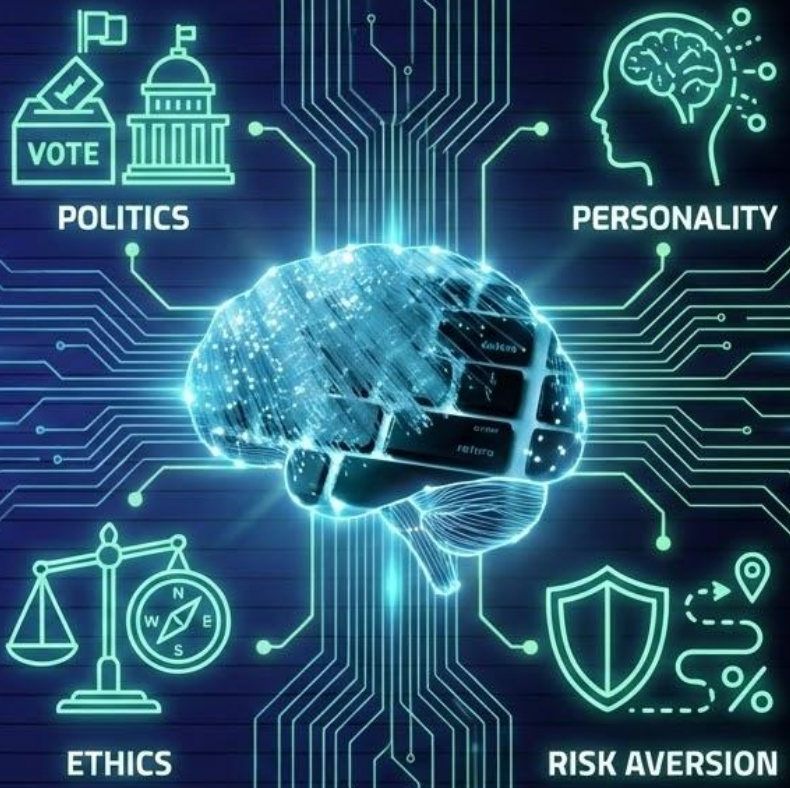

Scammer University
Final Presentation
AI
By: Kyle Pasieniuk, Ruiqi Sun,
Aidan Flanigan, Yi Wu, Yuxin Zhao,
Kyusub Shin, and Kyungsu Noh

TABLE OF CONTENTS

MISSION STATEMENT

This project analyzes the decision-making tendencies of Large Language Models (LLMs) **in politics, personality, ethics, and risk aversion.** The results allow users to better understand each model's diverse tendencies, the influence of language on responses, and the overall reliability of the models.




01
OUR SETUP


TEST SUBJECTS

POLITICS / PERSONALITIES

8Values

Traditional questionnaire containing 70 different political questions.

><

Myers-Briggs

Extroversion vs introversion, sensing vs intuition, thinking vs feeling, and judging vs perceiving

ETHICS / RISK PREFERENCE

8 Categories

- Lying
- Animal/Environment
- Race/Gender
- Health
- Age
- Theft
- Doomsday
- Other

7 Scenarios

- Low-Probability High-Reward Lotteries
- Investment Decisions
- Insurance
- Job and Income Uncertainty
- Medical Decisions
- Gambling Scenarios
- Loss-Recovery Decisions

METHODOLOGY FOR ANALYSIS

SCORING & VALUES

POINT ASSIGNMENT

Each question contributes points to four axes: econ, dipl, govt, and scty. Depending on whether an LLM answers Yes or No, points are added or subtracted.

BINARY

- Probability of “Yes” [=1]
- Probability of “No” [=0]
- Errors/Refusals [=-1]


03
ANALYSIS &
FINDINGS


04
LIMITATIONS

LIMITATIONS

Model Tiers

Use of cost-efficient or mini models opposed to full scale models

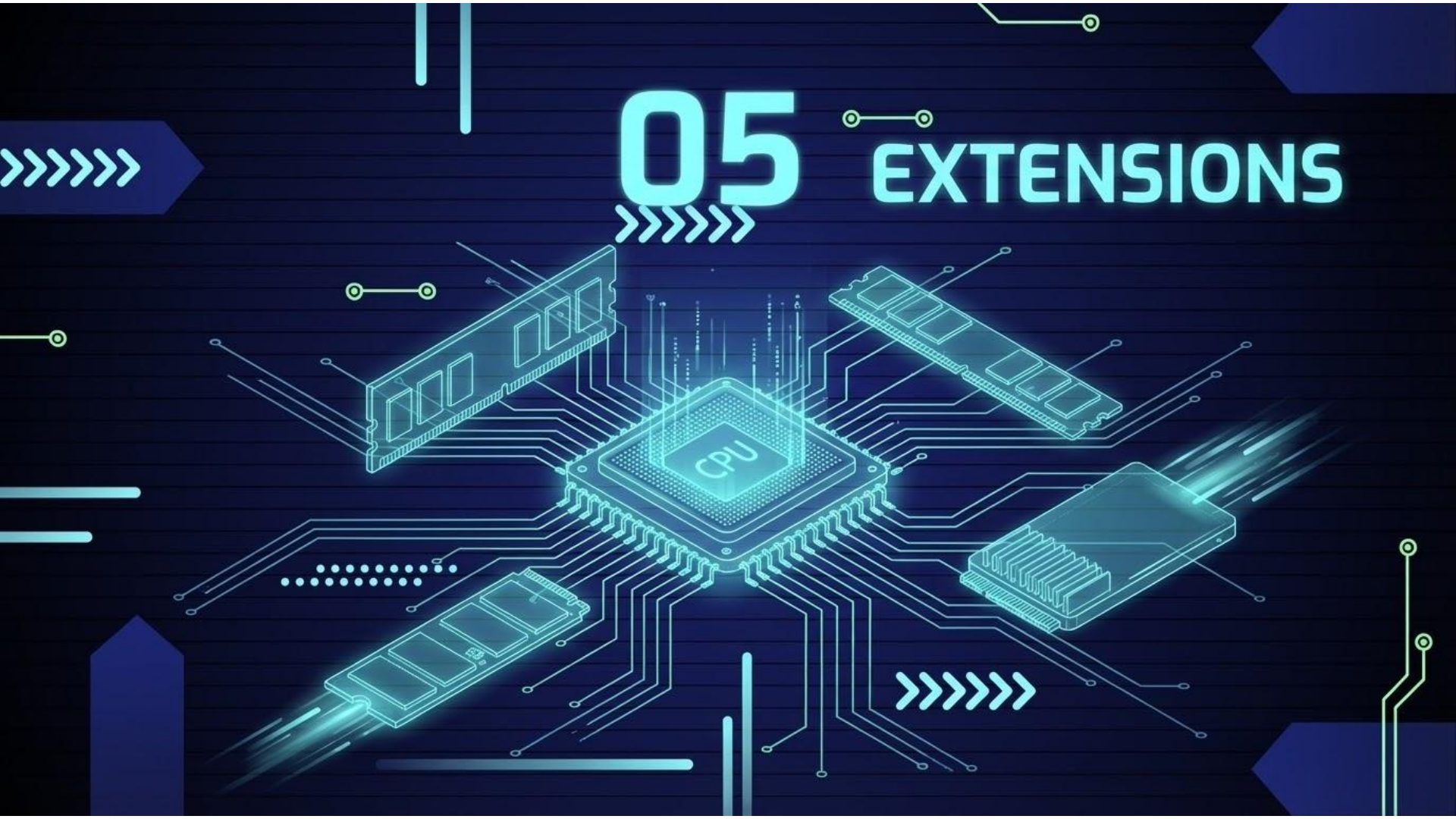
Ambiguity in Binary Constraints

Forcing yes/no outputs may obscure a model's true reasoning capabilities

Response Instability

High standard deviations on specific topics

05 EXTENSIONS



Extensions

Comparative Analysis to Flagship Models

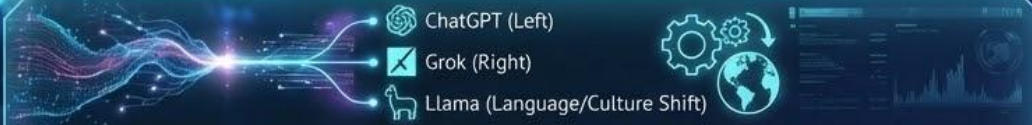
Incorporate more languages and test whether tendencies remain consistent in models with deeper reasoning

Diversification of Testing Domains

Expand to other domains such as finance, law, and cultural norms


06
SUMMARY

CONCLUSIONS



AI is not neutral. Our analysis reveals a clear political divide (e.g., ChatGPT leans Left, Grok leans Right). Furthermore, Llama demonstrates that 'Language is Culture' by shifting its stance from progressive in English to nationalistic in Korean.

As LLMs become increasingly prevalent in human production and daily life, their decision logic and behavioral patterns may potentially influence their users, and users of different languages may receive different outcomes. Although we find that most LLMs are progressively internationalist, the heterogeneity of the model regarding the prompt language might cause users in different countries, under its influence, to evolve toward different social trends.


THANK YOU