National College of Ireland

National College of Ireland

Project Submission Sheet

Student Name:	Sahana Lakshmikanthaiah, Aarthi Vijayaragavan, Lakshmi Meena Manivannan
Student ID:	X24172812, x23438533, x23426918
Programme:	MSCDAD_C **Year:** 2025
Module:	Analytics Programming and Data Visualization
Lecturer:	Furqan Rustam
Submission Due Date:	April 25, 2025
Project Title:	Analysing online discussion about AI and ethics
Word Count:	2933

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
<u>ALL</u> internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature:	Sahana, Aarthi, Lakshmi Meena
Date:	25th April 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. **Late submissions will incur penalties.**
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Analyzing Online Discussions About AI and Ethics

1st Aarthi Vijayaragavan
Data Analytics
National College of Ireland
Dublin, Ireland
x23438533@student.ncirl.ie

2nd Lakshmi Meena Manivannan
Data Analytics
National College of Ireland
Dublin, Ireland
x23426918@student.ncirl.ie

3rd Sahana Lakshmikanthaiah
Data Analytics
National College of Ireland
Dublin, Ireland
x24172812@student.ncirl.ie

***Abstract*—This study analyzes online discourse on AI ethics on Twitter and Reddit and sentiment trends on both sites. Reddit data, sourced from more current discussions around AI, was extracted through web scraping, while Twitter data is sourced from previous tweets. It cross-compares sentiment before and after 2024 and observes a shift from the early optimism for the promise of AI to growing worries regarding its impact on society, regulation, and algorithmic fairness. These trends are reflected in the evolving public sentiments about AI. The findings provide valuable information for guiding AI policy, regulation and scientific research, emphasizing the need for ethical considerations and responsible development of AI technology.**

***Index Terms*—Artificial Intelligence, AI Ethics, Sentiment Analysis, Social Media Analysis, Reddit, Twitter, Data Analysis, Text Mining, Temporal Trends, Public Opinion, Ethical Concerns, Web Scraping**

I. Introduction

Artificial intelligence (AI) is racing at lightning speed to transform the world today, sectors from healthcare to finance, education, and beyond. As AI systems take on more influence in decision making, society's operations, and individual behavior, a simultaneous phenomenon of growing interest and concern from the public about AI ethics has ensued. These ethics deal with data privacy, algorithmic fairness, transparency, and accountability and are crucial considerations in using AI responsibly and democratically. Social media discussion platforms such as Twitter and Reddit are useful sources of user-contributed content that capture public opinion and debate regarding many new technologies, including AI.

This project investigates the public discussion on the ethics of AI by analyzing posts and comments collected from Reddit and Twitter. Using web scraping methods, natural language processing (NLP), and sentiment analysis models, we aim to quantify and interpret the shifting sentiment and focus of these online discussions. The research not only captures the current state of public opinion but also follows significant long-term shifts in attitudes, particularly to big technology advances and policy actions. Ultimately, it hopes to enhance the capacity to examine how the state of public opinion on AI ethics shifts and help to guide effective responsive policy frameworks.

II. Related Work

Featuring impacts from computer science, law, philosophy, and social science, AI ethics is becoming a more multidisciplinary field. Researchers have investigated topics like the moral responsibilities of AI programmers, the dangers of self-governing systems, and the effects of automation on society. In their mapping of the main ethical issues with algorithmic systems, Mittelstadt et al. (2016) [2] emphasized concepts like accountability, transparency, and fairness.

As AI became more prevalent in daily life, public opinion toward technology moved from initial optimism to increasing skepticism, as demonstrated by Fast and Horvitz (2017) [1]. Few studies have contrasted platforms or looked at how sentiment changes over time, despite the fact that social media has been utilized to assess public opinion on emerging technology.

This study builds on previous work by combining data from two widely used platforms—Reddit and Twitter—and conducting a temporal analysis that distinguishes between pre- and post-2024 discussions. This approach provides a broader and more nuanced understanding of how perceptions of AI ethics are influenced by time, platform, and external events.

III. Data processing methodology

This section explains the technical framework and methodological approach used to analyze public opinion on the ethics of artificial intelligence (AI) using data from Twitter and Reddit. An extensive understanding of public opinion patterns over time has been made achievable by the pipeline, which consists of data collection, preprocessing, sentiment analysis, visualization, and data storage.

A. Data Sources and Justification

To ensure diversity in discourse and temporal coverage, two primary datasets were utilized:

- **Reddit Web-Scraped Dataset**: A specific web scraping process was used to collect data from 2023 to 2025, focusing on subreddits related to technology and artificial intelligence. Ongoing discussions and fluctuating public opinions regarding AI ethics have been captured in this source.
- **Twitter Dataset**: This dataset provides historical context for comparing the emotion of conversations around AI ethics from 2020 to 2022.

Justification:

- **Temporal Analysis:** Longitudinal tests of fluctuations in public opinion are made possible by a combination of both historical (Twitter) and contemporary (Reddit) data.

- **Platform Diversity:** The analysis has been improved by using both Reddit and Twitter, which capture a wider range of demographic information and methods of interaction.

B. Data Processing Activities

The data underwent several preprocessing stages to ensure quality and consistency for downstream analysis:

- **Data Cleaning:** Duplicate and redundant items are eliminated. Text standardization was performed by URL filtering, emoji removal, lowercasing, and punctuation stripping.
- **Natural Language Processing (NLP):** To eliminate noise, tokenization and stopword removal were applied. Lemmatization was used to combine word forms into their most basic versions (e.g., "running" to "run").
- **Metadata Formatting:** To facilitate trend analysis and temporal segmentation, posts were timestamped.
- **Sentiment Annotation:** Compound sentiment scores were assigned and items categorized as positive, neutral, or negative using the VADER sentiment analyzer, which is well-suited for short-form social text.
- **Segmentation and Visualization:** For contrast, the dataset was divided between pre-2024 and post-2024 periods. Thematic trends were explored using visual tools such as word clouds, time series plots, and sentiment distribution charts.

C. Data Analysis and Algorithms

The analysis was centered around understanding the public's emotional response and thematic focus:

- **Sentiment Analysis:** Utilised VADER for assessing social media posts' sentiment based on specific criteria. This allowed grouping into sentiment bands—positive, neutral, and negative.
- **Temporal Trends:** An overview of changes in public opinion, especially in response to global or regulatory AI events, was made possible through year-wise sentiment aggregation.
- **Theme and Keyword Extraction:** Developed word clouds and performed word frequency analysis to identify recurring issues such as accountability, fairness, and societal impact.

D. Data Storage

To efficiently manage the collected social media data, MongoDB was used as the primary storage solution.

- **MongoDB:** Chosen for its flexibility in managing unstructured and semi-structured data formats like JSON, which complements social media content. It enabled smooth archiving and retrieval of massive amounts of Twitter and Reddit data, retaining sentiment annotations, text content, and post metadata. MongoDB also supported effective querying for tasks such as keyword matching, sentiment classification, and time period filtering.

  This storage system provided the document-oriented structure and scalability required to handle user-generated content of varying lengths across multiple platforms.

E. Technology Stack

The project leveraged a Python-based stack tailored for natural language processing and data visualization:

- **Python:** Provided core support through its rich ecosystem for data science and NLP.
- **Pandas & NumPy:** Facilitated high-performance data wrangling and numeric operations.
- **NLTK & VADER:** Used for text preprocessing and sentiment classification.
- **Matplotlib & Seaborn:** Powered the generation of insightful visualizations including distribution and time series plots.
- **WordCloud:** Visualized frequent terms to aid in identifying thematic patterns.
- **Jupyter Notebook:** Served as the development environment for interactive and iterative analysis.

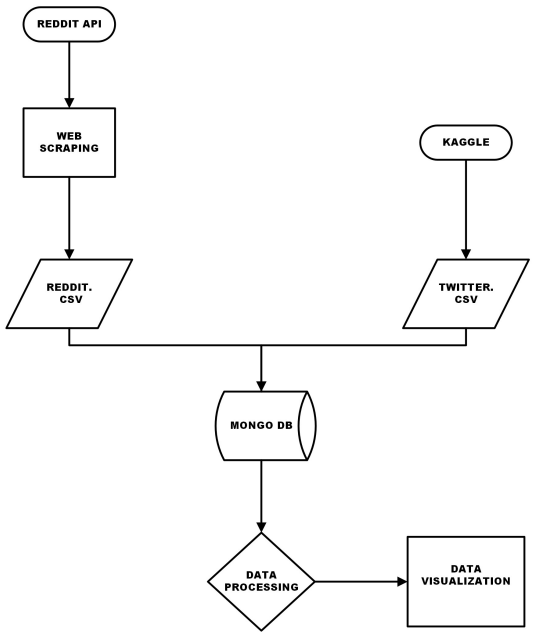

Fig. 1. Overall Data Flow Diagram

The project pipeline is shown in the data flow diagram (Figure 1), which begins with the collection of data from Kaggle and web-scraped Reddit articles. The data is saved in MongoDB for effective management of unstructured content after the text has been cleaned and standardized. To demonstrate public opinion and new developments in AI ethics, it is further processed using natural language processing (NLP) and sentiment analysis techniques. The outcomes are then displayed using tools such as word clouds and time-series charts.

F. Findings and Insights

The analysis highlighted significant shifts in public discourse regarding AI ethics:

- **Before-2024 (Twitter):** Most conversations were positive and focused on the advantages of AI, automation, and innovation. The excitement about the promise of technology was reflected in sentiment trends that leaned toward neutral and positive remarks.
- **After-2024 (Reddit):** A noticeable shift toward caution and critical participation emerged. Transparency, data privacy, moral dilemmas, and the necessity of regulation became increasingly prominent topics. Sentiment grew more politicized, with a rise in negative posts questioning the unconstrained development of AI.
- **Emerging Themes:** Algorithmic fairness, employment displacement, autonomous decision-making, and environmental sustainability of AI systems were among the major recurring topics. Discussions increasingly included calls for global governance of AI ethics and highlighted broader societal implications. These findings point to a growing public awareness of the need for responsible AI development and governance. This study provides valuable insights for AI ethicists, developers, and policymakers by reflecting evolving public sentiment.

IV. Data Visualization Methodology

The study employed a hybrid data visualization approach using Python libraries and Power BI to analyze the sentiment trends over time on different platforms like Twitter and Reddit. Plotly, Matplotlib, and Seaborn were the Python libraries used for custom visualizations, interactive plots, and in-depth analysis. These tools allowed for the creation of high professional figures and supported data preprocessing tasks. Plotly, in particular, enabled the creation of interactive chart layouts inside Jupyter Notebooks, making the analysis reproducible and customizable.

Besides Python, Power BI was used to create an interactive dashboard with bar charts, pie charts, and area graphs to suitably communicate findings from the sentiment analysis. The dashboard emphasized displaying post quantities over years, breakdown of sentiments by platform, general sentiment distribution, and time-based sentiment trends. Features of interactivity like tooltips, slicers, and filters greatly improved the user experience through enabling dynamic exploration of data by year, sentiment type, or platform.

Bar charts were used to show year-to-year change in post counts by positive, negative, and neutral sentiments, while pie charts illustrated the average sentiment ratios in the dataset. Stacked horizontal bar charts contrasted sentiment distribution across platforms (e.g., Reddit and Twitter), and area graphs depicted sentiment evolution over time.

To further analyze results, word clouds were generated from cleaned text data using Python to determine words most frequently used in discussions of AI ethics. Separate word clouds were generated for each sentiment and platform. Visualizations provided patterns and shifts in discourse, particularly when comparing Reddit and Twitter.

V. Results and Evaluation

The sentiment analysis revealed a complex and evolving picture of public sentiment toward AI usage and ethics. Across both platforms (Twitter and Reddit), the data was skewed toward neutral and negative sentiments, indicating a predominance of cautious or critical views. The key findings include:

- **Platform differences**: Twitter discussions from 2020–2022 generally showed a higher proportion of positive sentiment, with users often expressing enthusiasm about AI's potential. In contrast, Reddit discussions from 2023–2025 were more negative, reflecting growing awareness and concern over AI-related issues. Before the COVID-19 pandemic, there was relatively limited discussion and adoption of AI-related technologies among general users. However, following the pandemic—especially around 2023 and onward—there was a noticeable surge in AI usage across platforms, likely driven by increased digital adoption and rapid advancements in AI tools.
- **Temporal trends**: Comparing before-2024 and after-2024 data showed a significant rise in negative and neutral sentiment. This shift aligns with global events such as the emergence of AI-generated misinformation, high-profile data breaches, and increasing public discourse about AI governance.
- **Yearly breakdown**: A sentiment peak was observed in 2024, coinciding with several regulatory announcements and media coverage of AI ethics. This indicates that real-world events directly influence public perception.
- **Top themes and concerns**: Keyword analysis showed a shift from early focus on "innovation," "automation," and "efficiency" to more recent concerns like "bias," "accountability," "regulation," and "societal harm," reflecting growing awareness of AI's ethical and societal impacts.

These results suggest that public discourse is shifting from curiosity and excitement toward concern and a demand for oversight. This shift underscores the need for proactive regulatory and ethical frameworks that address the complexities of AI.

A. *VISUAL INSIGHTS FROM PYTHON ANALYSIS*

*1) **Sentiment and Volume Shift Over Time** :* As shown in Figure 2 and Figure 3, online discussions about AI shifted significantly before and after the COVID-19 pandemic. Before 2024, sentiment was largely positive, reflecting early optimism but favorable perception of AI. Post-2024, neutral and negative sentiments increased, pointing to growing awareness of AI's real-world impacts—ethical, employment-related, and privacy concerns. Discussion volume also surged after 2023, driven by pandemic-era digital transformation and increased AI use in healthcare, education, and remote work.

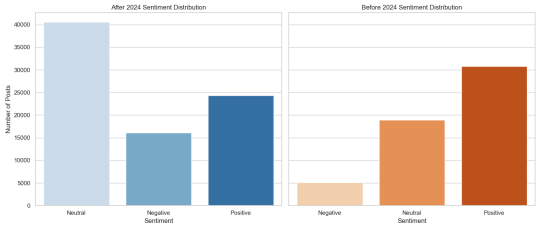

Fig. 2. AI Sentiment and Discussion Volume

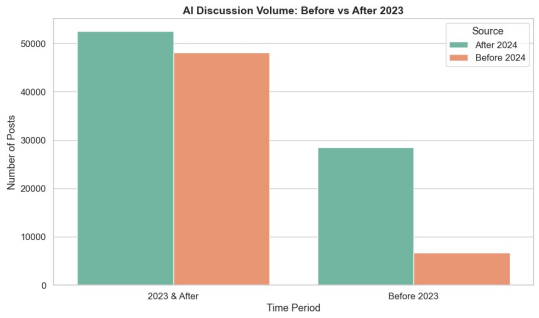

Fig. 3. AI Discussion Volume

*2) **Overall Sentiment Distribution**:* Figure 4's pie charts highlight a clear shift in sentiment before and after 2024. Before 2024, sentiment was largely positive (56.2%), with lower neutral (34.5%) and negative (9.3%) views. After 2024, positive sentiment dropped to 30.1%, neutral rose to 50.1%, and negative surged to 19.8%. This shift reflects growing public uncertainty or concern, likely influenced by technological, socioeconomic, and geopolitical developments.Also, the increasing presence of AI in everyday life and the accompanying issues, such as ethical problems and unemployment, likely contributed to driving the growth in negative sentiment. Also, widespread media reporting of AI problems has raised public awareness, resulting in a more risk-averse or negative attitude.

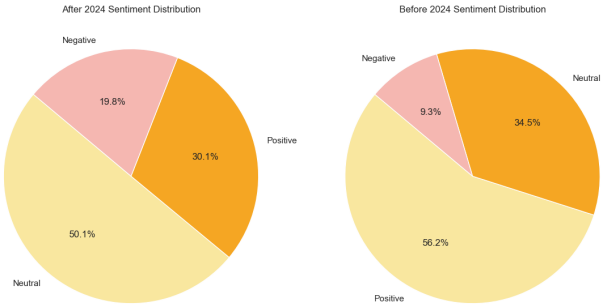

Fig. 4. AI Sentiment and Discussion Volume

Fig. 5. Overall Sentiment Distribution

*3) **Yearly Sentiment Trends**:* Figure 5 shows sentimental trends from 2020 to 2025. From 2020-2021, neutral sentiment remained steady at around 50%, positive sentiment at 30%, and negative sentiment at about 20%. In 2022-2023, there was a steep change, with positive sentiment at record highs of 45% in 2022 and 51% in 2023, likely due to global recovery efforts against the COVID-19 pandemic, economic reopening, and vaccine distributions. During this period, negative sentiment hit an all-time low at 11.5% in 2023. However, from 2024 to 2025, neutral sentiment dropped to 50%, positive sentiment to about 30%, and negative sentiment rose above 20%, showing potential socio-political or economic problems. This analysis is valuable for predicting society's response to major events, which can prove highly valuable to policymakers, analysts, and businesses.

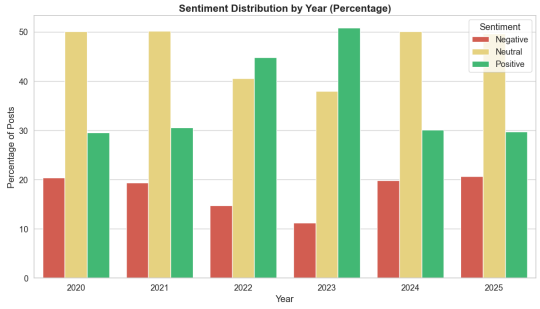

Fig. 6. Yearly Sentiment Trends

*4) **Sentiment Fluctuations Over Time** :*

- Figure 6 illustrates sentiment trends from the years 2020 to 2025. Sentiment levels remained steady with minimal activity from 2020 to mid-2022. There was, nevertheless, a sharp spike in late 2022 to early 2023, driven by positive sentiment, which was likely brought about by global events like vaccine rollouts and reopenings. After this peak, sentiment activity dropped off precipitously by mid-2023, plateauing throughout 2024–2025. Neutral sentiment dominated, while positive and negative sentiments decreased, reflecting decreased engagement and potential public exhaustion on pandemic topics post-pandemic.

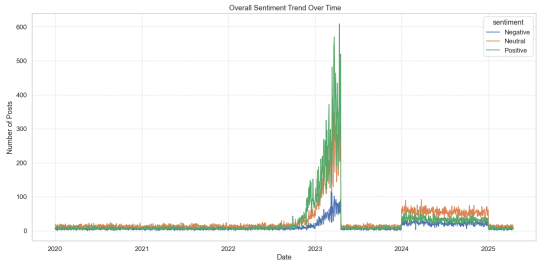

Fig. 7. Sentiment fluctuations over time

- Figure 7 illustrates from 2020 to 2025, sentiment trends regarding AI and ethics are displayed in the visual "Combined Sentiment Share Over Time (Reddit + Twitter)". Throughout, neutral emotion is predominant, but in 2023, optimism increases to a significant peak of 67.6% for positive sentiment. The smallest sentiment is still negative,

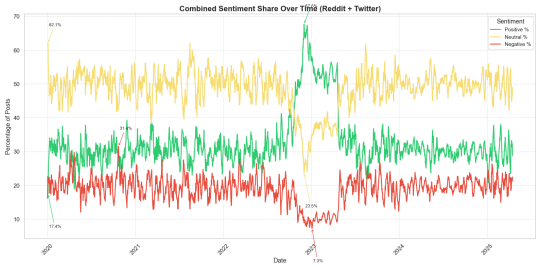

Fig. 8. Combined Sentiment over time

which fell to 7.3% in early 2023. The graph shows a significant shift in public opinion on Reddit and Twitter, particularly around 2023.

*5) **Social Media Discourse Dashboard: AI Keyword Frequency**:* The dashboard presents a comparative analysis of the top 20 most frequent words from Reddit and Twitter posts, focusing on discussions around AI and generative technologies. A word cloud and a bar chart are used to display the data from each platform, providing insights into the meaning and frequency of terms such as ”generativeai,” ”chatgpt,” ”aiart,” and ”stablediffusion.” Users can interactively explore platform-specific discourse patterns using this illustration.

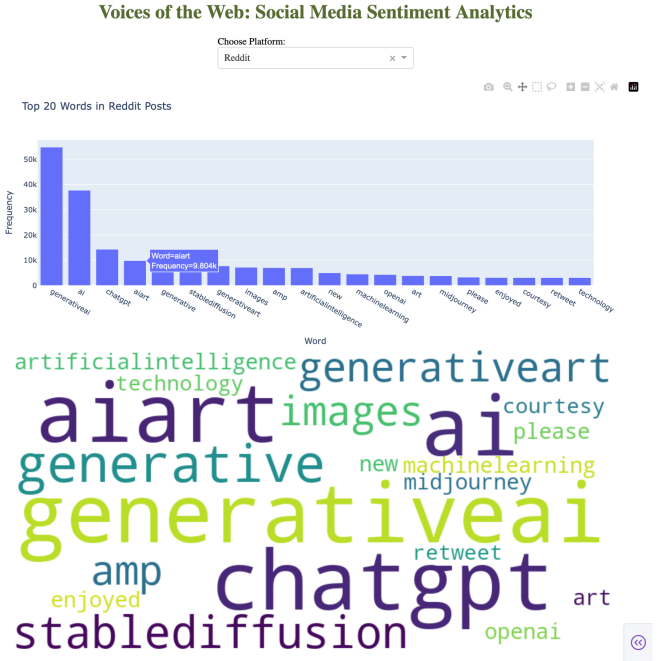

Fig. 9. Reddit Dashboard

B. *Dashboard-Based Sentiment Exploration*

The Power BI dashboard provides a comprehensive visualization of sentiment trends in online discussions surrounding Artificial Intelligence and Ethics from 2020 to 2025, drawing data from both Twitter and Reddit. A comprehensive analysis of how public opinion has changed over time and across platforms is made possible by this interactive dashboard.

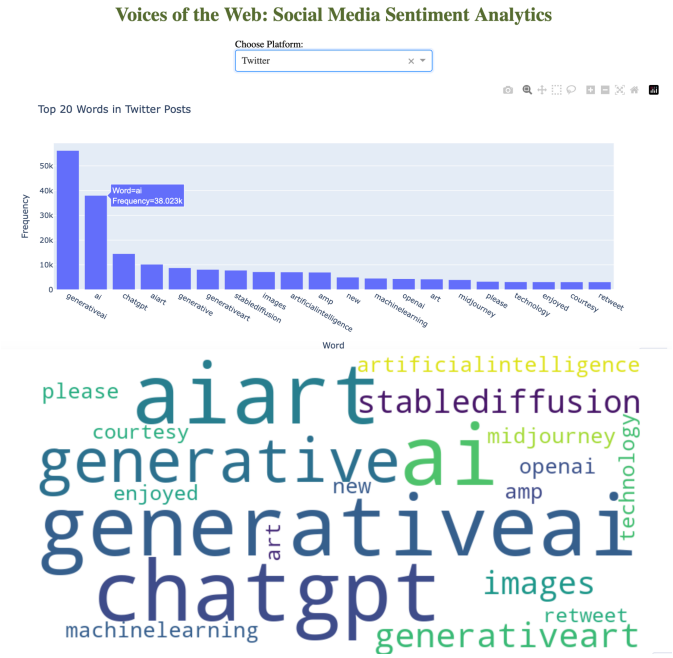

Fig. 10. Twitter Dashboard

Furthermore, post volume and positive sentiment increased significantly in 2023, indicating increased interest and hope in ethical AI discussion.

The overall sentiment distribution reveals insightful trends:

- Neutral sentiment dominates the conversation, accounting for 43.8% of total posts.
- Positive sentiment follows closely at 40.6%, indicating a generally hopeful or supportive public attitude.
- Negative sentiment remains least common at 15.6%, though still significant for tracking ethical concerns.

Platform-specific analysis shows:

- Twitter is more inclined towards positive sentiment, possibly due to broader reach and influence-driven discourse.
- Reddit, on the other hand, demonstrates a more balanced sentiment spread, reflecting diverse and critical engagement from niche communities.

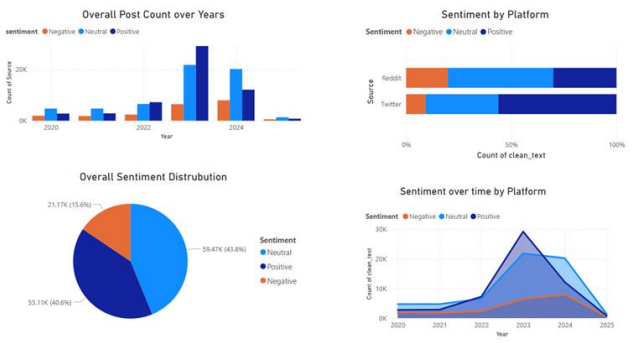

Fig. 11. Power BI Dashboard

These results demonstrate how the discussion on AI ethics is dynamic, with sentiment rising after significant events and reflecting broader social opinions. The dashboard's usability is further increased by the addition of word clouds and the top 20 most influential articles, which help in locating recurrent themes, phrases, and conversation hotspots that influenced the content of the article.

VI. CONCLUSIONS AND FUTURE WORKS

Analyzing discussions from Reddit and Twitter, this study examined public opinion on AI ethics and found that people are becoming more concerned about transparency, accountability, and regulations. The results illustrate how real-world events and changes in regulations, in addition to technical improvements, shape public opinion. Twitter and Reddit offered diverse but complimentary opinions, highlighting the need of multi-platform study in understanding the public's opinions. The rise in negativity stems from increased public awareness of AI risks—like deepfakes, misuse of generative tools, and weak regulation. As AI became more visible through viral incidents and media, online discussions reflected growing skepticism and concern.

These results demonstrate the importance of public disagreement in influencing future AI law and governance, as well as the growing desire for ethical standards in AI development. Having updated on these discussions will be crucial as AI continues to influence every aspect of daily life.

A. Future works

Future work can expand on this foundation in several keyways:

- Wider Platform Coverage: Integrating online forums, YouTube, LinkedIn, and other platforms would provide a more comprehensive understanding of popular opinion.
- Advanced Modeling: Sentiment and topic analysis may be improved by using transformer-based models like BERT or RoBERTa.
- Real-Time Tracking: Timely policy responses can be enhanced by developing live dashboards to track sentiment trends.
- Global Perspective: Including data from other languages and cultures may help us understand how people throughout the world perceive AI ethics.

By revealing how individuals view the ethical implications of AI, this study adds to the larger discussion on responsible AI and lays the groundwork for more knowledgeable and human-centered AI development.

REFERENCES

[1] Fast, E., & Horvitz, E. (2017). Long-Term Trends in the Public Perception of Artificial Intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*.

[2] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*.

[3] Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Eighth International AAAI Conference on Weblogs and Social Media*.

[4] , T. H. Davenport and R. Ronanki, "Artificial intelligence for the real world," *Harvard Business Review*, vol. 96, no. 1, pp. 108-116, 2018.

[5] G. C. Binns, "Understanding public sentiment toward artificial intelligence," *MIT Technology Review*, 2019. [Online]. Available: https://www.technologyreview.com. [Accessed: 24-Apr-2025].

[6] M. S. Zeng, "The impact of AI-generated misinformation on society," *IEEE Access*, vol. 8, pp. 12345-12352, 2020.