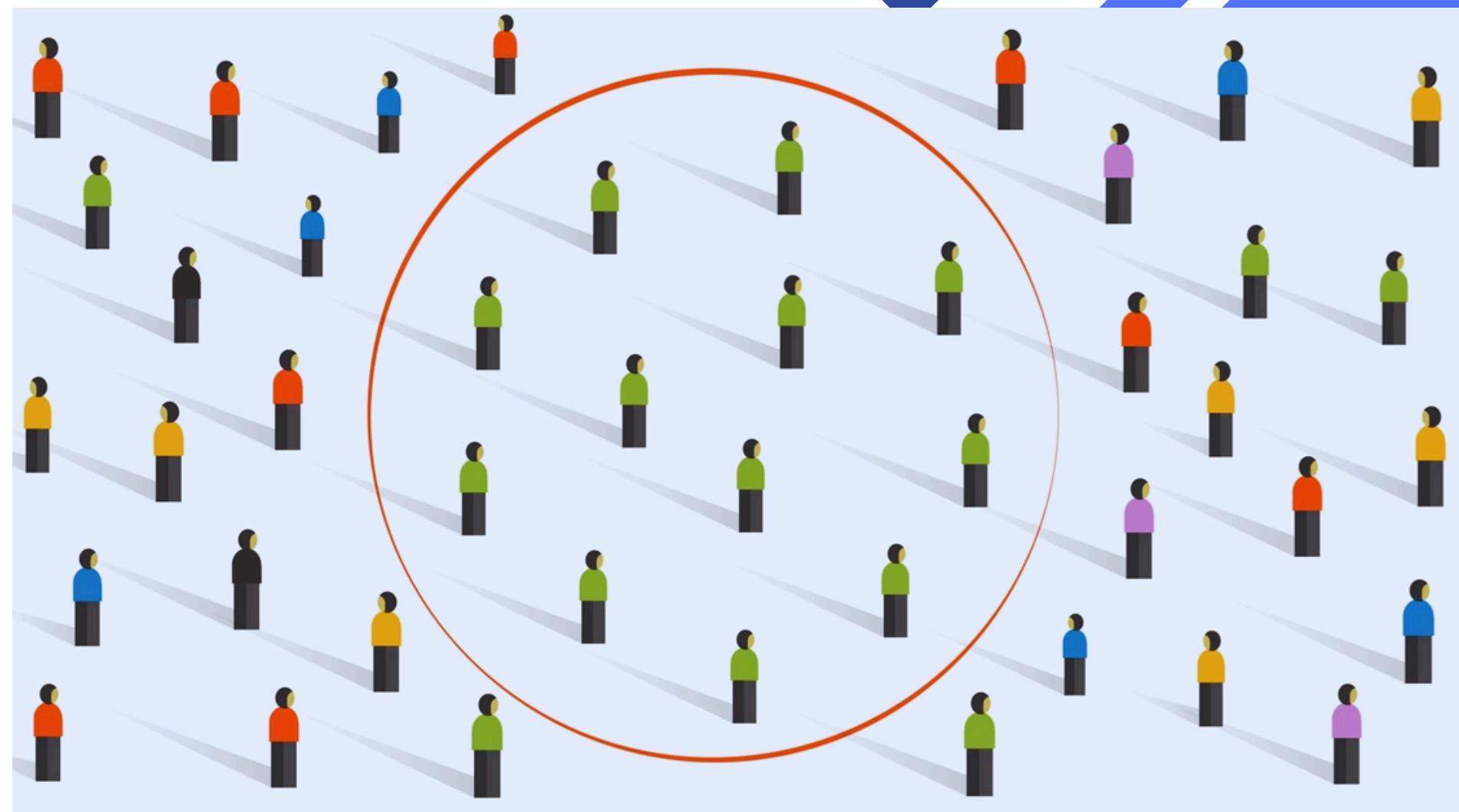


# Concerns

1

## ECHO CHAMBERS

- A *like-minded community*
- Curtails exposure to other perspectives & opinions
- Potential contribution to online extremism
- Makes you think you are always right 😊

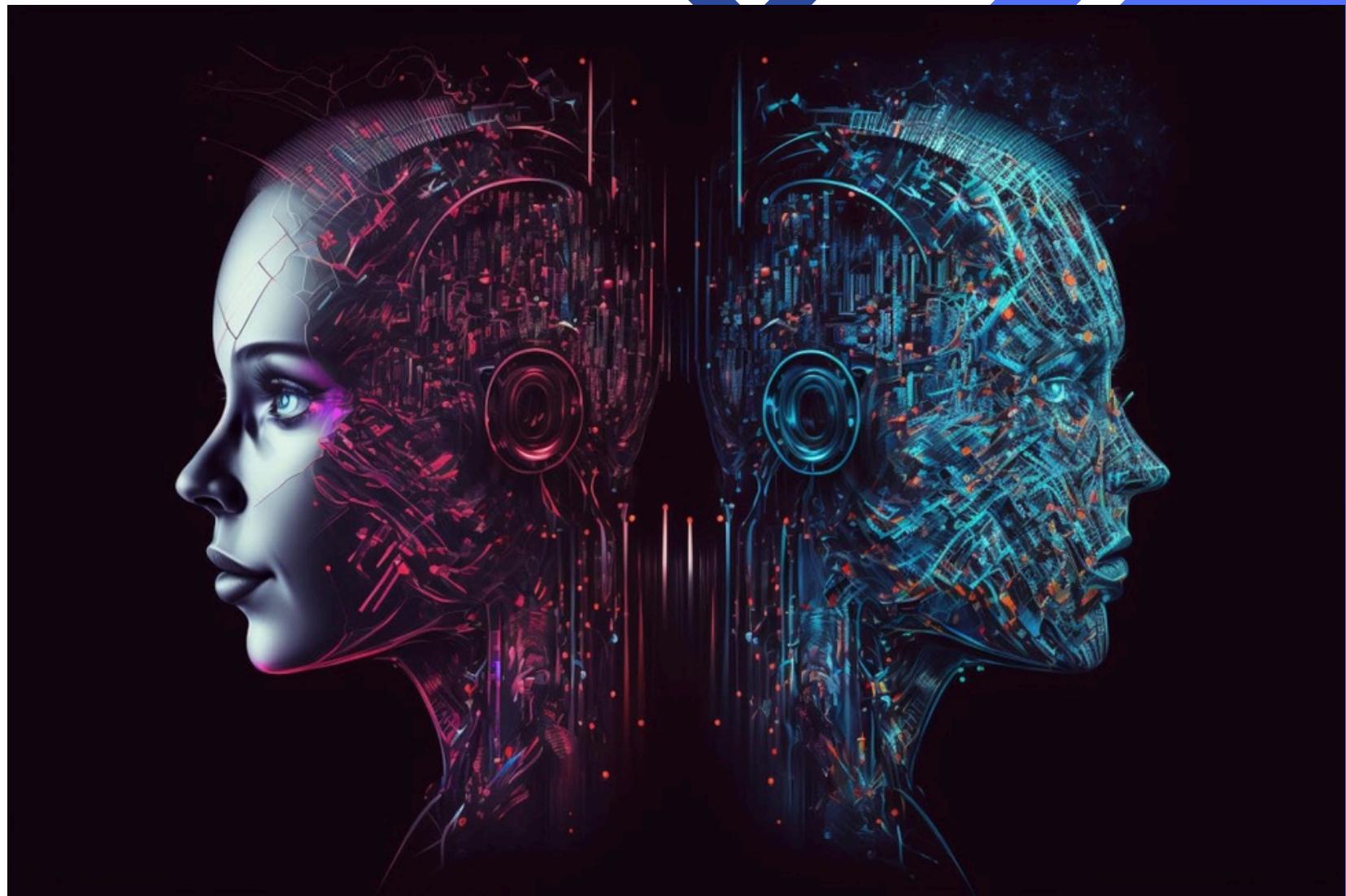


# Concerns

2

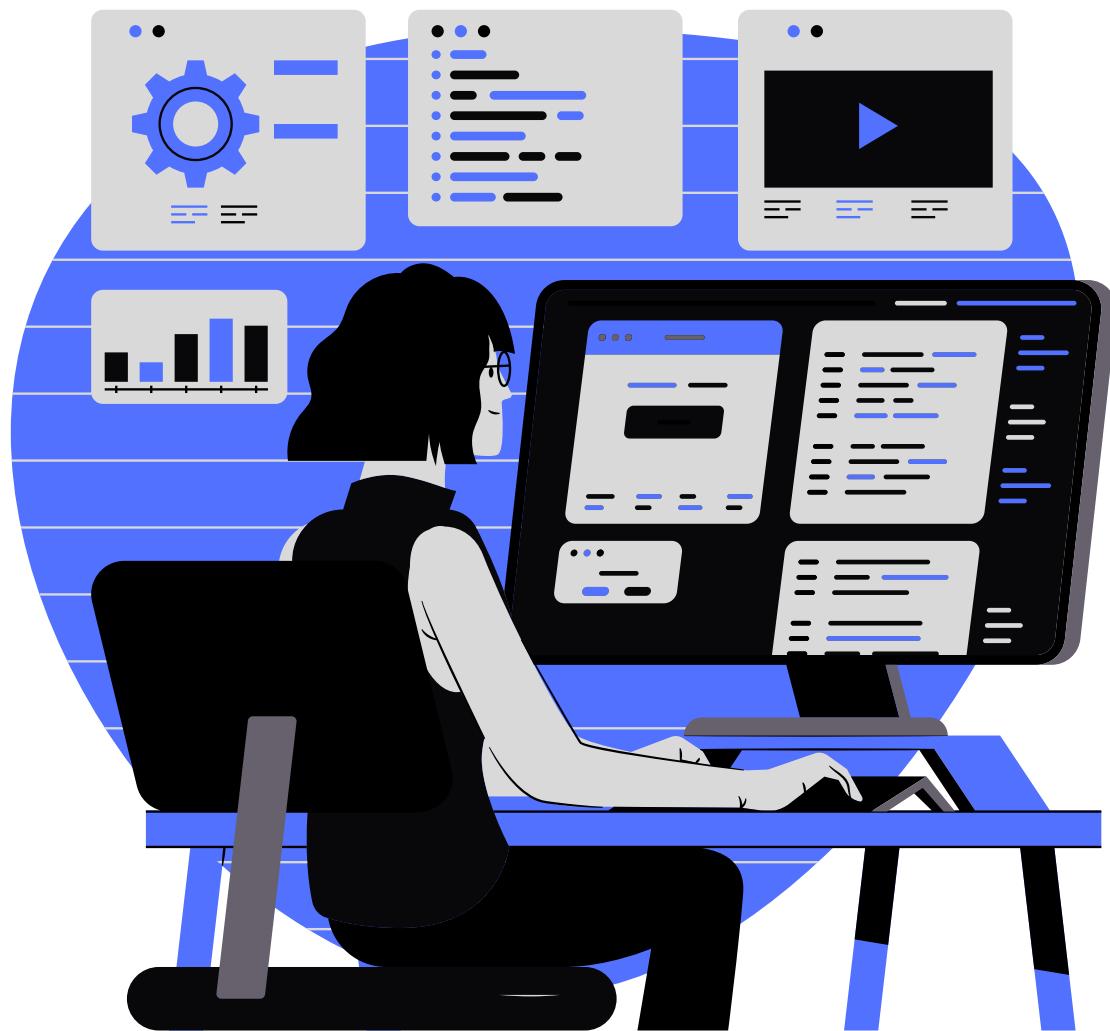
## AI HALLUCINATIONS

- Generates text that seems meaningful but on closer inspection is erroneous.
- Using chatbots may spread misinformation



# STAKEHOLDER

DEVELOPERS:  
EMPLOYEES IN CHARGE OF THE  
TWITTER ALGORITHM



How can developers strike a balance between:

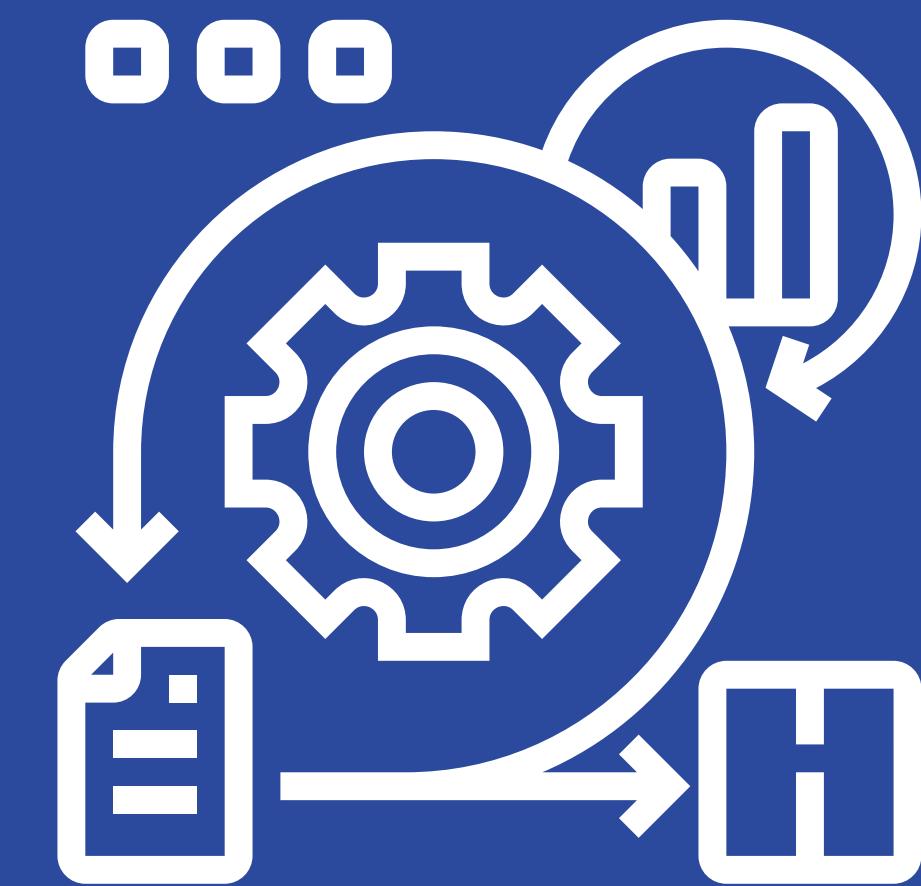
- Ethical obligations towards society
- Strategic goals of Twitter
- Their own career ambitions

# **MORAL PROBLEM**

“SHOULD TWITTER DEVELOPERS IMPLEMENT LARGE LANGUAGE MODELS (LLMs) FOR CONTENT RECOMMENDATIONS, CONSIDERING THE RISK OF EXACERBATING SOCIAL POLARIZATION AND SPREADING MISINFORMATION AMONG ITS USERS?”

# POSSIBLE SOLUTIONS

OVERCOMING CONCERNS :)



# MULTIFACETED APPROACH

## ALGORITHM

*Transparency: help users understand how content is selected for their feed*

*Diversity: prioritizing diverse perspectives*

## MEDIA LITERACY

*Educational initiatives: critical thinking and fact checking*

*Incentivize healthy dialogue: reward positive engagement and discourage abusive behavior*

## CONTENT RECOMMENDERS

*Collaborative filtering techniques*

*User-controlled content experience*

**REDUCES ECHO CHAMBERS AND POLARIZATION, AND PROMOTES USER AGENCY**

# ETHICAL INTUITION

UTILITARIANISM



# A CHOICE

## SOLUTION 1

*Not implementing Large Language Models*

### ADVANTAGES

- *Protects decisional privacy*
- *Attracts users who prioritize data privacy*

### CONCERNS

- *Twitter may lag behind competitors and lose customers.*
- *Measures to mitigate Echo Chambers still need to be put in place*

## SOLUTION 2

*Implementing Large Language Models along with the proposed possible solutions*

### ADVANTAGES

- *User satisfaction: better content recommendation and customer services*
- *Twitter: saves time on customer service, more businesses use Twitter for Advertising*

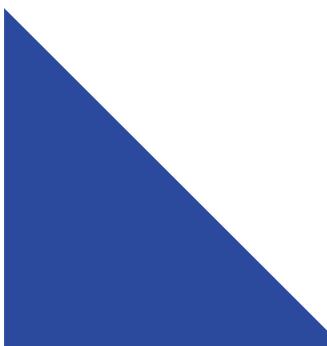
### CONCERNS

- *Does not eliminate the risk of harmful content generation or AI hallucinations completely.*

# CONCLUSION

OUR ADVICE TO DEVELOPERS





**ANY  
OBJECTIONS?**

**THANK YOU**

# REFERENCES

1. ALKAISI, HUSSAM, AND SAMY I. MCFARLANE. "ARTIFICIAL HALLUCINATIONS IN CHATGPT: IMPLICATIONS IN SCIENTIFIC WRITING." CUREUS, VOL. 15, NO. 2, FEB. 2023,
2. CINELLI, MATTEO, ET AL. "THE ECHO CHAMBER EFFECT ON SOCIAL MEDIA." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, VOL. 118, NO. 9, FEB. 2021,
3. ROUSE, MARGARET. "WHAT IS LARGE LANGUAGE MODEL (LLM)? - DEFINITION FROM TECHOPEDIA." TECHOPEDIA.COM, 20 MAR. 2023,
4. IBM. "WHAT IS NATURAL LANGUAGE PROCESSING? | IBM." WWW.IBM.COM, 2023,
5. "HOW TWITTER IS USING NLP TO IDENTIFY HATEFUL CONTENT." RE•WORK BLOG - AI & DEEP LEARNING NEWS, 2 MAR. 2022,
6. "USING DEEP LEARNING AT SCALE IN TWITTER'S TIMELINES." BLOG.TWITTER.COM,
7. KUMAR, AJITESH. "LARGE LANGUAGE MODELS: CONCEPTS & EXAMPLES." DATA ANALYTICS, 1 MAY 2023,
8. HONG, SOUNMAN. "POLITICAL POLARIZATION ON TWITTER: SOCIAL MEDIA MAY CONTRIBUTE TO ONLINE EXTREMISM." HARVARD.EDU, 2016,
9. GARIMELLA, KIRAN, ET AL. "POLITICAL DISCOURSE ON SOCIAL MEDIA: ECHO CHAMBERS, GATEKEEPERS, AND THE PRICE OF BIPARTISANSHIP." ARXIV.ORG, 2018,
10. "TWITTER'S AD REVENUE WILL FALL DRASTICALLY THIS YEAR, ACCORDING TO ESTIMATES." MARKETING BREW, ACCESSED 30 MAY 2023.