# Excitements and Concerns in the Post-ChatGPT Era: Deciphering Public Perception of AI through Social Media Analysis

Weihong Qi[1], Jinsheng Pan[2], Hanjia Lyu[2], Jiebo Luo[2]

[1]*Department of Political Science, University of Rochester*
[2]*Department of Computer Science, University of Rochester*

## Motivation


As AI continues to advance and integrate into society, it's essential to gain insights into how public perceive it. We aims to study public perception towards AI, particularly for generative AI.

- ❑ RQ1: What specific topics characterize the discussion of AI on Reddit? How do topics vary across subreddits?
- ❑ RQ2: What is the prevailing sentiment/opinions surrounding the most discussed topics, and do these sentiments differ among subreddits?

## Methods

**Data Collection**: We employ Reddit PRAW API with curated keywords to extract relevant subreddits. Then we retrieve the latest Reddit comments..

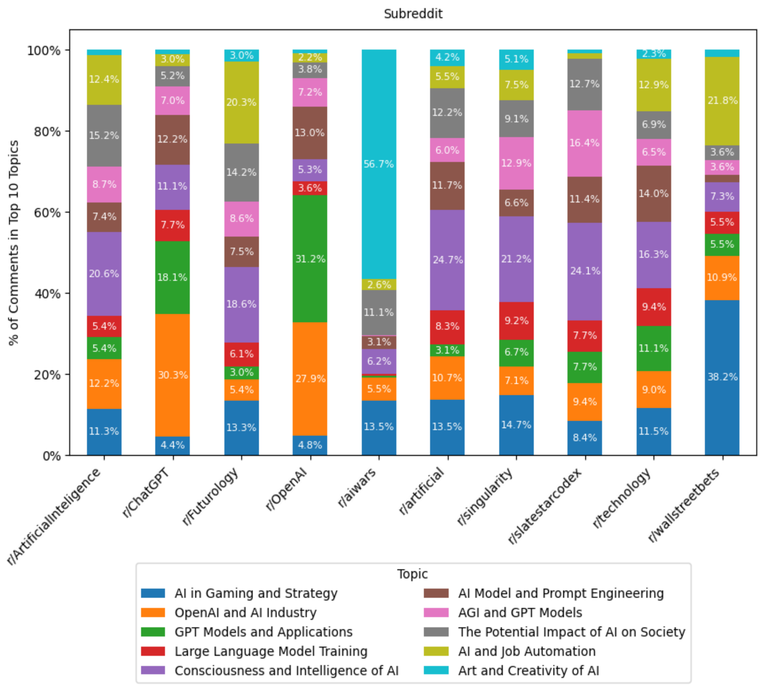
**Topic Modeling:** BERTopic [1] is used. For model architecture, we select all-MiniLM-L6-v2 as embedding component and DBSCAN as clustering component.

**Sentiment classification:** We leverage GPT-3.5-turbo [2] in zero-shot prompting, assigning labels (Positive, Negative, Neutral) to each Reddit Comment.

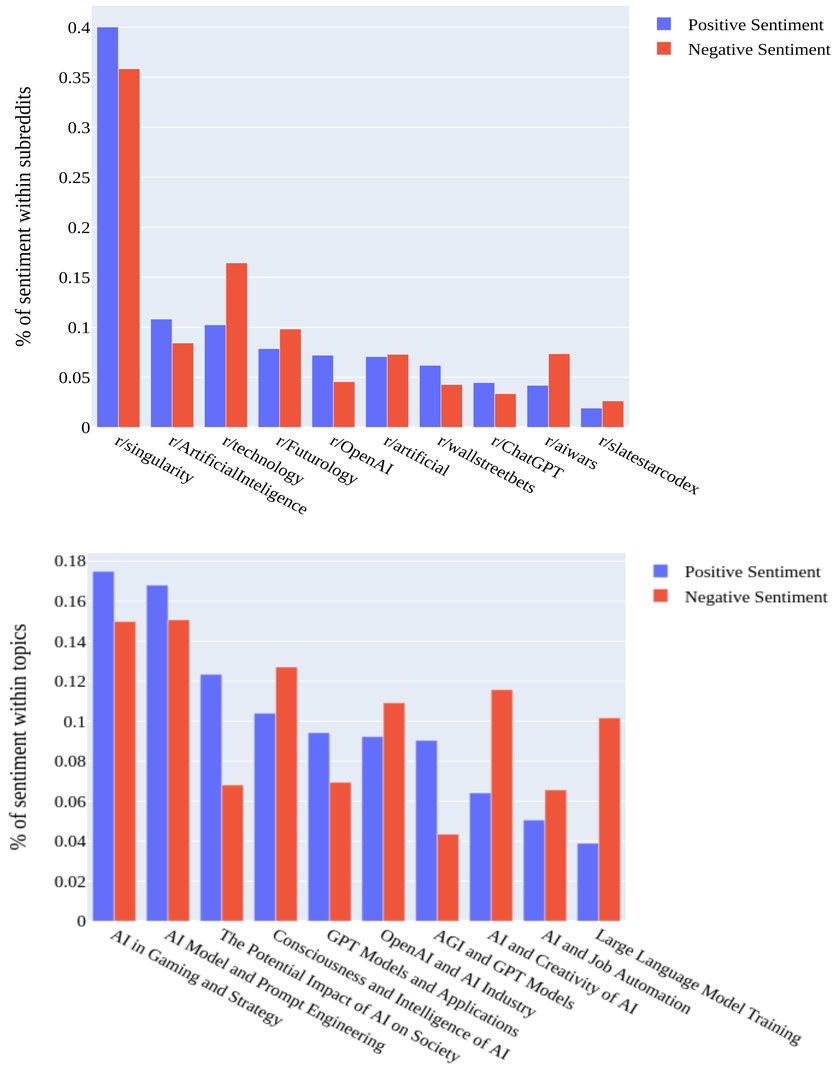
**Regression:** To explore whether perceptions of AI differ between tech-centric and non-tech groups, we run LIWC [3] and then regress emotion and tone on if comment belongs to tech group.

$$\text{LIWC}_{Attr} = \alpha_0 + \alpha_1 \cdot Tech + \epsilon$$

## RQ1




## RQ2




| | (1) Positive tone | (2) Negative tone | (3) Positive emotion | (4) Negative emotion | (5) Prosocial | (6) Conflict |
|---|---|---|---|---|---|---|
| *Tech* | 0.173** | -0.0160 | 0.0562* | 0.102*** | 0.0632* | 0.0153 |
| | (0.0630) | (0.0519) | (0.0282) | (0.0297) | (0.0287) | (0.0205) |
| Reference Group: Non-tech subreddits | | | | | | |
| *N* | 18764 | 18764 | 18764 | 18764 | 18764 | 18764 |

**Note:** This table presents the estimation coefficients of the regressions. Standard errors of each coefficient are in parentheses. The p-values indicating significance at the 90%, 95%, and 99% confidence levels have been adjusted using the Bonferroni correction.
$^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

## References

[1] *Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.*

[2] *Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. Advances in Neural Information Processing Systems, 35, 27730-27744.*

[3] *Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology, 29(1), 24–54. https://doi.org/10.1177/0261927X09351676*