Ollscoil na Gaillimhe | University of Galway

Scoil Ghnó agus Eacnamaíochta J.E. Cairnes

J.E. Cairnes School of Business and Economics

Module Name and Code:
Business Analytics Major Project
MS5115

Student ID and Name:
22222651
Andres Guillermo Covilla Martínez

Project Supervisor:
Dr. Anastasia Griva

The Public Perception of AI Ethics in the Context of Twitter #ChatGPT Conversations

Declaration of Originality

Project Details

Module Code: **MS5115** Assignment Title: **M.Sc. Business Analytics - Major Project**

Group Members: **(please use BLOCK CAPITALS)**

Student ID	Student Name	Contact Details (Email, Telephone)
22222651	ANDRES GUILLERMO COVILLA MARTINEZ	a.covillamartinez1@universityofgalway.ie +353876875253

I/We hereby declare that this project is my/our own original work. I/We have read the University *Code of Practice for Dealing with Plagiarism*[*] and am/are aware that the possible penalties for plagiarism include expulsion from the University.

Signature **Date 08/06/23**

[*] http://www.nuigalway.ie/plagiarism

Abstract

This study explores public sentiment and discourse surrounding AI ethics as captured within Twitter conversations labelled with the hashtag #ChatGPT. Leveraging the power of analytics techniques such as Sentiment Analysis and The Latent Dirichlet Allocation, the study first identifies the overall sentiment toward AI ethics in the dataset. The investigation reveals a generally positive yet nuanced public perception, with a lean toward neutrality. Additionally, through focused analysis on specific ethical keywords such as "bias", "transparency", "accountability", "privacy", "discrimination", "responsibility", and "job", a slightly enhanced positive sentiment emerges. The research further unveils key topics and ethical concerns prominent in the public discourse through Twitter and explores their sentiment and correlation with the sentiment revealing a generally positive perception. A descriptive analysis then unfolds the demographic and geographic distribution of these opinions, highlighting the diversity of perspectives. The study concludes with broad practical and theoretical implications, reinforcing the importance of dialogue, inclusivity, and transparency in AI ethics discussions and the positive sentiment of a forecasted implication of AI through the eyes of the general public. It also provides a springboard for future research, recommending the exploration of other social media platforms, deep-diving into demographics, and employing more advanced models for topic analysis. Through this comprehensive exploration, the study serves as a valuable resource for stakeholders involved in AI development and policymaking, highlighting the path for more ethical AI practices.

Contents

The Public Perception of AI Ethics in the Context of Twitter #ChatGPT Conversations 1
Declaration of Originality 2
Abstract 3
1. Introduction 5
2. Background 6
2.1 Historical Overview of Ethics 6
2.2 Ethics in the Age of AI 7
2.3 AI Ethics, Research and Public Perception 8
2.4 Analytics in AI Ethics Research and Research Gap 9
3 Methodology 10
3.1 Business Understanding 10
3.2 Data Understanding 11
3.3 Data Preparation 12
3.3.1 Data Cleaning 12
3.3.2 Data Pre-processing 13
3.4 Modelling 13
3.5 Evaluation and Deployment 15
4 Analysis 15
4.1 Sentiment Analysis 15
4.1.1 Sentiment Analysis of the Entire Dataset 15
4.1.2 Sentiment Analysis of Ethics-Related Tweets 17
4.2 Topic Modelling and Sentiment Analysis 17
4.2.1 Most Common Words in the Dataset Drawn from the Topic Modelling 18
4.2.2 Sentiment Analysis of the New Data frame Drawn from the Topic Modelling 20
4.2.3 Sentiment Analysis of Topics 22
4.3 Descriptive Analysis of Demographic and Sentiment Data 25
4.3.1 Gender Sentiment Analysis 25
4.3.2 Country Sentiment Analysis 27
4.3.3 Country Geographical Distribution of AI Ethics 28
4.3.4 Relationship Between User Influence and Sentiment 29
4.3.5 Conclusion .... **Error! Bookmark not defined.**
5 Conclusion and Discussion 30
5.1 Practical Implications 31
5.2 Theoretical Contribution 32
5.3 Challenges, Limitations and Future Research 33
5.4 Future Research 34
6 References 35

1. Introduction

The recent advancements in AI have ushered in a new era of technological breakthroughs, sparking interest among a wide range of stakeholders, including policymakers, technologists, and the general public. As AI continues to find its footing in various aspects of our daily lives with the capacity to disrupt and completely change many human endeavours, the need for responsible development and employment of these technologies becomes increasingly pressing. This is why these stakeholders should care deeply about the ethical concerns surrounding AI, which are taking centre stage in discussions about its future.

Existing literature has already highlighted some of these ethical dilemmas. For instance, concerns about biased decision-making processes and outcomes have been raised (Martínez and Fernández, 2019; Katznelson and Chan, 2021), with algorithms being susceptible to training on biased data or incorporating implicit biases, which may end in discriminatory outcomes, especially for marginalized groups. There are also potential threats to job security due to displacement by automation (Wang et al., 2021), such as AI's ability to automate numerous jobs. Lastly, the opaqueness about accountability for the actions of AI systems (Siau and Wang, 2020) is a major concern, with a clear difficulty in understanding AI decision-making processes (Bubeck et al., 2023). This makes it problematic to know the extent to which businesses and individuals are responsible for ensuring the accountable, precise, and ethical usage of such systems.

Despite these known issues, there is a research gap in understanding public perception of AI ethics, particularly in the context of novel and popular AI systems like ChatGPT. Addressing these issues implies formulating appropriate regulations so that all can benefit from ethically developed AI without fear of its misuse. This is considering that the growth and application of AI relentlessly continue, to the point that prominent people have asked publicly to stop its development (Metz & Schmidt, 2023), therefore it is crucial to establish sound ethical practices.

This study contributes to this field by gaining insight into the public perception of AI ethics, particularly in the context of Twitter conversations around ChatGPT due to its novelty and increasing popularity, and Twitter being and important platform for open discourse and exchange of opinions. Analysing public sentiment about the ethics governing this AI technology allows us to proactively address people's concerns more effectively, while at the same time promoting development that is reliable, responsible, and secure. This can be achieved by examining the diverse perspectives, sentiments, and topics expressed through the hashtag #ChatGPT, where prevailing ethical themes and potential areas for improvement can be identified. This public perception empowers us to shape AI technologies in a manner that aligns with societal values and expectations, ensuring a harmonious integration of AI systems into our lives.

In section 2 we will discuss about the background, we transition to detailing our methodology in section 3. Section 4 presents a thorough analysis of our findings, followed by a comprehensive discussion, conclusions, and suggestions for future research in Section 5.

2. Background

2.1 Historical Overview of Ethics

Ethics, at its core, is the philosophical and scientific inquiry into the nature of morality, principles, and values that govern human conduct (Maxim, 2014). From the earliest days of human civilization, ethical concerns have shaped social norms, laws, and cultural practices, evolving alongside our understanding of the world and our place within it. The origins of ethical thought can be traced back to the ancient world, where Eastern and Western philosophical traditions began to grapple with questions of morality and human purpose. In the West, Socratic philosophy and the teachings of Plato and Aristotle laid the groundwork for a rational approach to ethics. In the East, Confucius and Lao Tzu contributed their perspectives, emphasizing harmony, balance, and compassion as foundational principles.

As the centuries progressed, ethical inquiry continued to evolve, with the Enlightenment bringing a renewed emphasis on reason, individualism, and secularism. Key thinkers such as Immanuel Kant, John Stuart Mill, and Jeremy Bentham developed influential ethical theories, including deontology, consequentialism, and utilitarianism, which aimed to provide a systematic approach to determining right and wrong. The 20th century witnessed further diversification in ethical thought, with existentialism, phenomenology, and virtue ethics gaining prominence. These strands of ethical theory emphasized individual responsibility, the importance of subjective experience, and the cultivation of personal virtues as essential components of moral life.

2.2 Ethics in the Age of AI

In this new era, the main ethical questions take on new dimensions, as we are challenged to apply these principles to a rapidly changing technological landscape. Interestingly, a huge diversification of ethical thought came to light in an age when machines began replacing manual labour, resulting in significant social and economic changes and a fear that machines or automation may replace human labour and potentially lead to the domination of society by a small group of individuals who control the technology. These concerns lead to a growing debate about the impacts of automation on people's sense of self-worth, as work is often a major source of self-esteem in modern societies (Abrams & Hogg, 1988). On the other hand, some researchers argue that automation has historically created more jobs than it has displaced and that automation will complement rather than replace the workforce, making workers more productive and creating new, more rewarding occupations (Evans, 2017; Wright & Schultz, 2018). These debates have been evolving until today and they are more prevailing with the advent of advanced AI systems where new strands of machine ethics research focused on AI ushered due to the potentially great impact that the adoption of the technology can cause in lives (Stahl, 2021; Awad et al., 2018).

While there are ongoing ethically driven efforts to improve AI systems in areas such as accountability, privacy protection, and explainability (Li *et al.*, 2021), there are notable interest in addressing other crucial ethical aspects. These interests range from the potential danger of malevolent artificial general intelligence to a lack of diversity in the AI community and beyond. The previous and new preoccupations around the growth of AI calls for a transition from a deontological, rule-based ethical approach to a more nuanced, situation-sensitive ethical approach, focusing on virtues, personality dispositions, and responsible autonomy (Hagendorff, 2019).

The future of AI ethics needs to empower moral actors to act responsibly and empathetically in morally significant situations. This approach calls for a balance between a more detailed focus on technological aspects of AI and a shift towards genuine social and personality-related aspects. By doing so, AI ethics would offer a broader perspective, moving beyond purely technological phenomena and embracing a more holistic view of ethical considerations in the age of AI (Hagendorff, 2019).

2.3 AI Ethics, Research and Public Perception

The public's increased interest in AI has led to greater scrutiny of the ethical implications and unintended consequences of AI systems. The case of Microsoft's Twitter-based chatbot "Tay" is a prominent example of these concerns. Tay was taken down in less than 24 hours of its release due to the public teaching it to use misogynist, racist, and politically incorrect phrases, highlighting the urgent need for responsible AI principles (Neff & Nagy, 2016; Mikalef et al., 2022).

AI ethics research also have emphasized the significance of understanding the ethical consequences of AI systems as they gain increased autonomy and integration into society. Ethical concerns surrounding AI usage mostly revolve around the issues that are derived from the participation of the AI, the human counterpart, and society (Siau and Wang, 2020), as the "Tay" case displays, where the concurrence of the three stakeholders creates externalities that will ultimately end up being paid by the public. These issues can have such a big impact on the future of society as a whole that the manner in which society tackles them will shape our trust in AI and its impact on individuals (Dignum, 2018).

These incidents have prompted a more holistic understanding of what constitutes responsible AI, with a focus on eliminating bias, explainability of AI outcomes, and ensuring safety and security. Multiple initiatives across different entities and organizations are underway to establish ethical, transparent, and accountable use of AI technologies consistent with user expectations, organizational values, and societal laws and norms (Barredo Arrieta et al., 2020; Mikalef et al., 2022).

This growing consensus on responsible AI underscores the necessity to understand public perception of AI ethics. It provides a basis for examining how society perceives the principles that ensure AI technologies are used responsibly, and how these perceptions shape the discourse on AI ethics on platforms like Twitter (Mikalef et al., 2022).

As a result of the clear impact of ethical issues of AI on society, it is crucial to understand the real needs and address the research gap in public perception of ethical challenges for the development of a human-centric and beneficial AI that serves society as a whole. This includes discerning the key themes contemplated by various groups, such as different genders and countries, and the implications for each, as diverse groups assign varying weights to ethical principles (Kieslich, Keller, and Starke 2021).

2.4 Analytics in AI Ethics Research and Research Gap

Analytics has been used in addressing interesting similar topics, a study by Manikonda and Kambhampati (2018) employed sentiment analysis to understand public sentiment toward AI on Twitter. They found that the public tends to be more positive about AI, and women are generally more optimistic about AI impact than men. They also found that women often focus more on the ethical issues surrounding AI.

Another study by Taecharungroj (2023) used topic analysis to examine early reactions to ChatGPT on Twitter. They found that ethical discussions were a significant part of the conversation around this technology. This kind of analysis offers a more nuanced view of the public conversation on AI, as it allows for the identification of specific themes and sentiments expressed by the public.

Despite these advancements in the field using analytics, a gap in the literature exists, as no study has yet examined the public perception of ethics in a social media platform like Twitter with a specific focus on a language model such as ChatGPT. This study aims to bridge that gap and contribute to the existing body of knowledge by employing sentiment and topic analysis to explore public perception of AI ethics in the context of the hashtag #ChatGPT on Twitter. This approach not only expands the scope of AI ethics research but also provides valuable insights for stakeholders involved in AI development and policymaking with the potential to inform the creation of responsible AI policies and practices that align with societal values and expectations.

3 Methodology

The methodology section of this study outlines the systematic approach employed to achieve the study objectives. This approach is grounded in the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, a widely accepted standard for conducting data mining projects. The CRISP-DM methodology comprises six phases: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment (Shearer, 2000). Each phase is designed to ensure a comprehensive and rigorous approach to the study, from understanding the objectives and the available data to preparing the data, developing the model, evaluating the results, and deploying the insights. The following sections detail the specific methods and techniques used in each phase of the CRISP-DM methodology for this study.

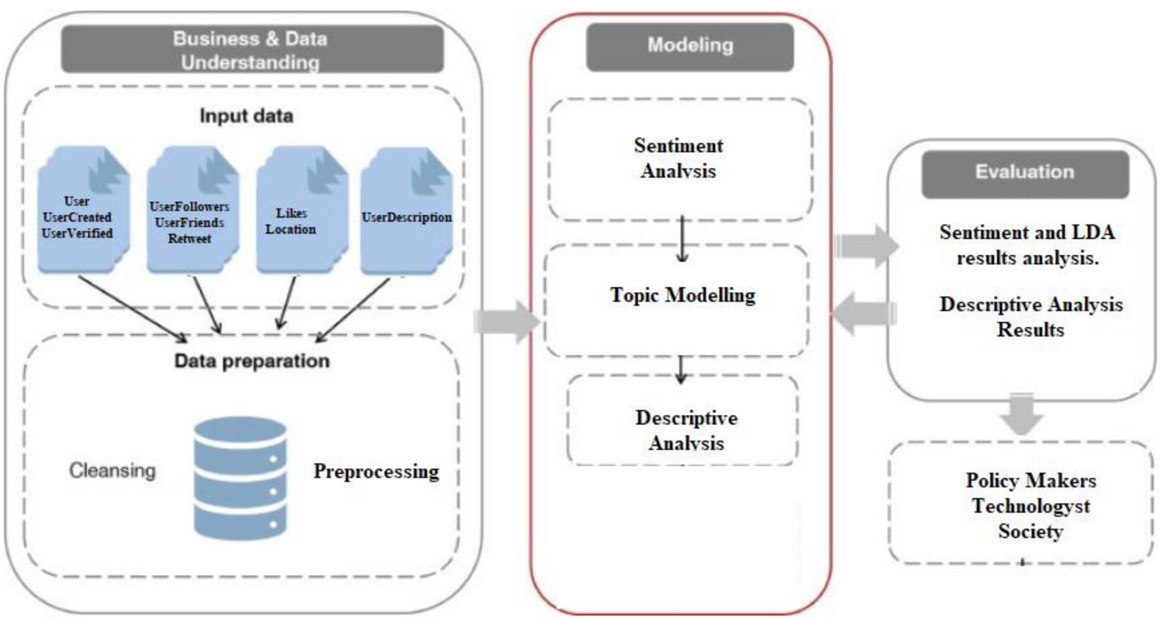

Fig 1: Data analysis operations. Based on (Griva et al., 2018).

3.1 Business Understanding

This study tackles the business problem of understanding public perception related to AI ethics encapsulated in Twitter conversations marked with the #ChatGPT hashtag. The aim here is not only to understand the sentiment and highlight the prevailing themes and concerns but also to closely examine the association between sentiment and the identified ethical topics. Moreover, it is critical to understand the distribution of these opinions across different demographics and geographic regions. The study intends to equip stakeholders, including technologists and policymakers, with robust insights. These insights could significantly contribute to formulating effective strategies to tackle ethical dilemmas emerging in AI.

The objectives of this study are to:

1. Investigate public sentiment towards AI ethics, as expressed in #ChatGPT tweets.
2. Identify key themes and concerns related to AI ethics discussed within the #ChatGPT dataset.
3. Analyse the relationship between sentiment and the identified ethical topics.
4. Understand the demographic and geographic distribution of opinions regarding AI ethics.
5. Provide insights to inform policymakers, technologists, and other stakeholders on addressing AI ethical concerns.

3.2 Data Understanding

The dataset utilized in this study comprises a collection of Twitter data associated with the hashtag #ChatGPT acquired from Kaggle*. The dataset includes the following features:

Date	The date when the tweet was published, providing a chronological context for the analysis.
Tweet	The content of the tweet, which serves as the primary source of textual data for sentiment and topic analysis.
Url	The link to the original tweet, offering a direct reference to the tweet in its native context.
User	The username of the individual who authored the tweet, allowing for the identification and potential analysis of user-specific patterns or trends.
UserCreated	The date when the user's account was created, which may provide insights into user experience or familiarity with the platform.
UserVerified	A Boolean field indicating whether the user is verified or not, which may be used to distinguish between influential users or experts and regular users.
UserFollowers	The follower count of the user, providing a measure of the user's influence or reach within the Twitter community.
UserFriends	The friends count of the user, which may be indicative of the user's level of engagement or connectedness on the platform.
Retweet	The count of retweets for a particular tweet, reflecting the tweet's reach and potential impact on the wider Twitter audience.
Likes	The count of likes for a tweet, serving as an indicator of the tweet's popularity or resonance among users.
Location	The geographical location of the user, enabling the exploration of regional variations in sentiment or ethical concerns.
UserDescription	The user's self-description, offering additional context and information about the user's background, interests, or perspective

Date	Tweet	Url	User	UserCreat	UserVerifi	UserFollov	UserFrien	Retweets	Likes	Location	UserDescription
2023-02-2	How to	https://tw	smnishad	2009-03-0	FALSE	2524	4966	0	0	New Delhi	Account Planning at /
2023-02-2	Chatgtp bı	https://tw	SevenKing	2010-05-0	FALSE	1322	174	0	0	Ilford, Red	All Through. Outstan
2023-02-2	@PiCore	https://tw	jad_alrabe	2012-03-0	FALSE	311	1822	0	0	Irbid, jord	Ø£Ù†Ø§
2023-02-2	Build	https://tw	yournotio	2023-02-0	FALSE	43	10	0	0		ðŸ¤“

Fig 2:
Sample of the dataset

*Dataset: https://www.kaggle.com/datasets/manishabhatt22/tweets-onchatgpt-chatgpt

3.3 Data Preparation

Data cleaning and pre-processing play a pivotal role in sentiment and topic analysis, as they contribute to the refinement of data quality, mitigation of noise, and establishment of consistency across the dataset (Palomino & Aider, 2022). These steps ensure the streamlining of the analytical procedure and heighten accuracy. Furthermore, they reduce computational complexity and prepare the data for natural language processing, which is crucial for sentiment and topic analysis.

3.3.1 Data Cleaning

Data cleaning ensures the quality and reliability of the dataset. The process for this study involved the following steps:

1. **Handling missing values**: The dataset was checked for null values in the “Tweet” column, which contained the primary textual data for our analysis. Initially, there were six null values in the column. To address this issue, we filled the null values with empty strings, ensuring that the column contained no null values in the subsequent analysis.
2. **Filtering irrelevant tweets**: The analysis in this study is carried out with deductive reasoning that entails drawing conclusions by moving from a broad premise to a detailed one. Therefore the first sentiment and topic analysis was drawn from the entire dataset and subsequently, to move to a more detailed premise, the “Tweets” column was filtered using relevant keywords and criteria. These criteria included keywords such as: "ethics", "bias", "fairness", "privacy", "transparency", "accountability", "safety", "security", "trust", "responsibility", "regulation", "discrimination", "responsibility" and "job". The later process reduced 96.5% of the dataset. In the third analysis based on a new dataset focused only on selected topics drawn from the topic analysis, the dataset was reduced by 97.7%

3.3.2 Data Pre-processing

Data pre-processing involves preparing the text data for analysis. The following text pre-processing steps were implemented for this study:

1. **Text cleaning**: A custom function was developed to clean the text data by removing URLs, special characters, and converting the text to lowercase. This step ensured that the textual data was uniform and free of unnecessary noise.
2. **Tokenization**: The cleaned tweets were tokenized, which involved breaking down the text into individual words or tokens. Tokenization facilitates the further processing and analysis of text data.
3. **Stopword removal and short token removal**: To reduce the computational complexity and improve the efficiency of the analysis, common stopwords, and short tokens were removed from the tokenized tweets.
4. **Stemming/Lemmatization**: Involves reducing words to their base or root form. This process helps in consolidating similar words and improving the accuracy of the analysis.
5. **Demographics extraction**: The "Location" and "Gender" data were cleaned and pre-processed to extract demographic information about the users, which was essential for understanding the geographic and demographic distribution of opinions regarding AI ethics.

3.4 Modelling

The methodology employed in this study aims to address the objectives outlined in the Business Understanding phase by utilizing sentiment analysis, topic analysis, and descriptive analysis using Python and the following libraries:

1. Pandas: A data manipulation library for data cleaning, transformation, and analysis.
2. NumPy: A library for scientific computing in Python.
3. TextBlob: An NLP library for common text processing tasks.
4. Matplotlib: A data visualization library.
5. Seaborn: A data visualization library based on Matplotlib.

6. Gensim: A library for unsupervised topic modelling and NLP.
7. NLTK: A library for NLP and computational linguistics.
8. Geopy: A geocoding library for Python.
9. GeoText: A library for extracting geographical entities from unstructured text.

. The following methods detail the approach to the objectives:

1. Sentiment Analysis

Sentiment analysis is a field of study that deals with the computational treatment of opinions, sentiments, and subjectivity in text. It involves using technology to analyse and understand the emotions and attitudes expressed in written or spoken language (Pang & Lee, 2008). This methodology allows us to determine the general sentiment of a tweet, whether it is positive, negative, or neutral. This technique can gauge the public's perception of AI ethics, identify areas of concern, and evaluate potential trends over time when applying it to the dataset of Twitter #ChatGPT conversations. This approach provides a valuable foundation for understanding the public's attitude towards AI ethics and informing further research and policy discussions.

2. Topic Analysis

The Latent Dirichlet Allocation (LDA) is a technique used to identify latent topics in a corpus of documents. It is a three-level hierarchical Bayesian model that infers the latent topic distribution of each document and the word distribution of each topic (Zhang, Luo and Tang, 2016). LDA is particularly useful in extracting meaningful insights from large, unstructured text data. In the context of this study, topic analysis enables us to identify the key themes and topics related to AI ethics that emerge from the Twitter #ChatGPT conversations to better understand the public's concerns and priorities regarding AI ethics, as well as identify areas that warrant further investigation or require the attention of policymakers and technologists.

3. Descriptive Analysis

Descriptive Analysis refers to the process of summarizing and examining data to describe the main features and patterns within it. This methodology often involves the use of summary statistics, such as measures of central tendency and dispersion, to provide a high-level overview of the data. In the context of this study, descriptive analysis allows us to contextualize the findings from demographic and geographic information to identify trends, patterns, and potential areas for intervention.

3.5 Evaluation and Deployment

The evaluation of the methodology will be based on the insights generated from the sentiment analysis, topic analysis, and descriptive analysis. The effectiveness of the methodology will be determined by its ability to meet the study objectives.

The insights generated from this study can be shared with policymakers, technologists, and society to inform the development of responsible AI policies and practices that align with societal values and expectations.

4 Analysis

4.1 Sentiment Analysis

This sentiment analysis is conducted in two parts: The first part examines the entire dataset to establish a baseline understanding of the overall sentiment related to ChatGPT conversations on Twitter and identify overarching trends and patterns. After gaining this broader context, we then narrow our focus to the dataset containing ethics-related keywords, which allows us to explore public sentiment specifically concerning AI ethics. This comprehensive and targeted approach not only enables comparisons between the general and ethics-related sentiment but also uncovers nuances in the public's perception of AI ethics in the context of Twitter ChatGPT conversations.

4.1.1 Sentiment Analysis of the Entire Dataset

To analyse the sentiment of the whole dataset, a total of 305,432 tweets were examined. The mean sentiment score of these tweets is 0.123198, with a standard deviation of 0.261039, reflecting a generally positive sentiment towards ChatGPT conversations on Twitter. This average is, however, closer to a neutral sentiment, which indicates that public opinion is not overwhelmingly positive. The minimum and maximum sentiment scores are -1 and 1, respectively, representing the full range of sentiment values, from highly negative to highly positive.

The quartile values provide additional insights into the distribution of sentiment scores. The 25th percentile (Q1) is 0, the median (50th percentile or Q2) is 0.033333, and the 75th percentile (Q3) is 0.263941. These results indicate that the majority of tweets exhibit a neutral to mildly positive sentiment towards ChatGPT conversations. The skewness towards positive sentiment in the upper quartile may suggest a more favourable view of AI ethics in the context of these discussions.

the subsequent figure provides an insightful perspective on the overall emotional tone embodied within the #ChatGPT tweets. The graph is a visualization that represents the distribution of sentiment scores throughout the entire dataset.

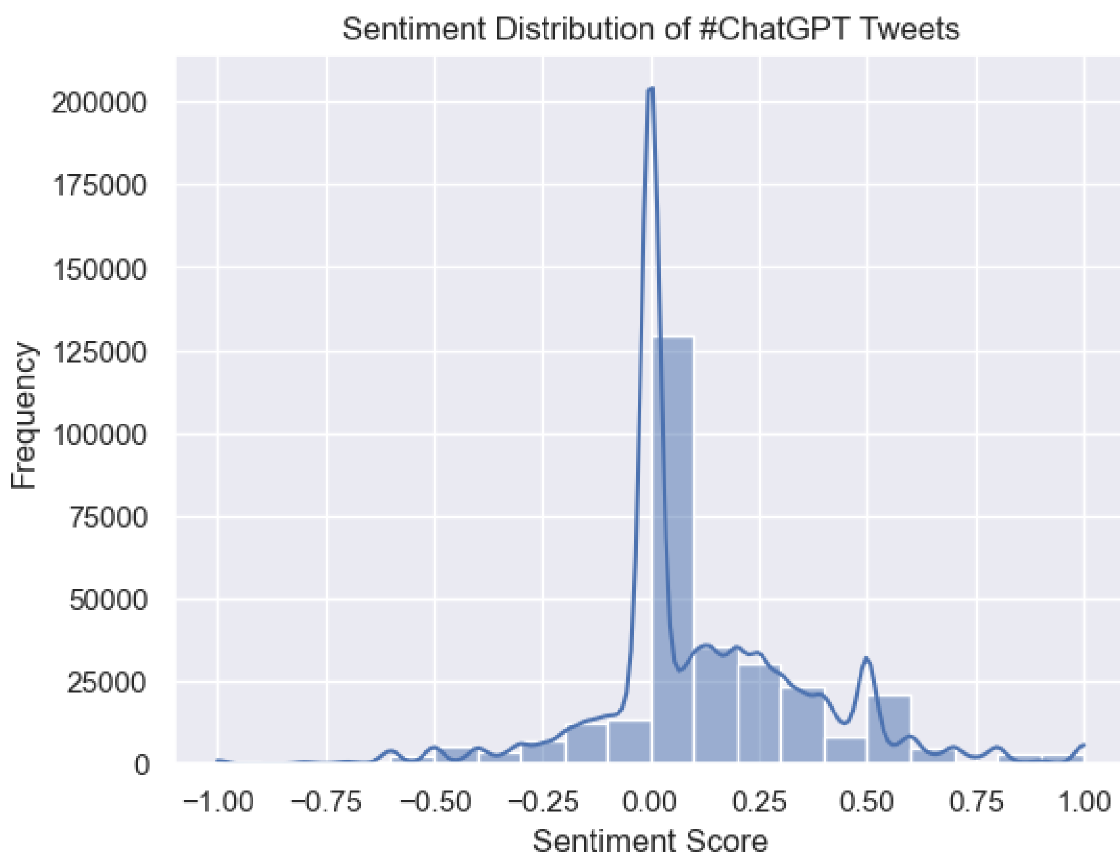

Fig 3: Sentiment distribution of #ChatGPT tweets in the entire dataset. Positive sentiment is above 0.00, negative sentiment below 0.00, and neutral at 0.

4.1.2 Sentiment Analysis of Ethics-Related Tweets

The second part of our sentiment analysis focuses on tweets containing ethics-related keywords, such as "bias", "transparency", "accountability", "privacy", "discrimination", "responsibility", and "job". In this subset, 10,453 tweets were analysed.

The mean sentiment score for these tweets is 0.131244, which is slightly higher than that of the entire dataset, with a standard deviation of 0.251641. This finding may suggest that users discussing AI ethics-related topics tend to have a somewhat more positive sentiment toward ChatGPT conversations. The minimum and maximum sentiment scores are -1 and 1, respectively, similar to using the entire dataset.

The quartile values for this subset are: the 25th percentile (Q1) is 0, the median (50th percentile or Q2) is 0.080952, and the 75th percentile (Q3) is 0.275000. Compared to the entire dataset, these results show a marginally more positive sentiment distribution among tweets with ethics-related keywords. This shift may indicate that users engaged in discussions about AI ethics appreciate the importance of addressing these concerns in the context of ChatGPT conversations.

The sentiment analysis reveals that the overall perception of AI ethics in ChatGPT conversations is generally positive, with a skew toward neutral sentiment. The sentiment becomes slightly more positive when focusing on tweets containing ethics-related keywords, suggesting that users discussing these topics recognize the significance of ethical considerations in AI-driven conversations. To maintain this positive perception, it is crucial to continue fostering an open dialogue on AI ethics and engaging various stakeholders in discussions.

4.2 Topic Modelling and Sentiment Analysis

A topic analysis was performed using the Latent Dirichlet Allocation (LDA) method with the objective of identifying key themes and concerns related to AI ethics discussed within the dataset. A set of keyword filters, namely: "ethics", "bias", "fairness", "privacy", "transparency", "accountability", "safety", "security", "trust", "responsibility", and "regulation" was used to determine five topics with the LDA method and a new dataset was created focusing on the best topics. The purpose of the new dataset was to analyse the most common words and create a new sentiment analysis around the main topics. The five generated topics painted a good picture and permitted a focus on the best ones for further inspection:

Topic 0: ChatGPT, machine learning, Python, cybersecurity, 100 days of code

Topic 1: ChatGPT, cybersecurity, security, malware, new

Topic 2: ChatGPT, security, cybersecurity, OpenAI, Google

Topic 3: ChatGPT, ethics, technology, artificial intelligence, OpenAI

Topic 4: ChatGPT, bias, trust, biased, data

Topics 3 and 4 encompass the most accurate ethics-related themes to create a new focused dataset. Topic 3 appears to focus on the intersection of ChatGPT, AI ethics, technology, and artificial intelligence, which are central to this study. Topic 4, on the other hand, revolves around the ethical concerns of bias and trust in the context of ChatGPT conversations.

4.2.1 Most Common Words in the Dataset Drawn from the Topic Modelling

The most common words in the new data frame, which focuses on topics 3 and 4, are ChatGPT, bias, ethics, cybersecurity, trust, data, OpenAI, biased, amp, and security. The prevalence of these words underscores the significance of AI ethics within the ChatGPT conversations, with particular emphasis on the challenges of bias and trust.

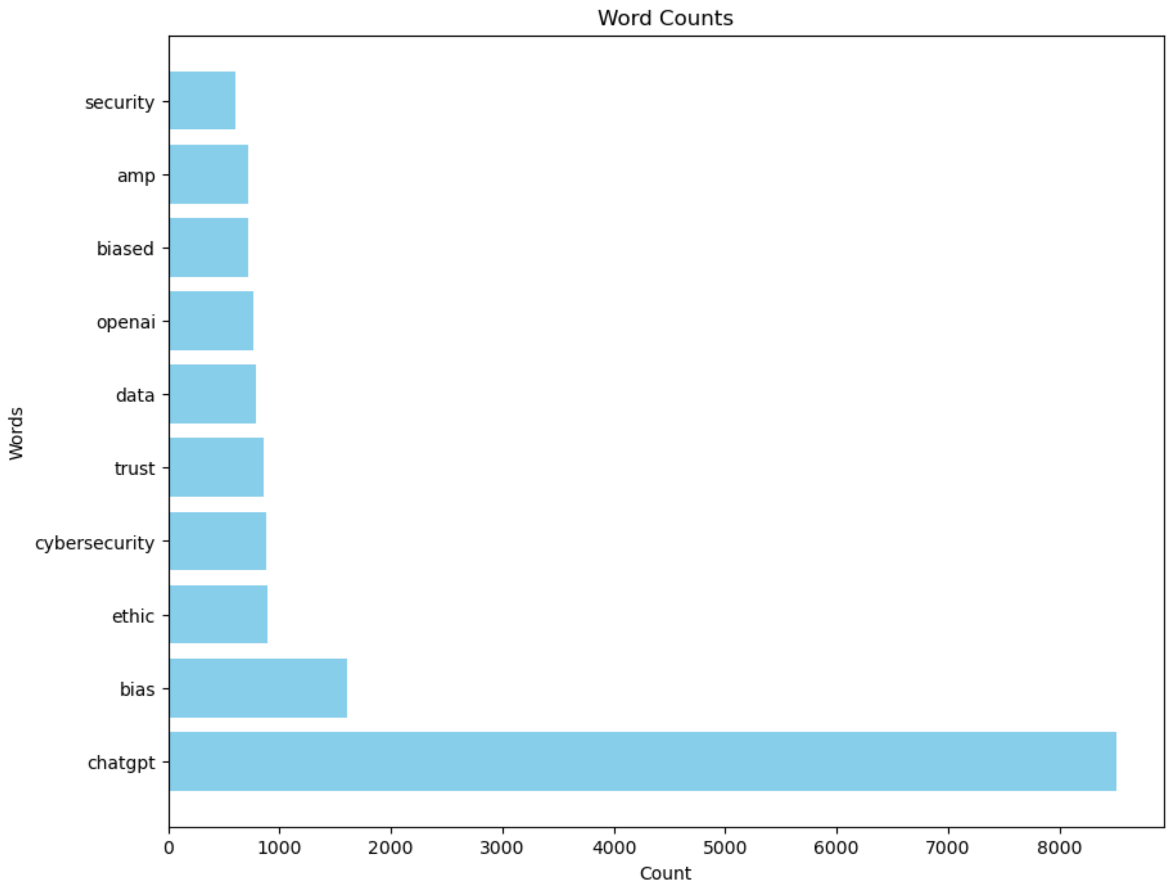

Fig 4: Most common words in the new dataset drawn from topic modelling

The prominence of "bias" and "biased" in the dataset reflects the growing concern among users about the potential for AI systems, such as ChatGPT, to perpetuate or even exacerbate existing biases present in the data they are trained on. This issue has critical implications for the fairness and equity of AI systems, as biased outcomes may lead to unintended negative consequences for certain individuals or groups.

The frequent occurrence of "ethic" and "trust" emphasizes the need for transparency, accountability, and responsibility in AI systems. Trust is crucial for users to feel confident in adopting and relying on AI-driven technologies like ChatGPT. To foster trust, developers, and organizations must address ethical concerns and demonstrate that their AI systems operate in a manner that aligns with the users' values and expectations.

The presence of "cybersecurity," "data," and "security" in the list of common words highlights the importance of data protection and privacy in AI ethics discussions. As AI systems process vast amounts of data, ensuring the security and confidentiality of user data becomes paramount. Users must be confident that their personal information is protected and used responsibly, further emphasizing the need for robust security measures and transparent data usage policies.

Lastly, the inclusion of "openai" in the list of most common words suggests that users recognize the organization as a key player in the development and deployment of AI systems like ChatGPT. This highlights the responsibility of companies such as OpenAI, the developer of ChatGPT's large language model, to address the aforementioned ethical concerns and work towards creating AI systems that are not only powerful and efficient but also fair, transparent, and secure.

4.2.2 Sentiment Analysis of the New Data frame Drawn from the Topic Modelling

The sentiment analysis output for the new data frame focused on topics 3 and 4 is as follows: A mean of 0.213744 and a Standard deviation of 0.489834, The minimum and maximum sentiment scores are -0.963800 and 0.978600, respectively.

the subsequent figure exemplifies these outcomes of sentiment analysis performed on the topic modelling. This diagram delivers a visual representation of the dispersion of sentiment scores associated with the topics identified through the LDA model.

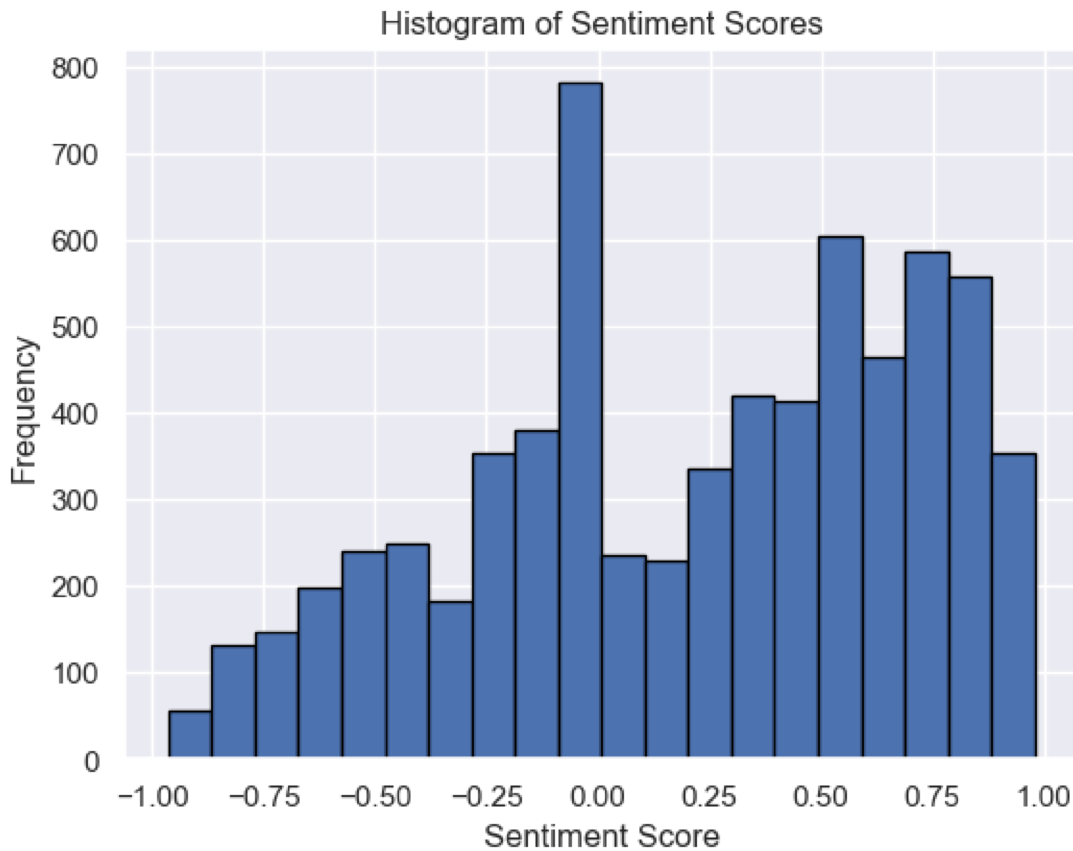

Fig 7: Sentiment distribution of #ChatGPT tweets in the dataset drawn from the topic modelling. Positive sentiment is above 0.00, negative sentiment below 0.00, and neutral at 0.

The mean sentiment score of 0.213744 in the data frame signifies an even more positive sentiment compared to the entire dataset and the dataset containing ethics-related keywords of point 4.1. This finding implies that users discussing the topics of ethics, bias, and trust in the context of ChatGPT generally have a more favourable outlook, which may be attributed to several factors.

First, the increased positivity in sentiment could be a reflection of users' appreciation for the growing awareness and open discussion of AI ethics issues. As more people engage in conversations about the ethical implications of AI systems, it may foster a sense of shared responsibility and collective effort to address these concerns. This collective effort could contribute to a more positive sentiment among users.

Second, the higher mean sentiment score may also suggest that users perceive progress in tackling AI ethics challenges. As organizations like OpenAI work to develop more transparent, accountable, and fair AI systems, users may view these advancements positively and recognize the potential for AI to benefit society while minimizing its negative consequences.

Moreover, the distribution of sentiment scores in the new focused data frame demonstrates a wider range, as indicated by the higher standard deviation and the larger gap between the minimum and maximum scores. This wider range can be interpreted as evidence of the diverse opinions and perspectives held by users discussing AI ethics topics. It also highlights the complexity and multifaceted nature of ethical concerns surrounding AI, with some users holding more optimistic views while others express skepticism or even criticism.

The broader distribution of sentiment scores may also signal a healthy and robust dialogue surrounding AI ethics, as it encompasses a variety of opinions and experiences. This diversity in perspectives is essential for addressing the ethical challenges posed by AI systems, as it allows for the identification and consideration of a wide array of potential issues and solutions.

4.2.3 Sentiment Analysis of Topics

Upon examining the average sentiment scores for the identified ethical topics 3 and 4 in the new dataset drawn for the topic modelling, the results reveal distinct differences between the two topics. This analysis separates the new dataset drawn from topic modelling, to understand the sentiment score of the themes from topic 3 (ChatGPT, ethics, technology, artificial intelligence, OpenAI) and the themes from topic 4 (ChatGPT, bias, trust, biased, data). The average sentiment score for topic 3 is 0.227727, while for topic 4, it is 0.201880. Both topics exhibit positive sentiments on average, suggesting that users discussing these ethical issues generally express a favourable view. However, the difference in the mean sentiment scores between the two topics indicates that users discussing topic 3 have a slightly more positive outlook compared to those discussing topic 4.

The following figure offers a view of the average sentiment score across different ethical topics. This diagram shows the emotional undertones associated with each topic, thereby facilitating an understanding of public sentiment regarding the facets of topics 3 and 4.

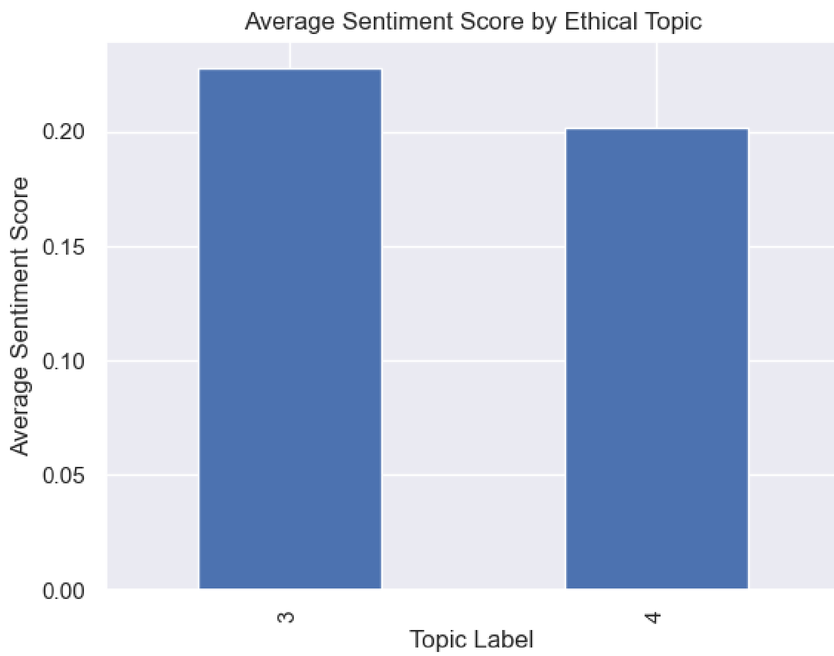

Fig 8: Average sentiment score of Topic 3 and Topic 4

Topic 3, which is characterized by keywords such as “ethics”, “technology”, “artificial” “intelligence”, and “OpenAI”, seems to evoke a more optimistic sentiment among users. This could be due to a perception that advances in AI technology and increased awareness of ethical concerns are leading to progress in addressing these issues. Users discussing this topic may appreciate the efforts made by organizations like OpenAI to develop more transparent, accountable, and fair AI systems.

On the other hand, topic 4, which is centred around keywords like bias, trust, and data, demonstrates a somewhat lower average sentiment score. While still positive overall, this lower sentiment score may reflect the challenges and complexities associated with addressing bias and trust in AI systems. Users discussing this topic might express concerns about the potential negative consequences of biased algorithms and the difficulties in ensuring fairness and trust in AI applications.

The correlation of -0.26 between topics distribution and sentiment scores suggests a weak negative relationship between the two topics. This correlation implies that as the proportion of discussions shifts more toward one topic, the sentiment score tends to decrease slightly. The negative correlation might indicate that users discussing the topics in a more balanced manner express more varied opinions, leading to a wider range of sentiment scores.

The subsequent figure presents a comparative view of the correlation of sentiment scores on topics 3 and 4.

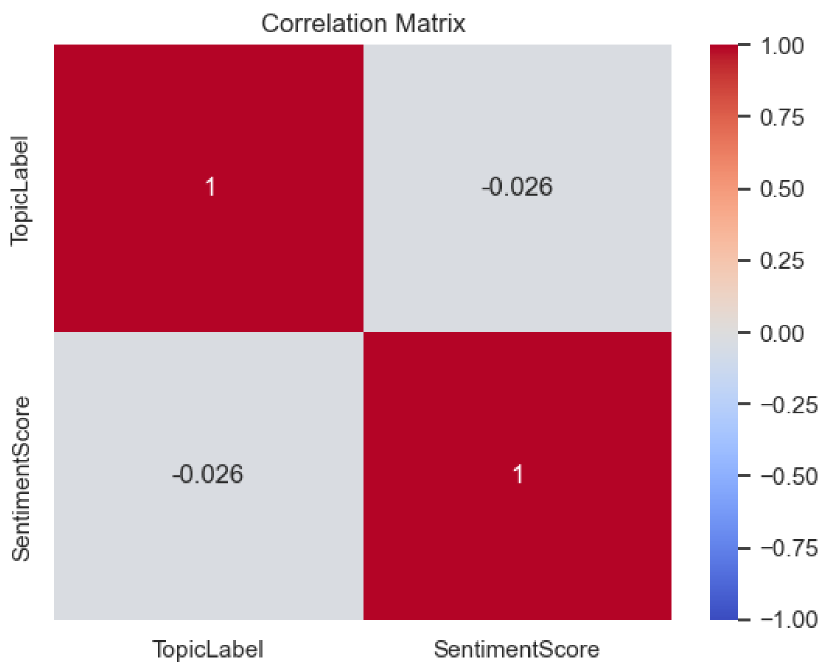

Fig 9: Topic distribution and sentiment score correlation on topics 3 and 4

The topic analysis highlights key themes and concerns related to AI ethics within the ChatGPT dataset, focusing on ethics, bias, and trust. Addressing these concerns is essential to ensure the responsible development and deployment of AI systems that positively impact society. The sentiment analysis of the new dataset demonstrates that users discussing these themes tend to have a more positive sentiment toward ChatGPT conversations. Addressing these ethical concerns effectively may contribute to a more favourable public perception of AI-driven interactions on Twitter and other platforms.

The higher mean sentiment score and wider distribution of sentiment scores in the new dataset indicate a more favourable view of users discussing ethics, bias, and trust in the context of ChatGPT, as well as a rich and diverse dialogue on AI ethics. The weak negative correlation between topic distribution and sentiment scores highlights the diverse range of opinions and the complex nature of ethical concerns surrounding AI systems. These findings emphasize the importance of continued conversations and collaboration among various stakeholders to ensure the responsible development and deployment of AI systems.

4.3 Descriptive Analysis of Demographic and Sentiment Data

A comprehensive examination of sentiment scores, geographic distributions and user status in the dataset drawn from the topic modelling, categorized by gender and country, is aimed to uncover disparities and patterns in the perception of AI ethics, shedding light on the varying perspectives that exist within diverse demographic and geographic contexts.

4.3.1 Gender Sentiment Analysis

Going further into the analysis, it is pertinent to understand the connection between gender and sentiment scores. The following figure provides a visual representation of the distribution of sentiment scores across different gender categories in the dataset.

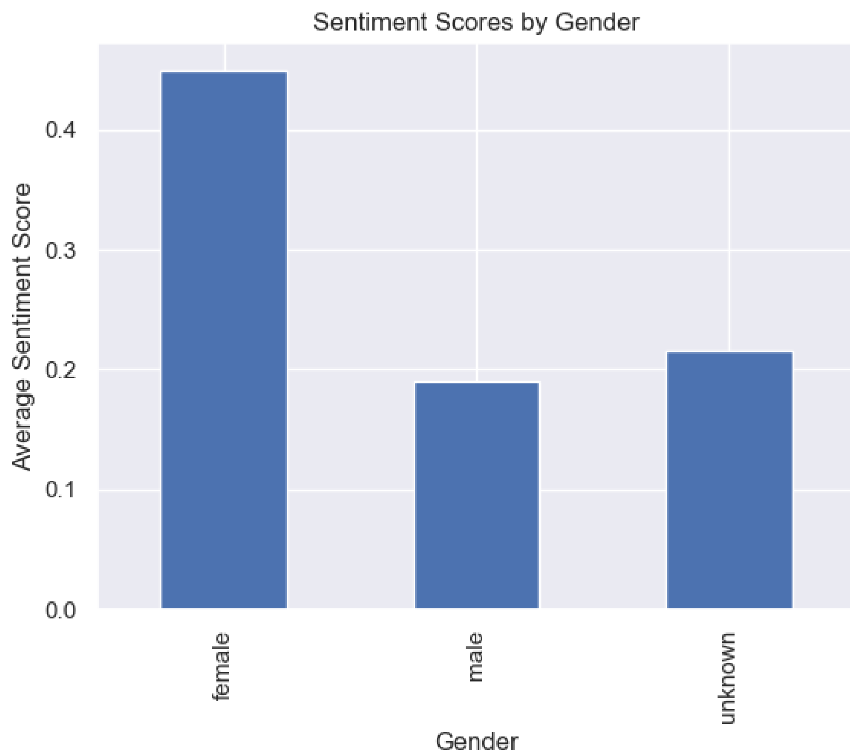

Fig 10: Sentiment scores by gender

The sentiment data indicate that females demonstrate a notably positive sentiment toward AI ethics (sentiment score of 0.449661) in comparison to their male counterparts (sentiment score of 0.190439) and individuals of unspecified gender (sentiment score of 0.215606), this is in line with Manikonda and Kambhampati (2018) insights. This observation implies that female users may harbour a more optimistic perspective on the ethical dimensions of AI, possibly exhibiting greater confidence in the capacity to tackle and resolve ethical dilemmas within AI systems. In contrast, the sentiment scores for males and users of unknown gender tend to be lower, suggesting that these individuals might adopt a more prudent or critical stance on AI ethics. Such a viewpoint could be driven by heightened awareness of the potential risks, challenges, and complexities associated with addressing ethical issues in AI, or by a desire to ensure that ethical considerations are thoroughly and rigorously examined. This gender-based discrepancy in sentiment underscores the importance of including diverse perspectives in AI ethics discussions, as it can lead to a more comprehensive understanding of the various concerns and potential solutions.

4.3.2 Country Sentiment Analysis

The sentiment scores for various countries show a diverse range of opinions about AI ethics. Some countries, such as Argentina (AR, 0.657450), Belgium (BE, 0.644500), and Ecuador (EC, 0.731500), exhibit high sentiment scores, indicating a positive outlook on AI ethics. These countries may have more optimism about the progress and potential of AI technology in addressing ethical concerns.

The following figure provides a visual representation of the distribution of sentiment scores across different countries in the dataset

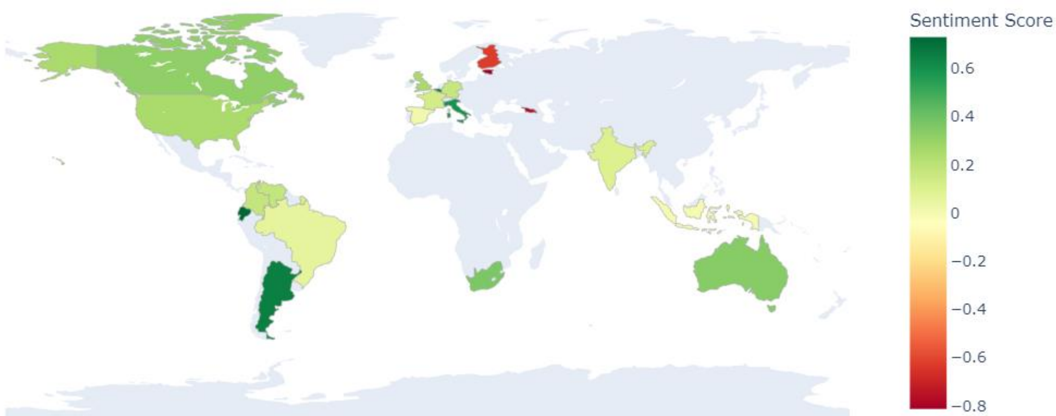

Fig 11: Sentiment scores by country

As can be visualised in Figure 11, there is a strong contrast in countries like Estonia (EE, -0.812600), Finland (FI, -0.624900), and Georgia (GE, -0.778300) where there are negative sentiment scores, reflecting a more pessimistic view on AI ethics. Users from these countries might express concerns about the challenges in addressing ethical issues in AI systems or the potential negative consequences of biased or unfair algorithms.

The sentiment scores for other countries, such as the United States (US, 0.253230), the United Kingdom (GB, 0.253168), and India (IN, 0.098664), are more moderate, indicating a balanced perspective on AI ethics. Users from these countries may recognize the potential benefits of AI technology while acknowledging the ethical challenges that need to be addressed.

4.3.3 Country Geographical Distribution of AI Ethics

The geographical distribution of opinions on AI ethics, as represented by the number of tweets per country, indicates that the majority of the conversation is concentrated in a few key countries. The visual representation in Figure 12 of the distribution of tweets across different countries in the dataset shows that the United States leads the discussion with 229 tweets, followed by the United Kingdom with 56, and India with 44. Canada and France also contribute a significant number of tweets, with 32 and 14 respectively. South Africa and Germany each account for 11 and 8 tweets respectively, while Australia, and Hong Kong each have fewer than 10.

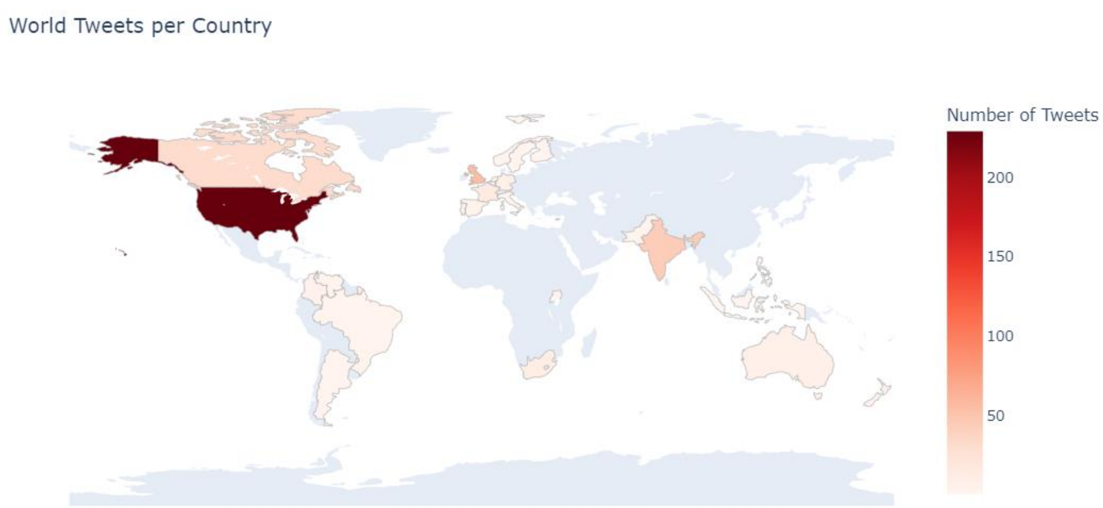

Fig 12: Tweets per country

The conversation about AI ethics appears to be more prevalent in English-speaking countries and those with strong technology sectors, such as the US, UK, Canada, and Australia. However, it is important to note that a considerable number of tweets also originate from non-English speaking countries, such as India, France, Germany, and South Africa. This suggests that AI ethics is a global concern, transcending language and cultural barriers.

In contrast, several countries have a relatively low number of tweets on AI ethics, which may indicate a lack of awareness or limited engagement in the topic. It is essential to foster dialogue and collaboration across nations to ensure that diverse perspectives and experiences are considered in discussions around AI ethics. This will help to promote a more inclusive and comprehensive understanding of the challenges and potential solutions related to ethical AI development and deployment.

4.3.4 Relationship Between User Influence and Sentiment

Examining the relationship between user influence and sentiment can provide valuable insights into how influential users shape the discussion on AI ethics. In this case, we assess user influence through the number of followers and compare it with the sentiment score of their tweets.

The following graph illustrates the correlation between the authority a user holds in the #ChatGPT discourse and the sentiment embodied in their tweets.

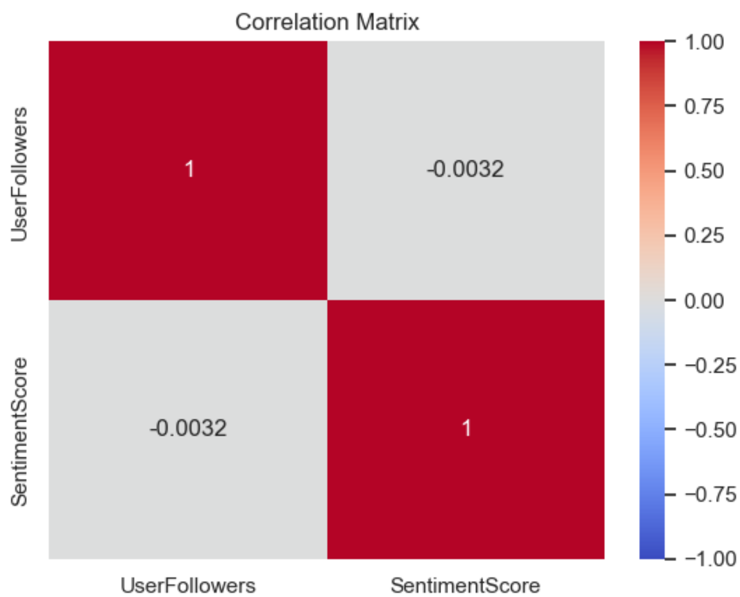

Fig 13: User influence and sentiment score correlation

The correlation matrix shows a negligible correlation of -0.003187 between user influence (UserFollowers) and sentiment score (SentimentScore). This result suggests that there is virtually no relationship between the influence of a user and the sentiment expressed in their AI ethics-related tweets.

In other words, the sentiment expressed by users with a large number of followers is not significantly different from those with fewer followers. It is important to consider that this lack of correlation indicates that the discussion around AI ethics is not significantly skewed or dominated by influential users. Consequently, it implies that the conversation around AI ethics is diverse and incorporates a wide range of perspectives, regardless of the level of influence a user possesses.

The analysis of demographic data reveals a diverse range of opinions regarding AI ethics, with differences in sentiment scores across gender and country. Understanding these variations in perspectives can help inform discussions and policy development surrounding AI ethics, ensuring that a wide range of voices is considered in addressing the ethical challenges posed by AI technology. It is important to promote inclusivity and guarantee that different perspectives are represented in AI ethics guidelines and policies.

5 Conclusion and Discussion

The analysis of AI ethics in the context of Twitter ChatGPT conversations has provided valuable insights into public perception, key themes, and demographic variations in opinions. This chapter consolidates the findings and presents a comprehensive discussion of the results, along with recommendations for addressing ethical concerns in AI systems.

Sentiment analysis unveils a mostly positive, leaning towards neutral, perception of AI ethics within ChatGPT conversations. A slight uptick in positivity appears when ethics-related keywords are involved, emphasizing the value placed on ethical discourse in AI conversations.

Topic modelling had the intention to uncover the main related topics within the dataset and it exposes similar insights as the sentiment analysis of the ethics-related keywords dataset, with a general positive to neutral sentiment around AI ethics in ChatGPT discussions. The importance of ethics is highlighted when these themes emerge in conversation, indicating a conscious engagement with ethical aspects of AI among users.

Diversity is key in demographic analysis, as sentiment scores on AI ethics differ across gender lines and national boundaries. Acknowledging these varying outlooks can help shape more comprehensive and inclusive ethical guidelines for AI, ensuring that our digital future is shaped by of diverse voices and inclinations.

5.1 Practical Implications

The analysis conducted offers various practical implications that could guide the engagement with AI systems in an ethical manner. One crucial insight gained from the study is the need for fostering an open dialogue and collaboration among diverse stakeholders. This approach would involve a diverse collective of academia, industry leaders, policymakers, and the general public in shaping the ethical frameworks surrounding AI. The aim is to build a well-rounded conversation that addresses potential concerns and maximizes the benefits of AI systems, which resonates with the finding of a generally positive perception of AI ethics in our data.

Moreover, inclusivity plays a central role in the way we envision the development of AI ethics. The representation of varied perspectives in the establishment of AI ethics guidelines and policies is critical to ensure fairness and avoid the marginalization of certain groups. The analysis showed differing sentiment scores across demographic groups, which highlights the importance of considering these variations in the formulation of AI ethics guidelines.

Education also appears as a key implication of the study. The development of resources and programs to increase awareness and understanding of AI ethics can lead to more informed discussions and decision-making. Conversations about AI ethics, as shown in the sentiment analysis, tend to be more positive, reinforcing the significance of education in this context.

Transparency is another significant aspect that emerges from the analysis. It's important to make AI systems' algorithms, data sources, and decision-making processes accessible and comprehensible to users. This transparency can foster trust in AI systems, which the analysis shows as a key ethical topic.

Finally, the study highlights the necessity of implementing strong mechanisms for overseeing and addressing various ethical concerns, such as bias, fairness, privacy, etc., in AI systems. Findings suggest that handling these issues effectively can help maintain a positive public perception of AI, which is vital for the successful integration of AI systems into society.

With these insights, developers can work towards addressing the ethical challenges posed by AI technology and ensure the responsible deployment of AI systems that have a positive impact on society.

5.2 Theoretical Contribution

This study contributes by exploring the sentiment and the main topics being conversed in social media, particularly Twitter, around the ethics of AI in the context of a large language model like ChatGPT. Echoing Munoko et al.'s (2020) approach, this study addresses a current gap in literature by endorsing a proactive approach to anticipating ethical implications of AI technology, specifically through the lens of societal discourse.

Moreover, the application of analytics techniques like Sentiment Analysis and The Latent Dirichlet Allocation underscores the potential of these methodologies to discover nuanced perspectives on AI ethics and provide insights into how AI is perceived from an ethical standpoint. These methodologies demonstrate the capacity to discern subtle shifts in sentiment and topic prominence, highlighting their value in the study of perceptions on rapidly advancing technologies like AI.

The findings from this study also contributes to valuable studies around sustainable development of AI such as (Floridi *et al.*, 2018; Hagendorff, 2019b; Li *et al.*, 2021; Vinuesa *et al.*, 2020), because the study provides a valuable theoretical counterpoint to concerns about public mistrust or fear of AI. This suggests a potential theoretical framework where the ethical implications of AI are not merely problems to be mitigated, but also opportunities for enhancing user trust and acceptance, shaping policy discussions, and refining AI technology.

In the broader context, the study contributes to theoretical discussions about the role of public sentiment in shaping the development and deployment of AI technology. It suggests that understanding public sentiment could be a key aspect of ethical AI development, indicating the need for a more democratized approach to AI technology.

5.3 Challenges, Limitations and Future Research

While this study offers valuable insights into the public perception of AI ethics in the context of Twitter #ChatGPT conversations, it is important to recognize the challenges and limitations inherent in the research process.

1. Processing Power and LDA Quality

A significant challenge faced during this study was the limited processing power available for performing Latent Dirichlet Allocation (LDA). As a result, the LDA model might not have provided the best possible quality of topics based on the keyword filter. The quality of LDA results depends on factors such as the size of the dataset, the complexity of the text, and the computational resources available. In future research, it would be beneficial to explore more advanced techniques or utilize more powerful computing resources to improve the quality and granularity of the topics generated.

2. Ambiguity and Subjectivity in Sentiment Analysis

Sentiment Analysis is inherently subjective and may not accurately capture the nuances and subtleties of human emotion. This limitation is particularly relevant when analysing text data from social media platforms like Twitter, where users often employ informal language, sarcasm, and emojis. Consequently, the sentiment scores derived from our analysis might not fully represent the true emotions and opinions of the users. Future research could explore the use of more advanced Sentiment Analysis techniques or incorporate additional contextual information to improve the accuracy of sentiment classification.

3. Incomplete or Biased Data

The dataset utilized in this study, comprising tweets containing the #ChatGPT hashtag, might not represent the complete range of opinions and sentiments related to AI ethics to do a statistically significant study. This language model es quite recent and by the date range where the information was captured new and improved models, such as ChatGPT 4 were barely open to the public. It is possible that certain demographics or viewpoints are underrepresented in the dataset, leading to biased or incomplete results. This is clearly visible when contrasting the sentiment analysis of topics with opinions of different communities such as “Reddit”, where forums such as r/ChatGPT or r/OpenAI seem to have a negative outlook on the efforts of Open AI to develop transparent and fair AI systems, showing their frustration when ChatGPT begins an answer with “As an AI language model… I cannot… I’m not programmed to…” after controversial conversations. Additionally, the use of specific keywords for filtering tweets could result in a skewed representation of the topic.

Future research should aim to incorporate more diverse data sources or implement different keyword filters to mitigate potential bias and achieve a more comprehensive understanding of the public perception of AI ethics.

5.4 Future Research

In contemplating the potential directions for future research, several possibilities emerge. A natural extension of this study would be to expand the analysis beyond Twitter and engage with data from other social media platforms like Reddit. Researchers could develop a richer, more diverse understanding of public discourses on AI ethics.

Additionally, an exploration of AI ethics in differing contexts could yield fascinating findings. While this study focused on ChatGPT, future research might examine the public perception of ethics in other AI applications, such as healthcare, autonomous vehicles, or algorithmic decision-making.

6 References

Abrams, D. and Hogg, M.A. (1988) “Comments on the motivational status of self-esteem in social identity and intergroup discrimination,” *European Journal of Social Psychology*, 18(4), pp. 317–334. Available at: https://doi.org/10.1002/ejsp.2420180403.

Arrieta, A. *et al.* (2019b) “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” *Information Fusion*, 58, pp. 82–115. Available at: https://doi.org/10.1016/j.inffus.2019.12.012.

Awad, E. *et al.* (2018) “The Moral Machine experiment,” *Nature*, 563(7729), pp. 59–64. Available at: https://doi.org/10.1038/s41586-018-0637-6.

Bubeck, S. *et al.* (2023) “Sparks of Artificial General Intelligence: Early experiments with GPT-4,” *arXiv (Cornell University)* [Preprint]. Available at: https://doi.org/10.48550/arxiv.2303.12712.

Dignum, V. (2018). “Ethics in artificial intelligence: introduction to the special issue” Ethics and Information Technology (2018) 20:1–3

Evans, G.W. (2017) “Disruptive Technology and the Board: the Tip of the Iceberg,” *The Poznań University of Economics Review*, 3 (17)(1), pp. 205–223. Available at: https://doi.org/10.18559/ebr.2017.1.11.

Floridi, L. *et al.* (2018) “AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations,” *Minds and Machines*, 28(4), pp. 689–707. Available at: https://doi.org/10.1007/s11023-018-9482-5.

Griva, A. *et al.* (2018b) “Retail business analytics: Customer visit segmentation using market basket data,” *Expert Systems With Applications*, 100, pp. 1–16. Available at: https://doi.org/10.1016/j.eswa.2018.01.029.

Hagendorff, T. (2019) “The Ethics of AI Ethics: An Evaluation of Guidelines,” *Minds and Machines*, 30(1), pp. 99–120. Available at: https://doi.org/10.1007/s11023-020-09517-8.

Katznelson G, Chan B. What Does Artificial Intelligence Mean for Oncologists?. UWOMJ [Internet]. 2021 Apr. 3 [cited 2023 Apr. 6];89(1). Available from: https://ojs.lib.uwo.ca/index.php/uwomj/article/view/9365

Kieslich, K., Keller, B. and Starke, C. (2021) “AI-Ethics by Design. Evaluating Public Perception on the Importance of Ethical Design Principles of AI.,” *arXiv (Cornell University)* [Preprint]. Available at: https://arxiv.org/pdf/2106.00326.pdf.

Li, B. *et al.* (2021) “Trustworthy AI: From Principles to Practices,” *ACM Computing Surveys*, 55(9), pp. 1–46. Available at: https://doi.org/10.1145/3555803.

Manikonda, L. and Kambhampati, S. (2018) “Tweeting AI: Perceptions of Lay versus Expert Twitterati,” *International Conference on Weblogs and Social Media*, 12(1). Available at: https://doi.org/10.1609/icwsm.v12i1.15061.

Martínez, M. and Fernández, A. (2019) “AI in Recruiting. Multi-agent Systems Architecture for Ethical and Legal Auditing,” *International Joint Conference on Artificial Intelligence* [Preprint]. Available at: https://doi.org/10.24963/ijcai.2019/903.

Maxim, S.T. (2014) “Ethics: Philosophy or Science?,” *Procedia - Social and Behavioral Sciences*, 149, pp. 553–557. Available at: https://doi.org/10.1016/j.sbspro.2014.08.209.

Metz, C., & Schmidt, G. (2023, March 29). *Elon Musk and Others Call for Pause on A.I., Citing ‘Risks to Society.’* The New York Times. https://www.nytimes.com/2023/03/29/technology/ai-artificial-intelligence-musk-risks.html

Mikalef, P. *et al.* (2022b) “Thinking responsibly about responsible AI and ‘the dark side’ of AI,” *European Journal of Information Systems*, 31(3), pp. 257–268. Available at: https://doi.org/10.1080/0960085x.2022.2026621.

Munoko, I. *et al.* (2020c) “The Ethical Implications of Using Artificial Intelligence in Auditing,” *Journal of Business Ethics*, 167(2), pp. 209–234. Available at: https://doi.org/10.1007/s10551-019-04407-1.

Neff, G. and Nagy, P. (2016b) “Automation, Algorithms, and Politics| Talking to Bots: Symbiotic Agency and the Case of Tay,” *International Journal of Communication*, 10, p. 17. Available at: https://ijoc.org/index.php/ijoc/article/view/6277.

Palomino, M. A., & Aider, F. (2022). Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis. *Applied Sciences*, *12*(17), 8765. https://doi.org/10.3390/app12178765

Pang, B. and Lee, L. (2008) “Opinion Mining and Sentiment Analysis,” *Foundations and Trends in Information Retrieval*, 2(1–2), pp. 1–135. Available at: https://doi.org/10.1561/1500000011.

Siau, K. & Wang, W. (2020). Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI. *Journal of Database Management (JDM)*, *31*(2), 74-87. http://doi.org/10.4018/JDM.2020040105

Shearer, C. (2000). The CRISP-DM model: the new blueprint for data mining. *Journal of data warehousing*, *5*(4), 13-22

Stahl, B.C. (2021) "From computer ethics and the ethics of AI towards an ethics of digital ecosystems," *AI And Ethics*, 2(1), pp. 65–77. Available at: https://doi.org/10.1007/s43681-021-00080-1.

Taecharungroj, V. (2023) "'What Can ChatGPT Do?' Analyzing Early Reactions to the Innovative AI Chatbot on Twitter," *Big Data and Cognitive Computing*, 7(1), p. 35. Available at: https://doi.org/10.3390/bdcc7010035.

Vinuesa, R. *et al.* (2020) "The role of artificial intelligence in achieving the Sustainable Development Goals," *Nature Communications*, 11(1). Available at: https://doi.org/10.1038/s41467-019-14108-y.

Wang, L. *et al.* (2021) "Artificial Intelligence and Economic Growth: A Theoretical Framework," *Scientific Annals of Economics and Business*, 68(4), pp. 421–443. Available at: https://doi.org/10.47743/saeb-2021-0027.

Wright, S.A. and Schultz, A.E. (2018) "The rising tide of artificial intelligence and business automation: Developing an ethical framework," *Business Horizons*, 61(6), pp. 823–832. Available at: https://doi.org/10.1016/j.bushor.2018.07.001.

Zhang, D., Luo, T.-Y. and Tang, B.Z. (2016) "Learning from LDA Using Deep Neural Networks," in *Lecture Notes in Computer Science*. Springer Science+Business Media, pp. 657–664. Available at: https://doi.org/10.1007/978-3-319-50496-4_59.