

**Investigating the Use of the Term ‘Narcissist’ on Twitter:
A Text Mining Approach**

Josie Vorhauer

s2079704

Faculty of Behavioural, Management and Social Sciences (BMS), University of Twente

Master Thesis Positive Clinical Psychology and Technology

First Supervisor: Dr. Peter ten Klooster

Second Supervisor: Dr. Jorge Piano Simões

25th September 2024

Acknowledgements

Writing this thesis was the last step I needed to take to become a psychologist. I would like to express my sincere gratitude towards everyone who has helped on this journey. First and foremost, I would like to express my gratitude to Dr. Peter ten Klooster and Dr. Jorge Piano Simões for their guidance throughout this thesis. I also would like to thank Paul Mittelstädt for his invaluable support during the last months of my master's programme. Furthermore, I would like to thank my grandmother, Sigrid Ullrich, for always believing in me and, quite frankly, supporting me financially throughout my studies. Last but certainly not least, I dedicate this thesis to the strongest woman I have ever met, Silvia Vorhauer, whom I am also lucky enough to call my mother. Without her unconditional love and support, I would have never been able to fulfil my dream of studying psychology.

Table of Contents

Abstract	4
Investigating the Use of the Term “Narcissist” on Twitter: A Text Mining Approach.....	5
Background and Research Relevance	6
Text Mining	8
Twitter	8
Present Study.....	9
Method	9
Data Collection.....	10
Data Pre-processing.....	10
Topic Modelling	11
Sentiment Analysis.....	13
Results	14
Topic Modelling	15
Antisocial & Pathological Traits	17
Donald Trump & U.S. Elections	20
Narcissistic Abuse & Relationships	21
Sentiment Analysis.....	23
Sentiment Analysis Per Topic.....	23
Discussion	24
Antisocial & Pathological Traits	24
Narcissistic Abuse & Relationships	25
Politics and Donald Trump.....	27
Implications.....	28
Strengths and Limitations.....	29
Conclusion.....	30
References	31
Appendix A	40
Appendix B	41
Appendix C	42
Appendix D	42

Abstract

Objective: Mental health care professionals have been increasingly discussing the misuse of the term ‘narcissist’ on social media, warning of a stigmatization of people with a narcissistic personality disorder (NPD). However, no empirical study has yet systematically explored the use of the term ‘narcissist’ on social media. Using a text mining approach, the aim of the current study was to investigate the use of and sentiment towards the term ‘narcissist’ on Twitter.

Method: The data consisted of all tweets that included the keyword ‘narcissist’ posted between the 19th and 26th of March 2023. The tweets were scraped and imported into the data mining toolkit *Orange*, where the tweets were pre-processed and analysed using LDA topic modelling and a VADER-based sentiment analysis. The final corpus consisted of 30,205 tweets.

Results: The word frequency analysis showed that *trump* was the most frequently used word in the corpus followed by *people*, *lie*, “cross mark” emoji, *man*, *malignant*, *love*, *liar*, *life* and *woman*. Topic modelling revealed that the most prevalent topic was Narcissistic Abuse & Relationships, with 28.56% of all tweets discussing this topic, followed by Politics (24.49%), Donald Trump & U.S. Elections (20.73%) and Antisocial & Pathological Traits (13.15%). Twitter users expressed on average a negative sentiment when they used the term ‘narcissist’, with nearly half of the tweets (47.53%) expressing a negative sentiment. The sentiments expressed across all topics were also on average negative. There was a statistically significant difference in mean compound sentiment scores between topics [$F(4, 30200) = 97.47, p < .001$]. Twitter users expressed the most negative sentiment in Tweets about Donald Trump & U.S. Elections.

Conclusion: The results of the current paper suggest that Twitter users frequently used the term ‘narcissist’ in personal contexts (e.g., abuse within relationships), and in political discussions (e.g., Donald Trump and elections). Furthermore, the overall negative sentiment of the tweets implies that the term ‘narcissist’ was mostly used in negatively connotated contexts. This colloquial and negative use of the term ‘narcissist’ may contribute to the stigmatization of NPD. Future research may investigate if negatively connotated tweets about narcissists (e.g., tweets about so-called narcissistic abuse) may lead to an increase in negative attitudes or change in intended behaviour towards people with NPD and whether clinical information about NPD can mitigate these effects.

Keywords: narcissistic personality disorder, social media, psychiatric terms, topic modelling, sentiment analysis, social media, LDA, VADER, Twitter

Investigating the Use of the Term “Narcissist” on Twitter: A Text Mining Approach

The term ‘narcissist’ is derived from the legend of Narcissus in Greek mythology (Dieckmann & Behary, 2015; Yakeley, 2018). Narcissus, a beautiful young man, cruelly rejects the advances of the mountain nymph Echo, leaving her heartbroken. As a punishment for rejecting Echo, Narcissus is lured to a pool of water by the Nemesis, the god of revenge. When Narcissus catches the image of himself in the water, he falls in love with it. Unable to take his eyes off his reflection, he eventually passes away from starvation and thirst (Dieckmann & Behary, 2015; Yakeley, 2018). This tale about Narcissus has influenced the idea of narcissists in the media, literature, and public for ages (Dieckmann & Behary, 2015; Yakeley, 2018). The use of the term ‘narcissist’ or ‘narcissism’ in a clinical context began in the early 20th century, when psychoanalysts described narcissism as a personality trait characterized by vanity and self-love (Yakeley, 2018). In these early publications, psychoanalysts, such as Otto Rank (1911) and Sigmund Freud (1914), conceptualized narcissism as a dimensional psychological state ranging from normal to pathological, and a normal part of human development as opposed to a disorder (Levy et al., 2011; Yakeley, 2018).

Subsequent papers on the concept of pathological narcissism by psychoanalysts, such as Otto Kernberg and Heinz Kohut, in the 1970s gave rise to the clinical interest in narcissism and eventually led to the inclusion of the narcissistic personality disorder (NPD) into the third edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1980; Levy et al., 2011; Yakeley, 2018). In the DSM-5, the most recent version of the DSM, NPD belongs to the Cluster B personality disorders, which are characterized by dramatic, emotional and erratic patterns of behaviour (American Psychiatric Association, 2013; Kacel et al., 2017). NPD is described as a “a pervasive pattern of grandiosity (fantasy or behaviour), need for admiration, and lack of empathy [...]” (American Psychiatric Association, 2013, p. 669). It is indicated by a grandiose sense of self-importance, preoccupation with fantasies of unlimited success or power, belief that one is special, need for excessive admiration, sense of entitlement, interpersonal exploitation, lack of empathy, envy of others, and arrogant behaviors or attitudes (American Psychiatric Association, 2013; Dieckmann & Behary, 2015; Skodol et al., 2014; Yakeley, 2018). In a systematic review of the prevalence of NPD in adult non-clinical settings, Dhawan et al. (2010) found a mean prevalence of 1.06% with a range from 0% to 6.2%. NPD is frequently comorbid with substance use, anxiety, mood, and other personality disorders (Stinson et al., 2008).

Notably, the DSM-5’s diagnostic criteria of NPD have been repeatedly criticized for

primarily describing grandiose narcissism, while excluding vulnerable aspects of narcissism (Skodol et al., 2014; Yakeley, 2018). According to Dieckmann and Behary (2015), the DSM-5 criteria may negatively impact the perception of people with NPD, as one may overlook that behind the display of grandiosity, people with NPD often experience psychological distress, feeling of emptiness and loneliness. Similarly, the diagnostic criteria mainly describe overt narcissism, while neglecting covert forms of narcissism, where narcissistic individuals are shy, hypersensitive and only live out their sense of grandiosity in their fantasies (Dieckmann & Behary, 2015; Skodol et al., 2014).

Within recent years, journalists and mental health care professionals have been increasingly discussing the misuse of the term ‘narcissist’ on social media (Dodgson & Colombo, 2022; Berman, n.d.; Lefort, 2022; Showfety, 2022; Webb, 2022). Social media users seem to use the term narcissist to refer to various people, ranging from friends or family, colleagues to public figures, such as influencers or politicians (Dodgson & Colombo, 2022; Berman, n.d.; Lefort, 2022; Webber, 2016). For instance, countless people use social media platforms, such as TikTok or Twitter, to talk about being abused by their allegedly narcissistic family members or friends (Showfety, 2022). It has been argued that social media users label other people as narcissists without understanding the psychiatric significance of term or the complexity of diagnosing NPD, which requires the evaluation by a licensed mental health professional (Dodgson & Colombo, 2022; Lefort, 2022). Furthermore, social media users attempt to use the term in a clinical context as well. Frequently shared pseudoscientific posts such as “15 signs you are dealing with a narcissist’ even seem to encourage social media users to “diagnose” their family, friends, or colleagues with narcissism (Lefort, 2022).

Background and Research Relevance

One academic paper has stated that the public (mis)use of the term has turned ‘narcissist’ into a meaningless and stigmatizing label, which is colloquially used to discredit any individual with self-serving characteristics, arguing that the term narcissist should be “reclaimed” for clinical and research purposes (Freestone et al., 2020). Nevertheless, no empirical studies have yet systematically investigated how the term narcissist is used on social media. However, there is some existing research on the use of other psychiatric terms on social media (Athanasopoulou & Sakellari, 2016; Delanys et al., 2022; Joseph et al., 2015; Lachmar et al., 2017; Li et al., 2020; Reavley & Pilkington, 2014). For instance, two studies investigated the use of the word schizophrenia on Twitter and found that tweets about schizophrenia were more likely to be medically inappropriate, sarcastic, express a negative sentiment, and used in a non-medical context compared to tweets about diabetes

(Athanasopoulou & Sakellari, 2016; Joseph et al., 2015). Similarly, Alvarez-Mon et al. (2019) compared the use of the term psychosis with physical diseases (breast cancer, diabetes, Alzheimer, and HIV) on Twitter. Their study found that tweets about psychosis were more likely to be misused (33.3% vs. 1.15%) and include pejorative content (36.2% vs. 11.3%) compared to tweets about physical diseases.

Investigating differences between the use of different psychiatric terms, Reavley and Pilkington (2014) found that tweets about schizophrenia express more often a stigmatizing attitude (5%) than tweets about depression (0.7%). Notably, however, Reavley and Pilkington (2014) found that the majority of tweets about both depression and schizophrenia included supportive content, such as services for people for depression and attempts to raise awareness about schizophrenia (Reavley & Pilkington, 2014). This is in contrast with findings from Delanys et al. (2022), who investigated the use of various psychiatric terms, such as schizophrenia and autism, on Twitter. They found that 45.33% of the analysed tweets misused psychiatric terms, while 42.37% used psychiatric terms in a medical context. Notably, more than half the tweets (52.14%) conveyed a negative sentiment. Furthermore, their study replicated the finding that the term depression was less often misused (5.36%) than the term schizophrenia (55.46%) (Delanys et al., 2022).

These research findings raise the question of how this may relate to the term 'narcissist'. Since various negative traits and behaviours (e.g., lack of empathy, interpersonal exploitation, arrogant behaviours) are symptomatic of NPD, it is to be expected that social media posts that use the term 'narcissist' are likely to include negative or possibly even stigmatizing content. However, while it is generally assumed that NPD is a highly stigmatized disorder, there are no empirical studies investigating the stigma attached to NPD (Penney et al., 2017). Research has argued that with the increasing use and popularity of social media, it is crucial to explore how psychiatric terms are used on social media, as it may provide important information about the stigmatization of mental disorders (Alvarez-Mon et al., 2019; Delanys et al., 2022). Therefore, scientific research on the use of the term 'narcissist' online is timely because if the term is used colloquial in an overly negative way on social media, it can lead to stereotypes of people with NPD, potentially contributing to a stigmatization of NPD.

Stigmatization of mental disorders is a major concern among mental health professionals, as it is associated with various negative effects, such as feelings of shame and increased reluctance to seek treatment (Da Silva et al., 2020; Sickel et al., 2014). This may be especially problematic for people with NPD, as the treatment of NPD already poses several challenges. People diagnosed with NPD are prone to only seek therapy after multiple failed

previous treatment attempts, prematurely end therapy and reject their diagnosis (Yakeley, 2018). Furthermore, people with NPD regularly have comorbid mental health problems, such as depression, and an increased suicide risk, especially after experiencing narcissistic injury (Gabbard, 2022; Yakeley, 2018). Narcissistic injury occurs when people with NPD experience a threat or attack on self-worth and pride, leading to emotional distress, feelings of being shamed or vilified (Gabbard, 2022).

Text Mining

Most of the research which investigated the use of psychiatric terms on social media has used some form of qualitative content analysis (Delanys et al. 2022; Lachmar et al., 2017; Reavley & Pilkington, 2014). Qualitative content analysis, however, requires researchers to manually read and analyse all the documents in the dataset, making it an unsuitable approach to analyse unstructured and large datasets (i.e., big data), such as the ever-growing number of posts on social media (Asmussen & Møller, 2019; Gandomi & Haider, 2015). For this reason, computer-assisted techniques, such as text mining, which allow for a (semi)automated analysis of textual documents, have become increasingly crucial for processing and analysing big data (Asmussen & Møller, 2019; Georgieva-Trifonova & Dechev, 2021).

Text mining is the process of transforming large quantities of unstructured textual data into a structured format to discover and extract meaningful information (Abbe et al., 2015; Namugera et al., 2019; O'Mara-Eves et al., 2015; Thakur & Kumar, 2021). Text mining is an interdisciplinary process, integrating methods from “statistical analysis, computational linguistics, and machine learning” (Gandomi & Haider, 2015, p.140). Natural language processing (NLP) is a core component of text mining. NLP employs machine learning to make computers able to process and analyse natural language text or speech. By using NLP, text mining identifies relationships between words or sentences in the data to extract meaningful information (Karami et al., 2020). There are various text mining techniques that researchers can use to extract information and patterns from their data, such as topic modelling and sentiment analysis.

Twitter

Twitter is a social networking and microblogging service, that is commonly used as a source for text mining studies (Karami et al., 2020). On Twitter, users can post so-called tweets, which are short messages consisting of up to 280 characters. With about 369 million active users posting more than 500 million tweets a day, Twitter provides a large amount of valuable data, that can be analysed to extract meaningful information about the use of the term ‘narcissist’ (Karami et al., 2020; Oscar et al., 2017; Statista, 2023; Wongkoblap et al., 2017).

Furthermore, Twitter is commonly used to share personal experiences, thoughts, opinions, and emotional status updates (Laing et al., 2017; Murthy et al., 2015). Lachmar et al. (2017) argue that this makes twitter a unique data source for exploring potentially stigmatized topics.

Moreover, Twitter provides its users with the option to stay anonymous, which may make individuals more inclined to express stigmatizing attitudes and beliefs compared to other social media platforms, such as Facebook (Joseph et al., 2015).

Present Study

The aim of the current study was to investigate the use of the term ‘narcissist’ on Twitter. This investigation was guided by the following research questions:

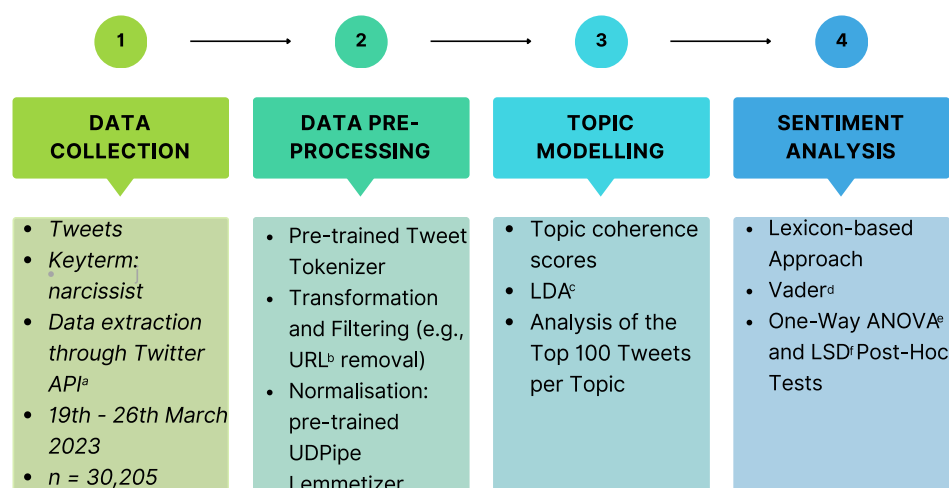
- 1.) What topics do Twitter users discuss when they use the term ‘narcissist’ in a tweet?
- 2.) What sentiment do Twitter users express when they use the term ‘narcissist’ in a tweet?
- 3.) What sentiments do Twitters users express regarding the topics they discuss when they use the term ‘narcissist’ in a tweet?

Method

Taking an exploratory approach, the present study utilized text mining techniques to analyse tweets that mention the term ‘narcissist’. The current study consisted of multiple consecutive steps to gain insights into how the term is used on Twitter (see Figure 1).

Figure 1

Pipeline of the Study Design



Note. This pipeline depicts the study design in four consecutive steps.

^a Application Programming Interface. ^b Uniform Resource Locator.

^c Latent Dirichlet Allocation. ^d Valence Aware Dictionary and sEntiment

reasoner. ^e Analysis of Variance. ^f Least Significant Difference.

Data Collection

The data collected consisted of tweets that included the key term ‘narcissist’. Only tweets written in the English language were included. Retweets were excluded. While retweets can provide insights about the engagement and the popularity of certain topics or opinions, retweets also produce noise in the dataset by repeating the same content and sentiment, hindering the model’s ability to extract meaningful information about the diversity of topics discussed on Twitter (Alizadeh, 2022). The tweets were scraped using Twitter’s Academic Application Programming Interface (API). Due to Twitter’s data limitation policy, the non-commercial API enables researchers to only access tweets of the last seven days (Karami et al., 2020; Wongkoblapp et al., 2017). The tweets were collected between the 19th and 26th of March 2023. The collected data included the content of the tweets but also metadata (e.g., username, time of publication, follower count etc.). The data was processed and analysed using Orange 3.35.0, which is a “machine learning and data mining suite for data analysis through Python scripting and visual programming” (Demšar et al., 2013, p.2349).

Data Pre-processing

As tweets are unstructured data, they contain vast amounts of elements which are irrelevant for the extraction of meaningful information by machine learning algorithms, such as natural text elements (e.g., punctuation marks, verb inflections, and capitalization) and unreadable data (e.g., URLs, HTML tags) (Agrawal, 2022). The aim of pre-processing, therefore, is to structure the data in a format that can be analysed by machine learning algorithms by removing or transforming irrelevant or unreadable text elements (Agrawal, 2022; Aydin, 2023; Wongkoblapp, 2017). The first step of the data pre-processing process was *tokenization*, which is the process of splitting a text into separate components (i.e., tokens), such as words, numbers, and punctuation (Wiegersma et al., 2022). *Orange’s* Pre-trained Tweet Tokenizer was used, as this tokenizer retains hashtags, emoticons and other symbols commonly used on Twitter.

The second step was *transforming* and *filtering*, which consisted of multiple different tools aimed at converting or removing certain text elements. First, all uppercase letters were converted into lowercase, as machine learning algorithms would categorize capitalized and uncapitalized versions of the same word (e.g., “ball” and “Ball”) as two different tokens (Agrawal, 2022). For the same reason, all accents were removed. Second, all URLs were removed, and HTML tags were parsed as neither added meaningful information to the text. Third, all punctuation marks were removed, as they would otherwise make up a substantial

share of the tweets without adding meaningful information (Agrawal, 2022). Numbers were not removed, as certain numbers (e.g., dates) might yield important insights. Stop words were removed. According to Alshanik et al., (2020) stop words are words that are frequently used in text data but add no or only little meaning and value to the analysis (e.g., “the”, “to”, “and”, “a”). *Orange’s* Open-Source Stop Words List for the English dictionary list was used. However, common stop word lists do not necessarily remove all frequently used but irrelevant words in a domain-specific corpus (Alshanik et al., 2020). Hence, after data visualization and inspection, a manually customized stop word list was added (see Appendix A). This list included, for instance, the key term ‘narcissist’ as this term occurred in all tweets by design and other irrelevant words in the corpus (e.g., typical, almost, either).

The final step was *normalization*, which refers to the process of transforming different variants of a term into a single canonical or standard form (Chitrapriya et al., 2017). Stemming and lemmatization are two commonly used techniques aimed at reducing inflected words to their base form to ensure that different forms of a word are not categorized as separate tokens (Bhoi, 2022). The current study employed lemmatization, as stemming only removes common suffixes from words, which often leads to meaningless words. For instance, the words “changing”, “changes” or “changed” would all be reduced to the non-existent word “chang”. Lemmatization, in contrast, would reduce the words to the correct form “change”, as lemmatization reduces words to their lemma (i.e., the canonical or dictionary form of a word) (Bhoi, 2022). *Orange’s* pre-trained UDPipe Lemmatizer for the English language was used, as manual inspection and visualization of the data showed that this model had the best interpretability.

Topic Modelling

To explore what topics Twitter users are discussing when they use the term ‘narcissist’, topic modelling was employed. Topic modelling is a “a statistical technique for revealing the underlying semantic structure in a large collection of documents” (Kherwa & Bansal, 2018, p. 1). It is used to discover what latent topics repeatedly occur in a set of text documents (Cho, 2019). To date, only the topic modelling algorithms Latent Semantic Indexing (LSI), Latent Dirichlet Allocation (LDA), Hierarchical Dirichlet Process (HDP) are available on *Orange*. Out of the three, LDA was deemed the most suitable algorithm for the current study. LDA is a generative probabilistic model used to extract hidden topics from textual data (Asmussen & Møller, 2019; Blei et al., 2003; Jelodar et al., 2018). The basic assumption of LDA is that documents are a mixture of various latent topics, and each topic is characterized by a probability distribution of words (Blei et al., 2003). LDA is the most popular and well-

established method for topic modelling (Kobayashi et al., 2017). In a study by Negara et al. (2019), LDA outperformed LSI in word indexing Sport topics on Twitter with an accuracy of 98%. Wold et al. (2016) found that the author-topic (AT) LDA model outperformed HDP in topic modelling tweets as part of a breaking news detection system.

To choose the topic model, the topic coherence score was computed. Topic coherence is a measure of “the degree of semantic similarity between high-scoring words in the topic” (Kapadia, 2019, What is Topic Coherence? section). In other words, each topic is represented by keywords that have the highest probability of belonging to the respective topic and the topic coherence scores shows how similar these high scoring words are to each other (Zvornicanin & Martin, 2024). The range of topic coherence scores is usually scaled from 0 to 1, with scores close to 1 indicating high similarity, and coherence scores close to zero indicating low similarity (Mimno et al., 2011). The coherence score normally increases with an increasing number of topics, so determining the topic solely based on the highest topic coherence score is not always practical. Zvornicanin & Martin (2024) suggest finding a balance between a manageable number of topics and a strong coherence score by finding the point after which there is only a minimal increase in coherence scores (i.e., the elbow method). After manual inspection, the final topic solution was chosen by considering both the topic coherence score and interpretability of the topics.

After the number of topics was determined, the generated topics were labelled and interpreted with the assistance of *ChatGPT 3.5*. *ChatGPT* was developed by OpenAI and launched in November 2022. *ChatGPT* is a generative “pre-trained AI mode designed to engage in natural language conversations, utilizing sophisticated techniques from Natural Language Processing (NLP), Supervised Learning, and Reinforcement Learning to comprehend and understand human-generated text” (Roumeliotis & Tselikas, 2023, p.1). *ChatGPT 3.5* is a recent and ever-changing development. However, there is preliminary research which suggests that *ChatGPT* is capable of interpreting topics accurately (compared to interpretation by domain experts) and can provide useful insights for topic modelling (Rijcken et al., 2023). The topics were first labelled by the researcher based on the top ten keywords of each topic, and then *ChatGPT* was prompted to label the topics. The final topic label was assigned by the researcher. The interpretation of the topics was also first conducted by the researcher. If some keywords were unclear or ambiguous, *ChatGPT* was further prompted to interpret the topics to minimize interpretation bias. As the ten most dominant keywords only give a limited insight into the topics, an additional word cloud of the 200 keywords that represent each topic was interpreted manually. To gain an even more in-depth

understanding of the topics, the 100 tweets with the highest loading on each respective topic were manually read and interpreted by the researcher.

Sentiment Analysis

To investigate what sentiment Twitter users express when they mention the term ‘narcissist’, a lexicon-based sentiment analysis was employed. Sentiment analysis uses methods from computer linguistics, machine learning, and artificial intelligence to identify the emotional valence of text documents (Gillis & Barney, 2024). That is, sentiment analysis provides insight into how people feel about a certain topic. The lexicon approach is an unsupervised method, which classifies the sentiment of tweets in predefined sentiment categories (i.e., negative, positive, or neutral sentiment) (Hota et al., 2021; Karami et al., 2020; Öztürk & Ayvaz, 2018). The sentiment is determined based on a sentiment lexicon, which is a list of words with an assigned sentiment polarity (i.e., positive or negative) (Kaity & Balakrishnan, 2020; Öztürk & Ayvaz, 2018).

The current study applied the Valance Aware Dictionary and sEntiment Reasoner (VADER) approach, a rule-based sentiment analysis tool. VADER sentimental analysis uses a sentiment lexicon that indicates not only sentiment polarity but also sentiment intensity on a scale from -4 to +4 (Hutto & Gilbert, 2014). So, for example, the word ‘good’ has a positive valence of 1.9, while the word ‘horrible’ has a negative valence of -2.5 (Hutto & Gilbert, 2014). The compound score is then calculated by summing the valence scores of each word, which is then normalized between -1 (most extreme negative) and +1 (most extreme positive) (Hota et al., 2021). To classify the compound score as either positive, neutral, or negative, Hutto (n.d.), one of the developers of VADER, recommends considering a compound score of ≥ 0.05 a positive sentiment and a compound score of ≤ -0.05 a negative sentiment. A compound score between > -0.05 and < 0.05 is considered a neutral sentiment. This scoring system has also been used in other research (Hota et al., 2021).

Compared to other lexicons, such as the commonly used Linguistic Inquiry and Word Count (LIWC), VADER is more sensitive to the sentiment expressed on social media (Hota et al., 2021; Hutto & Gilbert, 2014). For instance, the VADER lexicon includes special lexical features for microblogging contexts, such as emoticons (e.g., :-)) signals a positive sentiment), acronyms and initialisms (e.g., LOL, WTF), and slang words (e.g., meh or nah) (Hutto & Gilbert, 2014). Furthermore, VADER outperforms human raters at correctly classifying the sentiment of tweets (Hutto & Gilbert, 2014). To analyse the sentiment Twitter users express regarding the individual topics, every tweet was assigned the topic the particular tweet loaded the highest on and then the mean compound sentiment scores per topic were calculated using

the Software Statistical Package for the Social Sciences (SPSS) 27.0.0.0. Lastly, a One-way Analysis of Variance (ANOVA) with Post Hoc Least Significant Difference (LSD) test was conducted to check if there were statistically significant differences in mean compound sentiment scores between topics

Results

The present study analysed a corpus of 30,205 tweets posted by 22,998 unique users between March 19th and 26th, 2023. After pre-processing, the corpus consisted of 317,185 tokens in total, out of which 46,330 tokens were unique. The word frequency analysis revealed that the ten most frequent tokens in the corpus were: *trump*, *people*, *lie*, “cross mark” emoji, *man*, *malignant*, *love*, *liar*, *life* and *woman* (Table 1).

Table 1

The Ten Most Frequent Tokens in Tweets that Mention the Term ‘narcissist’ After Pre-processing

Term	Frequency (Percentage)
trump	3326 (1.05%)
people	2679 (0.84%)
lie	1919 (0.61%)
“cross mark” emoji	1683 (0.53%)
man	1498 (0.47%)
malignant	1490 (0.47%)
love	1428 (0.45%)
liar	1243 (0.39%)
life	1097 (0.35%)
woman	1086 (0.34%)
Total	17,449 (5.50% ^a)

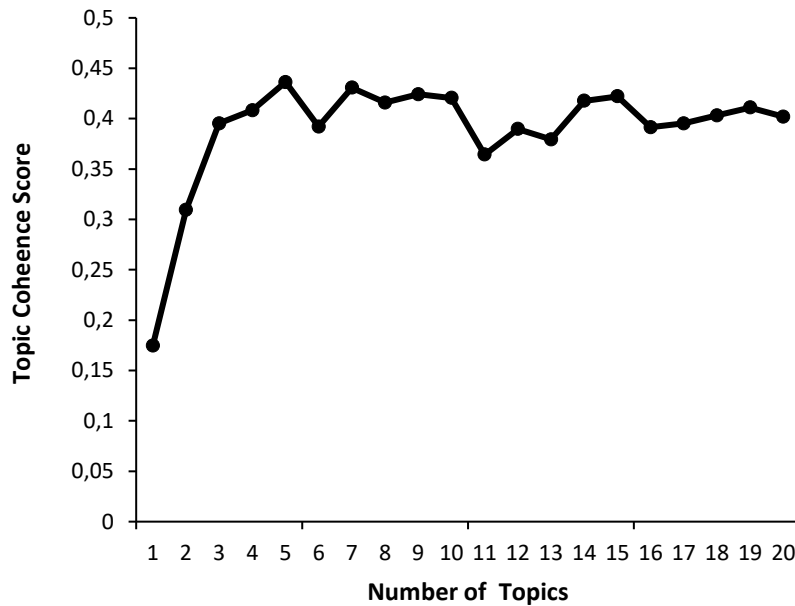
^aThe ten most prevalent tokens together made up 5.50% (n = 17,449) of all tokens.

A key finding was that *trump* was the most frequent word in tweets that mention the term ‘narcissist’. As can be seen in Figure 2, the word *donald* was also among the 200th most used words. That is, Twitter users frequently mentioned Donald Trump when they used the term ‘narcissist’. This may also be related to the frequent use of the “cross mark” emoji, as this symbol has been increasingly used to express support for Donald Trump on social media. Besides Donald Trump, Twitter users frequently mentioned three other politicians, namely Joe

0.44 ($k = 5$) (see Figure 3). The five-topic solution was deemed to be the most suitable topic model. It had the highest coherence score (0.44), after which the coherence scores decreased and flattened out. Also, manual inspection showed that the five-topic solution had the best interpretability, as it showed a high within-topic coherence and a low overlap between topics.

Figure 3

Topic Coherence Scores as a Function of Number of Topics (1-20)



The five generated topics were first labelled by the researcher and then by ChatGPT 3.5 and lastly, the label was manually adjusted by the researcher (see Appendices B and C). While ChatGPT 3.5. labelled the second topic ‘Emotional Reactions and Narcissism’, the researcher deemed the keywords too ambiguous to be interpreted properly and therefore, the second topic was removed from further analyses (see Appendices B and C). Table 2 shows the final four topic labels, the top ten keywords of each topic and the topic weights (i.e., the number of tweets that each topic loads the highest on and the percentage of all tweets).

Table 2

Keywords of the Four Extracted Topics with Topic Weights

Topic Label	Top Ten Keywords	Weight (%)
Antisocial & Pathological Traits	liar, pathological, self, sociopath, corrupt, serve, grifter, love, god	3971 (13.15%)
Politics	people, lie, trump, attention, 2020, right, real, man, foul, malignant	7396 (24.49%)

the words *self*, *serve*, *love*, and *god* fit into this topic, as they are, at first glance, a contrast to antisocial or pathological traits. However, the lemmatized words *serve* and *self* may belong together, indicating that Twitter users discussed individuals who self-serving, i.e., only care about their own interest. *God* may refer to the exclamation “oh my god”, which is frequently used on social media to express frustration, disbelief or, alternatively, excitement.

Alternatively, it may also suggest that Twitter users discussed pathological traits or behaviours in religious contexts. The use of the word *love* may indicate discussions about antisocial or pathological behaviour in romantic relationships.

According to *ChatGPT 3.5*, the use of the words *serve*, *love*, and *god* may also mean that Twitter users talked about people who exploit noble concepts, such as love and religion, to manipulate others (see Appendix D). Overall, the keywords suggest that Twitter users who mentioned the term narcissist in a tweet also often used terms referring to other antisocial and pathological traits. This topic interpretation of the top ten keywords is supported by other words representing the first topic, such as *conman*, *criminal*, *psychopath*, *borderline*, *racist*, *compulsive*, *adulterer*, *malignant* and *traitor*, which refer to additional antisocial and pathological traits (see Figure 4). The 100 tweets with the highest topic loading also supported the topic interpretation, as Twitter users predominately called other people names like narcissist, sociopath, or liar or accused them of engaging in various antisocial behaviours, such as corruption or fraud. For instance, the tweet with the highest topic loading of 0.9714 stated as follows:

[...]¹ SO LONG...DON THE CON
 *CHAOTIC NARCISSIST & SOCIOPATH.
 *PATHOLOGICAL LIAR.
 *TAX FRAUD.
 *POLITICAL & SOCIAL DIVIDER.
 *TRAITOR
 [...]
 *PANDERS TO RACISTS.
 *INCOMPETENT WITH COVID
 *SEDITIONOUS COUP ATTEMPT
 *CORRUPT, SELF SERVING CONMAN [...] (@Ink8Judy)

¹ Irrelevant (e.g., links, twitter handles) or inappropriate (e.g., swear words) text was omitted. Omitted text passages are by marked by a bracket ellipsis “[...]”.

Twitter users seemingly discussed the *results* of the elections in the United States. According to *ChatGPT 3.5*, the word *reality* suggests that Twitter users debated the facts and misinformation surrounding the U.S. Elections (see Appendix D). In general, Twitter users seemed to express a critical tone when discussing this topic. For one, the word *delusion* may suggest that Twitter users accused Donald Trump of being delusional. However, since *grandeur* was also among the 200 keywords of this topic, Twitter users may accused Donald Trump of having ‘*delusions of grandeur*’ (see Figure 6). Furthermore, Twitter users seemingly described Donald Trump as a *malignant* narcissist. Similarly, according to ChatGPT, the use of the word *fascist* suggests that some users perceived Donald Trump’s behaviour or rhetoric as undemocratic, authoritarian, or oppressive (see Appendix D). Twitter users also seemed to discuss the *criminal* charges brought against Donald Trump since his time in office or accused him of engaging in illegal behaviour.

This topic interpretation of the top ten keywords is supported by other keywords representing the topic (see Figure 6). For one, additional keywords, such as *interference*, *vote*, *potus* (i.e., President of the United States), *candidate*, *biden* (U.S. President Joe Biden), and *democracy* strengthen the idea that this topic is about U.S. Elections. Furthermore, the mention of the Twitter handles *@realDonaldTrump* and *@donaldtrumpjr* (the Twitter handle of Donald Trump’s son), as well as the mention of Donald Trump’s campaign slogan ‘Make America Great Again (*maga*)’ support the notion that Twitter users predominately discussed Donald Trump in the context of U.S. Elections. Moreover, the additional the keywords *law*, *indictment*, *violence*, *corrupt*, and *crime* further highlight that Twitter users discussed Donald Trump’s alleged *criminal* behaviour (see Figure 5). The analysis of the tweets supported the interpretation of the keywords, as the tweets mostly centred around Donald Trump’s role in elections, with one tweet with a loading of 0.9666 saying:

[...] Trump’s a self-absorbed narcissist with delusions of grandeur. In reality he only became POTUS due to his pal Putin’s interference in our 2016 election. In typical Fascist fashion; Trump cried foul over the 2020 election results to deflect away from his criminality [...] (@VenusianAndroid)

Narcissistic Abuse & Relationships

The final topic was labelled ‘Narcissistic Abuse & Relationships’ and consisted of the following keywords: *love*, *people*, *life*, *feel*, *use*, *woman*, *man*, *abuse*, and *narcissistic*. Narcissistic Abuse & Relationships was the most frequently discussed topic, with 28.56% (n = 8627) of all tweets discussing this topic. Based on the words *narcissistic* and *abuse*, Twitter seemed to discuss their experiences with psychological, emotional or physical abuse by

One Twitter user wrote in a tweet with a 0.9435 topic loading:

[...] My ex-wife is a narcissist, I lived a nightmare with her for many years. I'm divorced since, very difficult to have confidence after that [...]. (@MC_Tagada94566)

Sentiment Analysis

The mean compound sentiment score was -0.1112. Based on the interpretation guidelines by Hutto (n.d.) for a VADER-based sentiment analysis, a compound score of ≤ -0.05 is scored as a negative sentiment. That is, tweets mentioning the term 'narcissist' expressed on average a negative sentiment. More specifically, nearly half of the tweets (47.53%; $n = 14,356$) expressed a negative sentiment. The tweet that expressed the most negative sentiment had a compound score of -0.9946 and was stated as follows:

HOW THE USA [...] UGLY WAR CRIMINAL, DICK PRICK CHENEY
SOLD THE ILLEGAL WAR TO US AND THE WORLD [...]: A Liar, A Narcissist, A
Minuplater, A war Criminal, A Murderer, A psychopath, including BUSH AND
BLAIR: I will Never Forgive them for Murdering over one Million Iraqi [...] children
[...] (@Shabskk)

Moreover, only 19.31% of the tweets ($n = 5833$) had a neutral compound sentiment score (i.e. a compound score between > -0.05 and < 0.05). About a third of all tweets (33.16%, $n = 10,016$) had a positive compound sentiment score (i.e. a compound score of ≥ 0.05). The most extreme positive compound sentiment score of a tweet was 0.9870 and stated as follows:

#KangDaniel "You might think I'm a narcissist, but I believe I've done well. I love who I am now, and I am very proud of my work. And I'm confident about the things that I am going to accomplish in the future, so I'm having happy days right now" [...]
(ChompooOK)

Sentiment Analysis Per Topic

A one-way ANOVA revealed a significant overall effect of topic on compound sentiment score [$F(4, 30200) = 97.47, p < .001$]. This indicates there was a statistically significant difference in mean compound sentiment scores between topics. As the mean compound sentiment score across all topics was ≤ -0.05 , Twitter users expressed on average a negative sentiment across all topics (see Table 3). Twitter users expressed the most negative sentiment in the topic surrounding Donald Trump & U.S. Elections ($M = -0.2025, SD = 0.0069$), while Narcissistic Abuse & Relationships had the least negative mean compound sentiment score ($M = -0.0602, SD = 0.0058$). Post Hoc LSD comparisons indicated that the mean compound sentiment scores were significantly different between all topics, except for

between Antisocial & Pathological Traits ($M = -0.1231$, $SD = 0.0082$) and Donald Trump & U.S. Elections with an insignificant mean difference of -0.0096 .

Table 3

Descriptive Statistics for Compound Sentiment Score By Topic

Topic	n ^a	Compound Sentiment Score			
		M	SD	Min.	Max.
Antisocial & Pathological Traits	3971	-0.1231	0.0082	-0.9931	0.9803
Politics	7396	-0.1328	0.0061	-0.9924	0.9860
Donald Trump & U.S. Elections	6260	-0.2025	0.0069	-0.9946	0.9760
Narcissistic Abuse & Relationships	8627	-0.0602	0.0058	-0.9923	0.9870

^an = number of Tweets

Discussion

The current paper was the first study to systematically explore the use of the term ‘narcissist’ on social media. Using a text mining approach, the current study analysed 30,205 tweets posted between March 19th and 26th, 2023. After pre-processing, the ten most frequent words in the corpus were: *trump*, *people*, *lie*, “cross mark” emoji, *man*, *malignant*, *love*, *liar*, *life* and *woman*. Topic modelling revealed that the most prevalent topic was Narcissistic Abuse & Relationships, followed by Politics and then Donald Trump & U.S. Elections. The least discussed topic was Antisocial & Pathological Traits. The sentiment analysis revealed that Twitter users expressed on average a negative sentiment when they used the term ‘narcissist’ in a tweet, with nearly half of the tweets expressing a negative sentiment. The sentiments expressed across all topics were also on average negative, with Twitter users expressing the most negative sentiment when they tweeted about Donald Trump & U.S. Elections. Overall, the findings suggest that Twitter users used the term narcissist in personal contexts (i.e., abuse within relationships), and political discussions (e.g., Donald Trump, Elections), but not in clinical or psychiatric context.

Antisocial & Pathological Traits

Overall, the topic ‘Antisocial and Pathological Traits’ indicated that Twitter users who used the term ‘narcissist’ seemed to associate allegedly narcissistic individuals with a wide

array of other antisocial and pathological traits, such as lying, self-centredness, sociopathy, and corruption. Notably, some of the discussed traits do indeed overlap with NPD, which is essentially a disorder that is characterized by a self-centred personality style. More specifically, the finding that narcissists are often associated with being self-serving or self-centred may be related to the diagnostic criteria of having a grandiose sense of self-importance (American Psychiatric Association, 2013; Skodol et al., 2014). Similarly, the use of the word corrupt may be related to the diagnostic criteria that people with NPD are often exploitative and take advantage of others for personal gain (Skodol et al., 2014). This suggests that the understanding on Twitter of what the term ‘narcissist’ means might be to some extent clinically accurate.

On the other hand, some of the traits Twitter users associated with alleged narcissists are not features of NPD and may be better explained by other mental disorders. For instance, the word sociopath is a mostly outdated term for people with an antisocial personality disorder (ASPD), which is characterized by a disregard for and violation of other people’s rights, lack of empathy and remorse, and manipulative behaviour (DeLisi et al., 2019; Walsh & Wu, 2008). And while individuals with ASPD and NPD share certain characteristics (e.g., exploitative nature and lack of empathy), NPD and ASPD are different mental disorders with divergent symptoms (American Psychiatric Association, 2013). For example, people with NPD often do not display impulsivity or aggression and usually do not have a history of criminal behaviour like individuals with ASPD do (American Psychiatric Association, 2013).

Relatedly, while there is a correlation between a narcissistic personality and increased lying (Elaad et al., 2022; O’Reilly & Doerr, 2020), lying is not a symptom of NPD, while deception is, in fact, a symptom of ASPD (American Psychiatric Association, 2013). Lastly, the word *malignant* was one of the 200 keywords representing this topic and, moreover, the sixth most used word in the entire corpus. These frequent references to the concept of malignant narcissism further support the idea that the understanding of narcissism on social media is not entirely accurate. Malignant narcissism is a proposed subtype of NPD, which is characterized by a mixture of narcissism, antisocial behaviour, sadism, and paranoia (Faucher et al., 2022; Golder-Vukov et al., 2010, Kernberg, 2014; Lenzenweger, 2018; Russ et al., 2008). While being a popular construct, malignant narcissism is not clinically validated (Villines, 2024).

Narcissistic Abuse & Relationships

The topic ‘Narcissistic Abuse & Relationships’ indicated that Twitter users shared their experiences with psychological, emotional and/or physical abuse by allegedly narcissistic

individuals. While the initial interpretation of the top ten keywords suggested that this topic centred only on romantic relationships, a closer look at all keywords and at tweets with a high topic loading revealed that this topic was also about abuse within families and among friends. The tweets showed that Twitter users discussed specific forms of psychological abuse (e.g., gaslighting, love-bombing, and trauma bonding) but also referred to physical violence (e.g., “#domesticviolence”), which they discussed under the umbrella term ‘narcissistic abuse’ while frequently referring to the alleged offenders as narcissists.

Unfortunately, the empirical research on the concept of narcissistic abuse is sparse (Howard, 2019; Howard, 2022; Howard & Adan, 2022). However, narcissistic abuse is described as “a form of abuse inflicted by individuals exhibiting severe narcissistic traits and associated patterns of abusive behaviour” (Howard & Adan, 2022, p.1). It is said to differ from other forms of abuse, such as intimate partner violence, as the perpetrator uses complex deceptive tactics, and the abuse can both be direct and indirect (Howard & Adan, 2022). This form of abuse is said to consist of characteristic behaviour, such as love-bombing, pathological lying, and gaslighting (Howard, 2019). Love-bombing is described as a manipulation tactic whereas the perpetrator “bombs” the victim with excessive displays of affection and romantic gestures (e.g., frequent phone calls, premature professing one’s love, flattery (Howard, 2019; Strutzenberg et al., 2016). Gaslighting is a type of psychological abuse where the perpetrator distorts facts to make their victims doubt their own perception (e.g., telling the victim that they forgot to turn off the oven, while in reality they did) (Sweet, 2019). Notably, both gaslighting and love bombing have been specifically mentioned in tweets about Narcissistic Abuse & Relationships. This is not surprising, as both terms have received public attention in recent years, especially among social media users (Howard & Adan, 2022; Strutzenberg et al., 2016)

Mental health care professionals have been warning about the increased use of terms like narcissistic abuse or gaslighting on social media arguing that an overuse of these terms may lead to a weaponization of clinical terms, or it may wrongly pathologize normal behaviour (e.g., conflicts in relationships) (Berman, n.d.; Dodgson & Colombo, 2023; Haupt, 2023; Showfety, 2022; Trujillo, 2023). Furthermore, while using terms like narcissistic abuse may help raise awareness, it could also diminish the seriousness of abuse (Dodgson; 2023; Trujillo, 2023). If every ex-partner is a narcissist and every disagreement is narcissistic abuse, these terms may become meaningless. Academic papers have similarly warned of a misuse of terms like gaslighting, arguing that it is important to differentiate “between a personality disorder as a medical condition, pathological or subclinical narcissism and simple ‘mean’

behaviour” (Freestone et al., 2022, p.55). Furthermore, Freestone et al. (2022) state that a narcissistic personality structure of diagnosis of NPD cannot be inferred without understanding the intention behind the abusive behaviour. On the other hand, Howard (2019) states that victims of narcissistic abuse are often not aware of what happened to them until they stumble upon information about narcissistic abuse on social media or online communities accidentally, arguing that information on social media can be breakthrough for many victims. More specifically, a study by Howard & Adan (2022) found that memes on social media about narcissistic abuse facilitate help-seeking discussions, psychosocial support, and emotional expression.

Politics and Donald Trump

Another topic that emerged in tweets that mentioned the term ‘narcissist’ was ‘Politics’. In the context of politics, Donald Trump was an especially prominent theme. A key finding was that *trump* was the most frequently mentioned word in the corpus. Furthermore, in topic modelling, ‘Donald Trump & U.S. Elections’ emerged as a topic. The topic had an overall critical tone, with Twitter users seemingly describing Donald Trump as a narcissist, fascist, delusional and criminal. However, as the frequent use of the “cross mark” emoji indicates, at least part of the discussion on Twitter may have been supportive of Donald Trump, as the “cross mark” emoji has been increasingly used by social media users to express their support for Donald Trump or the alt-right more generally (Rawlings, 2024). Combined the topics ‘Politics’ and ‘Donald Trump & U.S. Elections’ made up 45.22% of all analysed tweets, suggesting that discussion about politics and political actors are a significant topic in tweets that mention the term narcissists.

The finding that Donald Trump is a prominent theme in tweets that mention the term ‘narcissist’ is not surprising. Donald Trump has often been described both in the news media and on social media as a narcissist despite there being no official reports of Donald Trump being diagnosed with NPD (Frances, 2017; Longman, 2024; Pengelly, 2023; Truong, 2017). Notably, attempting to diagnose Donald Trump with NPD is not only limited to lay people. In fact, the American Psychiatric Association issued a warning to its members and other psychiatrists to stop trying to diagnose Donald Trump because it is considered unethical to diagnose someone without conducting a proper evaluation or official treatment, which is referred to as the Goldwater Rule (Frances, 2017; Oquendo, 2016; Yakeley, 2018). Yakeley (2018) argued that it is the uncertainty around the definition of NPD that makes this disorder especially likely to be attributed to public figures, who are usually labelled as having a NPD based on a single trait (Yakeley, 2018).

Implications

With the increasing use of social media, investigating how psychiatric terms are used online may provide important information about the misuse of psychiatric terms, which may have implications for the stigmatization of mental disorders (Alvarez-Mon et al., 2019; Delanys et al., 2022). While it was outside this study's scope to investigate if the term 'narcissist' is misused or even stigmatized on social media, the current paper's findings may provide some preliminary indications about the alleged misuse of the term 'narcissist' on social media. For one, the finding that the term was used over 30,000 times over a period of seven days suggests that the term 'narcissist' was colloquially used on Twitter. Similarly, topic modelling revealed that the term narcissist was mostly used in either negatively connotated contexts (e.g. Narcissistic Abuse), in other non-clinical contexts (e.g., Politics, Donald Trump), but not in a clinical or psychiatric context.

Second, the current study found that the term 'narcissist' was often used together with terms referring to other antisocial and pathological traits, such as *sociopath*. This suggests that Twitter users may associate 'narcissist' with a wide array of antisocial or pathological traits, which are, however, better explained by other mental disorders, such as ASPD. The finding that social media users sometimes misuse a psychiatric term by mistaking it for a different mental disorder is not new. For instance, Reavley & Pilkington (2014) found that the majority of the tweets (52%) that expressed a stigmatizing attitude towards schizophrenia were inaccurate beliefs about schizophrenia being dissociative identity disorder. If the term narcissist becomes intertwined with characteristics of other mental disorders, this may lead to an increased stigmatization of people with NPD, as 'narcissist' may become an umbrella term for individuals exhibiting a wide array of antisocial and pathological traits, even though these traits are not (exclusive) features of NPD.

Third, the current study found that the term narcissist is often used in tweets about politics and especially Donald Trump, with Twitter users often referring to Trump as a narcissist, even though Donald Trump has never been officially diagnosed with NPD. This suggests that the term 'narcissist' may be misused for name-calling on Twitter. Lastly, the present study's finding that narcissistic abuse within relationships was the most prevalent topic may imply that the use of the term 'narcissist' on social media is strongly associated with abusive behaviour. In fact, the National Domestic Violence Hotline (2022) has reported that many of their clients refer to their abusive partners as narcissists. However, labelling every abuser as a narcissist may pose the risk that narcissism becomes a popular explanation for abusive behaviour, while the perpetrator may be wrongfully labelled as a narcissist, or the

actual narcissistic traits of the perpetrator are not causal factors for the abuse (National Domestic Violence Hotline, 2022).

As stigmatization can have detrimental consequences, such as feelings of shame and an increased reluctance to seek treatment (Da Silva et al., 2020; Sickel et al., 2014), future research may investigate whether the (mis)use of the term ‘narcissist’ on social media has an effect the stigmatization of people with NPD. For example, Battaglia et al. (2022) found that the exposure to Instagram posts about stories of violent offending by individuals with schizophrenia led to an increase in negative attitudes, intended social-distancing behaviours towards and perception of dangerousness of individuals with schizophrenia. Some of these effects could be mitigated by providing the participants with clinical information (Battaglia et al., 2022). Future research may therefore investigate if negatively connotated tweets about narcissists (e.g., tweets about so-called narcissistic abuse) may lead to an increase in negative attitudes or change in behaviour towards people with NPD and whether clinical information about NPD can mitigate these effects.

Strengths and Limitations

Previous studies on the use of other psychiatric terms on social media have used qualitative methods of analysis (Athanasopoulou & Sakellari, 2016; Delanys et al., 2022; Joseph et al., 2015; Lachmar et al., 2017; Reavley & Pilkington, 2014). These methods, however, are not suitable to analyse the ever-growing number of tweets on Twitter (i.e., big data) due to the manual work they require (Abbé et al., 2015; Mohanty, 2015). To overcome this, the current study used computer-assisted text mining techniques, which enabled the analysis of all tweets that mention the term ‘narcissist’ over a period of seven days. Notably, however, the current paper also relied on qualitative techniques, such as the manual analysis of the 100 tweets with the highest topic loading. By using this combination of computer-assisted techniques and traditional qualitative forms of analysis, the current study was able to analyse a large dataset, while stilling obtaining a detailed understanding of the topics.

However, the current study also had some limitations. Due to some technical difficulties caused by a lack of computational power, some design choices had to be made that limited the scope of the research. For instance, the number of keywords were restricted to only the term ‘narcissist’. However, it would have been interesting to also include the related words ‘narcissism’ and ‘narcissistic’. Delanys et al. (2022) found that tweets that mentioned the term ‘schizophrenia’ mainly had a mixed or neutral polarity, while tweets that mention the terms ‘schizo’ or ‘schizophrenic’ expressed on average a negative sentiment. Future research may investigate if tweets that mention the term ‘narcissist’ also tend to express a more

negative sentiment than tweets about ‘narcissism’, assuming that the word ‘narcissist’ is more often used for name-calling. Second, to further limit the corpus, tweets were only collected over a period of seven days. However, given that both ‘Politics’ and ‘Donald Trump & U.S. Elections’ are topics dealing with current events, the content of these topics can be expected to change rapidly depending on the development of current news. Thus, it would have been interesting to investigate the development of the use of the term ‘narcissist’ over time.

Another limitation was the choice of topic modelling method. The present study used LDA topic modelling, as it was the most suitable of the three-topic modelling techniques integrated into *Orange*. However, there are more recent and more sophisticated topic models available, which have not yet been adapted by *Orange*. One example is transformer-based topic modelling techniques, such as BERTopic (Grootendorst, 2022). LDA views a document as a bag-of-words representation and mixture of latent topics, which disregards the semantic relationship among words, which may lead to a misrepresentation of the corpus. In contrast, BERTopic uses text embedding techniques, which enable BERTopic to take context into account, leading to a more accurate topic representation than LDA (Egger & Yu, 2022; Grootendorst, 2022). Future research could therefore investigate the use of the term ‘narcissist’ on social media more comprehensively by using transformer-based text mining techniques. For instance, Moriceau et al. (2022) used FlauBERT, the French version of the BERT language model, to automatically detect how psychiatric terms are used on social media (i.e., medical use, misuse, or irrelevant use). Therefore, future research may use BERT to automatically detect to what extent the term narcissist is misused or used in a psychiatric context.

Conclusion

The results of the current paper suggest that Twitter users frequently used the term ‘narcissist’ in personal contexts (e.g., abuse within relationships), and in political discussions (e.g., Donald Trump and elections), but not in a clinical or psychiatric context. Furthermore, the overall negative sentiment of the tweets implies that the term ‘narcissist’ was mostly used in negatively connotated contexts. Likewise, in the current study the term ‘narcissist’ was often used together with terms referring to other antisocial and pathological traits, used for name-calling individuals, such as Donald Trump, and associated with abusive behaviour. This colloquial and negative use of the term ‘narcissist’ on Twitter may lead to increased stereotypes of people with NPD, potentially contributing to the stigmatization of the disorder. Future research may investigate the (mis)use of the term ‘narcissist’ on social media further and the effect this has on people with NPD.

References

- Abbe, A., Grouin, C., Zweigenbaum, P., & Falissard, B. (2015). Text mining applications in psychiatry: a systematic literature review. *International Journal of Methods in Psychiatric Research*, 25(2), 86–100. <https://doi.org/10.1002/mpr.1481>
- Agrawal, R. (2022, August 5). *Must Known Techniques for text preprocessing in NLP*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2021/06/must-known-techniques-for-text-preprocessing-in-nlp/>
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.).
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- Alizadeh, K. (2022, September 16). *Limitations of Twitter Data - towards data science*. Medium. <https://towardsdatascience.com/limitations-of-twitter-data-94954850cacf>
- Alshani, F., Apon, A., Herzog, A., Safro, I., & Sybrandt, J. (2020). Accelerating Text Mining Using Domain-Specific Stop Word Lists. *2020 IEEE International Conference on Big Data (Big Data)*. <https://doi.org/10.1109/bigdata50022.2020.9378226>
- Alvarez-Mon, M. A., Llaveró-Valero, M., Sánchez-Bayona, R., Pereira-Sánchez, V., Vallejo-Valdivielso, M., Monserrat, J., Lahera, G., Del Barco, A. A., & Alvarez-Mon, M. (2019). Areas of interest and stigmatic attitudes of the general public in five relevant medical conditions: Thematic and quantitative analysis using Twitter. *Journal of Medical Internet Research*, 21(5), e14110. <https://doi.org/10.2196/14110>
- Asmussen, C. B., & Møller, C. (2019). Smart literature review: a practical topic modelling approach to exploratory literature review. *Journal of Big Data*, 6(1). <https://doi.org/10.1186/s40537-019-0255-7>
- Athanasopoulou, C., & Sakellari, E. (2016). “Schizophrenia” on Twitter: Content analysis of Greek language tweets. *PubMed*, 226, 271–274. <https://pubmed.ncbi.nlm.nih.gov/27350523>
- Aydin, A. (2023, October 5). *Text Preprocessing Techniques for NLP*. Medium. <https://ayselaydin.medium.com/1-text-preprocessing-techniques-for-nlp-37544483c007#:~:text=So%20text%20preprocessing%20is%20a,to%20better%20results%20and%20insights.>
- Berman, L. (n.d.). *Why we shouldn't be so quick to label people 'Narcissists.'* <https://drlauraberman.com/why-we-shouldnt-be-so-quick-to-label-people-narcissists/>

- Bhoi, N. (2022, December 14). *Stemming vs Lemmatization in NLP*. Medium.
<https://nirajbhoi.medium.com/stemming-vs-lemmatization-in-nlp-efc280d4e845>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022. <https://doi.org/10.5555/944919.944937>
- Chitrapriya, N., Islam, M. R., Roy, M., & Pradhan, S. (2017). A study on different normalization approaches of word. In *Lecture notes in electrical engineering* (pp. 239–251). https://doi.org/10.1007/978-981-10-4765-7_25
- Cho, H. (2019). Topic modeling. *Osong Public Health and Research Perspectives*, 10(3), 115–116. <https://doi.org/10.24171/j.phrp.2019.10.3.01>
- Da Silva, A. G., Baldaçara, L., Cavalcante, D. A., Fasanella, N. A., & Palha, A. P. (2020). The impact of mental illness stigma on psychiatric emergencies. *Frontiers in Psychiatry*, 11. <https://doi.org/10.3389/fpsy.2020.00573>
- Delanys, S., Benamara, F., Moriceau, V., Olivier, F., & Mothe, J. (2022). Psychiatry on Twitter: Content analysis of the use of psychiatric terms in French. *JMIR Formative Research*, 6(2), e18539. <https://doi.org/10.2196/18539>
- DeLisi, M., Drury, A. J., & Elbert, M. J. (2019). The etiology of antisocial personality disorder: The differential roles of adverse childhood experiences and childhood psychopathology. *Comprehensive Psychiatry*, 92, 1–6.
<https://doi.org/10.1016/j.comppsy.2019.04.001>
- Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., Starič, A., Štajdohar, M., Umek, L., Žagar, L., Žbontar, J., Žitnik, M., & Zupan, B. (2013). Orange: data mining toolbox in python. *Journal of Machine Learning Research*, 14(1), 2349–2353. <http://eprints.fri.uni-lj.si/2267/1/2013%2DDemsar%2DOrange%2DJMLR.pdf>
- Dhawan, N., Kunik, M. E., Oldham, J., & Coverdale, J. (2010). Prevalence and treatment of narcissistic personality disorder in the community: a systematic review. *Comprehensive Psychiatry*, 51(4), 333–339.
<https://doi.org/10.1016/j.comppsy.2009.09.003>
- Dieckmann, E., & Behary, W. (2015). Schematherapie: Ein Ansatz zur Behandlung narzisstischer Persönlichkeitsstörungen. *Fortschritte Der Neurologie · Psychiatrie*, 83(08), 463–478. <https://doi.org/10.1055/s-0035-1553484>
- Dodgson, L., & Colombo, C. (2023, August 22). *Influencers are using psychological terms like “narcissist” and “gaslighting” to fuel drama and bring in clicks*. Business

- Insider. <https://www.insider.com/narcissist-abuse-gaslighting-trend-trisha-paytas-gabbie-hanna-shane-dawson-2021-7>
- Egger, R., & Yu, J. (2022). A topic modeling comparison between LDA, NMF, Top2VEC, and BERTopic to demystify Twitter posts. *Frontiers in Sociology*, 7. <https://doi.org/10.3389/fsoc.2022.886498>
- Elaad, E. (2022). Deceptive behavior: effects of rational thinking, narcissism, and Self-Assessed lie- and truth related abilities. *SAGE Open*, 12(2), 215824402210850. <https://doi.org/10.1177/21582440221085012>
- Faucher, J., Savard, C., Vachon, D. D., Payant, M., & Gamache, D. (2022). A scoring procedure for malignant narcissism based on personality inventory for DSM-5 facets. *Journal of Personality Assessment*, 104(6), 723–735. <https://doi.org/10.1080/00223891.2021.2019052>
- Frances, A. (2017). Misdiagnosing Donald Trump. *Journal of Mental Health*, 26(5), 394. <https://doi.org/10.1080/09638237.2017.1371845>
- Freestone, M., Osman, M., & Ibrahim, Y. (2020). On the uses and abuses of narcissism as a public health issue. *The British Journal of Psychiatry*, 220(2), 54–57. <https://doi.org/10.1192/bjp.2020.70>
- Freud, S. (1914). On narcissism. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud* (vol 14) (ed J Strachey): 67–102. Hogarth Press.
- Gabbard, G. O. (2022). Narcissism and suicide risk. *Annals of General Psychiatry*, 21(1). <https://doi.org/10.1186/s12991-022-00380-8>
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Georgieva-Trifonova, T., & Dechev, M. (2021). Applying text mining methods to extracting information from news articles. *IOP Conference Series Materials Science and Engineering*, 1031(1), 012054. <https://doi.org/10.1088/1757-899x/1031/1/012054>
- Gillis, A. S., & Barney, N. (2024, June 5). *What Is Sentiment Analysis?* Business Analytics. <https://www.techtarget.com/searchbusinessanalytics/definition/opinion-mining-sentiment-mining>
- Goldner-Vukov, M., & Moore, L. J. (2010). Malignant Narcissism: from fairy tales to harsh reality. *PubMed*, 22(3), 392–405. <https://pubmed.ncbi.nlm.nih.gov/20856182>
- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2203.05794>

- Haupt, A. (2023, March 15). *Gaslighting, narcissist, and more psychology terms you're misusing*. TIME. <https://time.com/6262891/psychology-terms-misused-gaslighting-toxic-narcissist/>
- Hota, H., Sharma, D. K., & Verma, N. (2021). Lexicon-based sentiment analysis using Twitter data. In *Elsevier eBooks* (pp. 275–295). <https://doi.org/10.1016/b978-0-12-824536-1.00015-0>
- Howard, V. (2019). Recognising narcissistic abuse and the implications for mental health nursing practice. *Issues in Mental Health Nursing*, 40(8), 644–654. <https://doi.org/10.1080/01612840.2019.1590485>
- Howard, V. (2022). (Gas)lighting Their Way to Coercion and Violation in Narcissistic Abuse. *Journal of Autoethnography*, 3(1), 84–102. <https://doi.org/10.1525/joae.2022.3.1.84>
- Howard, V., & Adan, A. (2022). “The end justifies the memes”: A feminist relational discourse analysis of the role of macro memes in facilitating supportive discussions for victim-survivors of narcissistic abuse. *Cyberpsychology Journal of Psychosocial Research on Cyberspace*, 16(4). <https://doi.org/10.5817/cp2022-4-10>
- Hutto, C. J. (n.d.). *cjhutto/vaderSentiment: VADER Sentiment Analysis. VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media, and works well on texts from other domains*. GitHub. <https://github.com/cjhutto/vaderSentiment>
- Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. <https://doi.org/10.1609/icwsm.v8i1.14550>
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2018). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11), 15169–15211. <https://doi.org/10.1007/s11042-018-6894-4>
- Joseph, A. J., Tandon, N., Yang, L. H., Duckworth, K., Torous, J., Seidman, L. J., & Keshavan, M. S. (2015). #Schizophrenia: Use and misuse on Twitter. *Schizophrenia Research*, 165(2–3), 111–115. <https://doi.org/10.1016/j.schres.2015.04.009>
- Kacel, E. L., Ennis, N., & Pereira, D. B. (2017). Narcissistic Personality Disorder in Clinical Health Psychology Practice: Case studies of Comorbid Psychological Distress and Life-Limiting Illness. *Behavioral Medicine*, 43(3), 156–164. <https://doi.org/10.1080/08964289.2017.1301875>

- Kaity, M., & Balakrishnan, V. (2020). Sentiment lexicons and non-English languages: a survey. *Knowledge and Information Systems*, 62(12), 4445–4480.
<https://doi.org/10.1007/s10115-020-01497-6>
- Kapadia, S. (2019, August 19). Evaluate topic models: Latent Dirichlet Allocation (LDA). *Medium*. <https://towardsdatascience.com/evaluate-topic-model-in-python-latent-dirichlet-allocation-lda-7d57484bb5d0>
- Karami, A., Lundy, M., Webb, F., & Dwivedi, Y. K. (2020). Twitter and Research: A Systematic Literature Review through text mining. *IEEE Access*, 8, 67698–67717.
<https://doi.org/10.1109/access.2020.2983656>
- Kernberg, O. F. (2014). An overview of the treatment of severe narcissistic pathology. *The International Journal of Psychoanalysis*, 95(5), 865–888.
<https://doi.org/10.1111/1745-8315.12204>
- Kherwa, P., & Bansal, P. (2018). Topic Modeling: A Comprehensive review. *ICST Transactions on Scalable Information Systems*, 7(24), 159623.
<https://doi.org/10.4108/eai.13-7-2018.159623>
- Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihók, G., & Hartog, D. N. D. (2017). Text mining in organizational research. *Organizational Research Methods*, 21(3), 733–765.
<https://doi.org/10.1177/1094428117722619>
- Lachmar, E. M., Wittenborn, A. K., Bogen, K. W., & McCauley, H. L. (2017). #MyDepressionLooksLike: Examining public discourse about depression on Twitter. *JMIR Mental Health*, 4(4), e43. <https://doi.org/10.2196/mental.8141>
- Laing, A. (2017). Authors using social media: Layers of identity and the online author community. *Publishing Research Quarterly*, 33(3), 254–267.
<https://doi.org/10.1007/s12109-017-9524-5>
- Lefort, R. (2022, August 17). *How psychiatric terms are constantly misused and hijacked by the media*. *Medium*. <https://medium.com/in-fitness-and-in-health/how-psychiatric-terms-are-constantly-misused-and-abused-by-the-media-6e277da04328>
- Lenzenweger, M. F., Clarkin, J. F., Caligor, E., Cain, N. M., & Kernberg, O. F. (2018). Malignant narcissism in relation to clinical change in borderline personality Disorder: an exploratory study. *Psychopathology*, 51(5), 318–325.
<https://doi.org/10.1159/000492228>
- Levy, K. N., Ellison, W. D., & Reynoso, J. S. (2011). A Historical Review of Narcissism and Narcissistic Personality. In *The handbook of narcissism and narcissistic personality*

- disorder: Theoretical approaches, empirical findings, and treatments* (pp. 2–13). Wiley. <https://doi.org/10.1002/9781118093108.ch1>
- Li, A., Jiao, D., Liu, X., & Zhu, T. (2020). A comparison of the psycholinguistic styles of Schizophrenia-Related stigma and Depression-Related stigma on social media: Content analysis. *Journal of Medical Internet Research*, *22*(4), e16470. <https://doi.org/10.2196/16470>
- Longman, M. (2024, September 4). *Is “Narcissist” being overused? Experts weigh in.* Popsugar. <https://www.popsugar.com/relationships/narcissist-label-overused-49358385>
- Mimno, D., Wallach, H. M., Princeton University, Miriam Leenders, E. T., National Institutes of Health, McCallum, A., & University of Massachusetts, Amherst. (2011). Optimizing semantic coherence in topic models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 262–272). <http://dirichlet.net/pdf/mimno11optimizing.pdf>
- Mohanty, H. (2015). Big Data: an Introduction. In *Studies in big data* (pp. 1–28). https://doi.org/10.1007/978-81-322-2494-5_1
- Moriceau, V., Benamara, F., & Boumadane, A. (2022, June 20). *Automatic detection of stigmatizing uses of psychiatric terms on Twitter.* <https://ut3-toulouseinp.hal.science/hal-03707226/>
- Murthy, D., Bowman, S., Gross, A. J., & McGarry, M. (2015). Do we tweet differently from our mobile devices? A study of language differences on Mobile and Web-Based Twitter platforms. *Journal of Communication*, *65*(5), 816–837. <https://doi.org/10.1111/jcom.12176>
- Namugera, F., Wesonga, R., & Jehopio, P. (2019). Text mining and determinants of sentiments: Twitter social media usage by traditional media houses in Uganda. *Computational Social Networks*, *6*(1). <https://doi.org/10.1186/s40649-019-0063-4>
- National Domestic Violence Hotline (2022, April 21). *Narcissism and Abuse.* <https://www.thehotline.org/resources/narcissism-and-abuse/>
- Negara, E. S., Triadi, D., & Andryani, R. (2019). *Topic Modelling Twitter Data with Latent Dirichlet Allocation Method.* <https://doi.org/10.1109/icecos47637.2019.8984523>
- O’Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., & Ananiadou, S. (2015). Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic Reviews*, *4*(1). <https://doi.org/10.1186/2046-4053-4-5>

- Oquendo, M. A. (2016, August 3). The Goldwater Rule: Why breaking it is Unethical and Irresponsible. *APA Blogs*. <https://www.psychiatry.org/news-room/apa-blogs/apa-blog/2016/08/the-goldwater-rule>
- O'Reilly, C. A., & Doerr, B. (2020). Conceit and deceit: Lying, cheating, and stealing among grandiose narcissists. *Personality and Individual Differences, 154*, 109627. <https://doi.org/10.1016/j.paid.2019.109627>
- Oscar, N., Fox, P. A., Croucher, R., Wernick, R., Keune, J., & Hooker, K. (2017). Machine Learning, sentiment analysis, and tweets: An examination of Alzheimer's disease stigma on Twitter. *The Journals of Gerontology Series B, 72*(5), 742–751. <https://doi.org/10.1093/geronb/gbx014>
- Öztürk, N., & Ayvaz, S. (2018). Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis. *Telematics and Informatics, 35*(1), 136–147. <https://doi.org/10.1016/j.tele.2017.10.006>
- Pengelly, M. (2023, December 14). *Trump is a 'populist, authoritarian narcissist', says ex-speaker Paul Ryan*. The Guardian. <https://www.theguardian.com/us-news/2023/dec/14/paul-ryan-trump-not-a-conservative>
- Penney, E. S., McGill, B. C., & Witham, C. (2017). Opinion Piece: Therapist Stigma towards Narcissistic Personality Disorder: Lessons Learnt from Borderline Personality Disorder. *Australian Clinical Psychologist, 3*(1), 1745. <https://acp.scholasticahq.com/article/1745.pdf>
- Qomariyah, S., Iriawan, N., & Fithriasari, K. (2019). Topic modeling Twitter data using Latent Dirichlet Allocation and Latent Semantic Analysis. *AIP Conference Proceedings*. <https://doi.org/10.1063/1.5139825>
- Rank, O. (1911). Ein Beitrag zum Narcissismus. *Jahrbuch für Psychoanalytische und Psychopathologische Forschungen, 3*, 401–426.
- Reavley, N. J., & Pilkington, P. D. (2014). Use of Twitter to monitor attitudes toward depression and schizophrenia: an exploratory study. *PeerJ, 2*, e647. <https://doi.org/10.7717/peerj.647>
- Rijcken, E., Scheepers, F., Zervanou, K., Spruit, M., Mosteiro, P., & Kaymak, U. (2023). *Towards Interpreting Topic Models with ChatGPT*. Paper presented at The 20th World Congress of the International Fuzzy Systems Association, Daegu, Korea, Republic of.
- Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and Open-AI models: A preliminary review. *Future Internet, 15*(6), 192. <https://doi.org/10.3390/fi15060192>

- Russ, E., Shedler, J., Bradley, R., & Westen, D. (2008). Refining the construct of Narcissistic Personality Disorder: diagnostic criteria and subtypes. *American Journal of Psychiatry*, *165*(11), 1473–1481. <https://doi.org/10.1176/appi.ajp.2008.07030376>
- Showfety, S. (2022, February 23). *You're probably using these psychological terms wrong on social media*. Lifehacker. <https://lifelifehacker.com/youre-probably-using-these-psychological-terms-wrong-on-1848582251>
- Sickel, A. E., Seacat, J. D., & Nabors, N. A. (2014). Mental health stigma update: A review of consequences. *Advances in Mental Health*, *12*(3), 202–215. <https://doi.org/10.1080/18374905.2014.11081898>
- Skodol, A. E., Bender, D. S., & Morey, L. C. (2014). Narcissistic personality disorder in DSM-5. *Personality Disorders Theory Research and Treatment*, *5*(4), 422–427. <https://doi.org/10.1037/per0000023>
- Statista. (2023, November 15). *X/Twitter: number of worldwide users 2019-2024*. <https://www.statista.com/statistics/303681/twitter-users-worldwide/>
- Stinson, F. S., Dawson, D. A., Golstein, R. B., Chou, P., Huang, B., Smith, S. M., Ruan, W. J., Pulay, A. J., Saha, T. D., Pickering, R. P., & Grant, B. F. (2008). Prevalence, correlates, disability, and comorbidity of DSM-IV Narcissistic Personality Disorder. *The Journal of Clinical Psychiatry*, *69*(7), 1033–1045. <https://doi.org/10.4088/jcp.v69n0701>
- Strutzenberg, C. C., Wiersma-Mosley, J. D., Jozkowski, K. N., & Becnel, J. N. (2016). Love-bombing: A narcissistic approach to relationship formation. *Discovery, the Student Journal of Dale Bumpers College of Agricultural, Food and Life Sciences*, *18*(1), 81–89. <https://scholarworks.uark.edu/cgi/viewcontent.cgi?article=1013&context=discoverymag>
- Sweet, P. L. (2019). The Sociology of Gaslighting. *American Sociological Review*, *84*(5), 851–875. <https://doi.org/10.1177/0003122419874843>
- Thakur, K., & Kumar, V. (2021). Application of text mining techniques on Scholarly research Articles: Methods and Tools. *New Review of Academic Librarianship*, *28*(3), 279–302. <https://doi.org/10.1080/13614533.2021.1918190>
- Trujillo, S. (2023, November 21). *From Gaslighting to Narcissist: Commonly Misused Pop Psychology Terms (and What They Really Mean)*. Sondermind. <https://www.sondermind.com/resources/articles-and-content/commonly-misused-pop-psychology-terms/>

- Truong, K. (2017, February 1). *This doctor nailed the problem with diagnosing Donald Trump with mental illness*. Refinery29. <https://www.refinery29.com/en-us/2017/01/138691/donald-trump-mental-illness-point-twitter>
- Villines, Z. (2024, January 31). *What to know about malignant narcissism*. MedicalNewsToday. <https://www.medicalnewstoday.com/articles/malignant-narcissist#definition>
- Walsh, A., & Wu, H. (2008). Differentiating antisocial personality disorder, psychopathy, and sociopathy: evolutionary, genetic, neurological, and sociological considerations. *Criminal Justice Studies*, *21*(2), 135–152. <https://doi.org/10.1080/14786010802159814>
- Webb, C. H. (2022, April 20). *Narcissism: considered diagnosis or easy insult?* Psychology Today. <https://www.psychologytoday.com/us/blog/drawing-the-curtains-back/202204/narcissism-considered-diagnosis-or-easy-insult>
- Wiegersma, S., Hidajat, M., Schrieken, B., Veldkamp, B., & Olf, M. (2022). Improving Web-Based treatment intake for multiple mental and substance use disorders by text mining and machine learning: algorithm development and validation. *JMIR Mental Health*, *9*(4), e21111. <https://doi.org/10.2196/21111>
- Wold, H. M., Vikre, L., Gulla, J. A., Özgöbek, Ö., & Su, X. (2016). Twitter Topic Modeling for Breaking News Detection. *Proceedings of the 12th International Conference on Web Information Systems and Technologies (WEBIST 2016)*. <https://doi.org/10.5220/0005801902110218>
- Wongkoblap, A., Vadillo, M. A., & Curcin, V. (2017). Researching Mental Health Disorders in the Era of Social Media: Systematic review. *Journal of Medical Internet Research*, *19*(6), e228. <https://doi.org/10.2196/jmir.7215>
- Yakeley, J. (2018). Current understanding of narcissism and narcissistic personality disorder. *BJPsych Advances*, *24*(5), 305–315. <https://doi.org/10.1192/bja.2018.20>
- Zvornicanin, E., & Martin, E. (2024, March 18). *Evaluate Topic Models: Latent Dirichlet Allocation (LDA)*. Baeldung. <https://www.baeldung.com/cs/topic-modeling-coherence-score>

Appendix A

Manually Added Stopwords List

sound	wasnot	fit	sense	yeah
=	wish	bigger	get	play
course	less	biggest	via	find
get	don't	mind	mind	called
one	better	u	matter	understand
like	deal	say	yes	can't
sign	do	see	no	another
thing	ways	total	nope	year
sorry	,	typical	need	got
bit	nobody	he's	see	must
see	one	that's	i'm	since
better	definitely	higher	take	sure
important	lundy	highest	nothing	narcissist'
dont	day	try	also	thought
lack	take	worse	ever	know
don't	happen	worst	everyone	even
upon	oh	never	anything	way
do	think	would	talk	realize
care	want	plus	old	absolutely
there's	he'll	i	could	narcissist
what	%	make	much	#narcissist
what's	go	doesnot	come	literally
word	really	away	actually	put
type	look	let	exactly	might
watch	b	she'	last	give
etc	please	every	whole	turn
10	fit	throw	question	seem
wait	may	well	full	read
3	he'	opinion	person	everything
bit	twiste	mf	tho	honestly
little	yet	yes	issue	i've
either	lot	im	w	point
I	little	cannot	clearly	without
l	else	truly	two	with
time	probably	call	finally	around
always	cause	ur	complete	become
isnot	first	still	fuck	already
tell	though	someone	part	something
back	maybe	continue	name	its

many	long	anyone	hold	best
ask	more	after	didn	lol
you're	m	doesn	behind	check
other	bc	am	absolute	lmao
pal	overwrite	same	least	mean
hi	should	again	1	po
anyway	could	such	fill	true
somehow	would	you're	wow	pretty
anyhow	guess	too	almost	very
re	then	anything	dude	false
flie	far	claim	shit	who's
highly	lack	reason	wonot	i'll
except	personally	guy	ye	bro
they're	2	r	i'd	gonna
x	due			

Appendix B

Topic Labels Generated by The Researcher and ChatGPT 3.5.

Topic	Researcher's Label	<i>ChatGPT 3.5's</i> Label	Final Topic label
1	Pathological Traits and Behaviour	Deceit and Selfish Traits	Antisocial & Pathological Traits
2	-Unlabeled-	Emotional Reactions and Relationships	-Unlabeled-
3	Politics	Political Figures and Narcissism	Politics
4	Donald Trump Politics	Trump and Election Fraud	Donald Trump & U.S. Elections
5	Narcissistic Abuse	Personal Relationships and Abuse	Narcissistic Abuse & Relationships

Appendix C

Prompt Used to Generate Topic Labels with ChatGPT 3.5.

I am currently writing my master thesis about the use of the term "narcissist" on social media. To achieve this, I am using text mining techniques. One of the techniques I used was topic modelling. Based on the coherence scores, five topics were extracted. The topics included the following keywords:

1. Topic 1: liar, pathological, self, sociopath, corrupt, serve, grifter, love, god
2. Topic 2: 😂, @clownworldld_, 🤖 emoji, contact, alone, worth, "heart emoji", leave, next
3. Topic 3: people, lie, trump, attention, 2020, right, real, man, foul, malignant
4. Topic 4: trump, election, malignant, fascist, result, delusion, 2016, criminal, trump', reality
5. Topic 5: Love, people, life, feel, use, woman, man, abuse, narcissistic

Based on these keywords, how would label these topics in a maximum of four words? ↓

Appendix D

Screenshots of Topic Interpretations Generated by ChatGPT 3.5.

I am currently writing my master thesis about the use of the term "narcissist" on social media. To achieve this, I am using text mining techniques. One of the techniques I used was topic modelling. The first topic was labelled 'Antisocial and Pathological Traits' based on the following keywords: liar, pathological, self, sociopath, corrupt, serve, grifter, love, and god. How would you interpret the words serve, love, and god in in a maximum of five sentence?



In the context of "Antisocial and Pathological Traits," the terms "serve," "love," and "god" illustrate how narcissists manipulate these concepts for their own gain. "Serve" suggests a pretense of helpfulness or altruism, used to gain trust and leverage over others while masking selfish motives. "Love" implies the exploitation of emotional bonds, where expressions of affection are insincere and aimed at controlling or manipulating people. "God" indicates the misuse of religious or spiritual beliefs to present a facade of piety, thereby gaining trust and justifying manipulative behaviors. Together, these terms highlight the deceptive strategies used by narcissists to exploit trust and relationships. ↓

I am currently writing my master thesis about the use of the term "narcissist" on social media. To achieve this, I am using text mining techniques. One of the techniques I used was topic modelling. The first topic was labelled 'politics' based on the following keywords: people, lie, trump, attention, 2020, right, real, man, foul, malignant. How would you interpret the words real, foul and man in a maximum of five sentences?



In the context of the topic labeled "politics," the words "real," "foul," and "man" reflect social media discourse around authenticity, ethical behavior, and male political figures.

1. **Real:** This term suggests debates about the authenticity or truthfulness of political statements, actions, and personas, often questioning the sincerity and legitimacy of political figures.
2. **Foul:** "Foul" implies a perception of unethical, corrupt, or morally reprehensible behavior, indicating strong disapproval and condemnation of certain actions or rhetoric in the political sphere.
3. **Man:** This term likely refers to prominent male political figures, particularly Donald Trump, emphasizing discussions around their conduct, influence, and role in politics.

Together, these keywords highlight critical discussions about the genuineness, ethical conduct, and impact of male politicians in the political landscape.

I am currently writing my master thesis about the use of the term "narcissist" on social media. To achieve this, I am using text mining techniques. One of the techniques I used was topic modelling. The topic was labelled 'Donald Trump and U.S. Elections' based on the following keywords: trump, election, malignant, fascist, result, delusion, 2016, criminal, trump' and reality How would you interpret the word fascist in a maximum of five sentences?



In the context of the topic labeled "Donald Trump and U.S. Elections," the word "fascist" likely reflects a strong, critical stance taken by social media users against Trump and his political actions or ideology. It suggests that some users perceive Trump's behavior, policies, or rhetoric as authoritarian, oppressive, or undemocratic, drawing parallels with fascist regimes. The use of "fascist" indicates a deep level of disapproval and concern about the potential erosion of democratic norms and values. This term highlights the intensity of the political and ideological conflict surrounding Trump and his influence on the U.S. political landscape. Overall, it underscores the polarized nature of discussions related to Trump and the U.S. elections.



I am currently writing my master thesis about the use of the term "narcissist" on social media. To achieve this, I am using text mining techniques. One of the techniques I used was topic modelling. The fourth topic was labelled 'Donald Trump & U.S. Elections' based on the following keywords: trump, election, malignant, fascist, result, delusion, 2016, criminal, trump' and reality. How would you interpret the word reality?



In the context of the topic labeled "Donald Trump & U.S. Elections," the word "reality" refers to the contested nature of the facts and narratives surrounding Trump's involvement and the election outcomes. It highlights the divide between what different groups perceive as the truth, particularly in terms of the legitimacy of the election results and the events that followed. "Reality" in this sense emphasizes the conflict between accepted facts and perceived delusions or misinformation, reflecting the broader societal struggle to establish a common understanding of truth in a highly polarized political landscape. This term captures the essence of the debates over what is real and what is fabricated in the discourse around the U.S. elections involving Trump.