Exploring and Comparing Climate Change Discussions on Reddit:

Unveiling the Discrete Emotions Expressed by Believers and Deniers

Felix Bonekamp

Faculty of Behavioural, Management and Social sciences (BMS), University of

Twente

Master's Thesis

First supervisor: dr. Alejandro Dominguez Rodriguez

Second supervisor: dr. Anneke Sools

External advisor: dr. Shenja van der Graaf

Abstract

Introduction. Despite the growing evidence of the climate crisis and its increasing presence as a societal issue, many people still doubt anthropogenic climate change. Social media platforms like Reddit function as locations for climate change discussions, in which emotions significantly influence opinions, risk assessments, and decision-making of the involved. This study used text mining to analyse Reddit discussions among climate change believers and deniers, focusing on comparing the emotional expressions of both groups.

Methods. We selected subreddits focusing on climate change belief and denial, using further web scraping and keyword searches to gather relevant posts and comments. Furthermore, Sentiment analysis and emotion detection were performed using VADER and the NRC Word-Emotion Association Lexicon (EmoLex).

Results. Our findings revealed no differences in emotional expression between the two groups. Both groups expressed fear most frequently, with trust and anger as secondary emotions. However, sentiment analysis yielded mixed results. Namely, VADER indicated more positivity overall, especially among believers, while EmoLex suggested similar levels of negativity in both groups. Additionally, deniers showed higher engagement levels than believers.

Discussion. The results highlight the function of cognitive dissonance and social identity in shaping the emotional expressions and behaviours of climate change deniers. The expressed emotions of both groups may, in turn, influence their risk perception and decision-making. The levels of fear expressed by both groups may motivate them to seek reassurance and increase their community engagement

Conclusion. Overall, the study offers insights into the subtle differences and similarities in emotional expression and online interaction between climate change believers and deniers. We identified many similarities in the emotions expressed by both groups, while fear can be highlighted as the most expressed emotion. Future research is needed to gain further insights

into the motivations and topics expressed by both climate change believers and deniers, as well as into complex emotions like sarcasm and hope, which may further influence their decisions.

Exploring and Comparing Climate Change Discussions on Reddit: Unveiling the Discrete Emotions Expressed by Believers and Deniers

The climate crisis is present, and, over the last few years, we have witnessed a noticeable increase in its negative consequences. Hereby, the overall population, especially the inhabitants of developing countries, are prone to suffer from these adverse effects (Huckelba & Van Lange, 2020). For example, greater rainfall and warmer temperatures increase the risk of malaria infection, which affects roughly 50% of the world's population. Moreover, nature is suffering from the effects of climate change, as European forests become increasingly exposed to disturbances such as fires, storms, and droughts (Vacek, et al., 2023). Namely, in European forests, approximately 33.4 billion tons of forest biomass may be severely impacted by those various disturbances (Forzieri, et al., 2021). Yet we have seen a 41% increase in global greenhouse gas emissions since 1990, with the private sector playing an important role in influencing the level of emitted emissions (Coen, et al., 2020; Sørheim, 2021). Therefore, it is crucial to grasp the motives of the population towards climate change, sustainable lifestyles, and environmental policies.

It can be observed that rising stress and anxiety concerning the climate crisis negatively impact the mental health and quality of life of the population (Huckelba & Van Lange, 2020). Consequently, it is an important worry of our current generation. As the young generation will most prominently be affected by the climate crisis in the future, a large majority of young individuals from various countries are becoming increasingly worried about it (Hickman, et al., 2021). This specific worry can better be named climate anxiety, which is defined by the American Psychological Association (2017, p.68) as "a chronic fear of environmental doom". Hereby, it could lead the suffering individual to experience distress in the forms of panic attacks, insomnia, and irritability (APA, 2017). Furthermore, individuals who have experienced the negative consequences of climate change, both indirectly or directly, are prone to suffer

from various mental disorders like depression, anxiety, or post-traumatic stress disorder (Sharpe & Davison, 2021). Moreover, most of the research concerns individuals worried about the negative consequences of climate change, however, less is known about the experienced emotions of those people who are termed 'climate change deniers'. Therefore, it is crucial to better grasp the motives of climate change deniers to foster sustainable lifestyles and agreement concerning environmental policies in society to, finally, counteract climate change and its negative impacts on humankind and planet Earth. Despite an overwhelming amount of evidence of climate change and its consequences, there remains a considerable group of people who are sceptical about climate change, either as a whole or as a result of human impact on the environment (Capstick & Pidgeon, 2014; Treen, et al., 2022).

Climate Change Deniers

Approximately 18% of questioned US citizens reported that they are doubtful and dismissive concerning our society's responsibility for climate change, with similar prevalence rates found in other countries (Beiser-McGrath & Bernauer, 2021; Maibach, et al., 2009). In addition, the results of a European Social Survey indicated that 3-10% of the respondents of western European countries are sceptical about climate change (Poortinga, et al., 2019). However, when accounting for a social desirability bias, often the actual number of climate change deniers in the general population may be higher (Beiser-McGrath & Bernauer, 2021). Climate change has become a political issue, and the majority of deniers are conservatives (McCright & Dunlap, 2011). Additionally, climate deniers are likely to be older, white, male and from the working class (Poortinga, et al., 2011). They frequently attribute their disbelief to a perceived lack of weather changes, natural climate variations, or inconclusive and contradictory scientific evidence (Jones, et al., 2014, as mentioned in Haltinner & Sarathchandra, 2018). However, some deniers still experience a degree of fear regarding environmental tragedies, despite their beliefs that climate change does not exist.

A sizeable number of deniers are concerned about the negative impact of environmental tragedies on the planet, its nature, and its population (Haltinner, et al., 2021). Particularly, they are worried about specific matters such as, for example, pollution or animal extinction. Thereby, it is argued that despite their disbelief in the concept of anthropogenic climate change, deniers are more likely to support pro-environmental policies the greater their experienced worry and anxiety caused by these negative events. In that respect, research argues that people are motivated to deny climate change as a way to cope with feelings of anxiety and uncertainty (Dodds, 2021; Haltinner & Sarathchandra, 2018). Consequently, the deniers mostly engage with information that affirms their views, which causes them to experience lower levels of anxiety, worries, and dread than climate change believers (Haltinner, et al., 2021). In an upcoming section the phenomenon called *echo chambers* (Farrell, 2015) will be explored further, which is a place where people of similar view interact to approve and reinforce their perspectives.

Generally, climate change is frequently discussed on social media platforms such as Reddit or Twitter, as these platforms are easily accessible to the majority of the population and allow their users to freely express their opinions, even if they contradict the general public's view (Hemsley, et al., 2021; Peace, et al., 2018, Villanueva, 2021). Furthermore, deniers tend to engage in these discussions, hereby, frequently doubting the opinion of the general public on climate change, while engaging with other individuals who are like-minded (Haltinner, et al., 2021; Treen, et al., 2022).

Using Social Media Data to Examine Climate Change Discussions

Amongst its typical applications, social media is increasingly being used as a tool to easily engage in discussions concerning activism, political action, or societal issues such as climate change (Chon & Park, 2020; Villanueva, 2021). Researchers extensively examined the contents of climate change discussions present on social media platforms, specifically on

Twitter (Hemsley, et al., 2021; Pearce, et al., 2018; Treen, et al., 2022). Recent research indicates that the individuals partaking in these discussions frequently hold polarized opinions (Pearce, et al., 2018). As most of the recent research on climate change discussions has been carried out using content from Twitter, it is of particular interest to examine discussions on the social media site Reddit, which is still relatively little-studied and has a specific architecture and mechanisms that may encourage discussions among its users (Pearce, et al., 2018; Treen, et al., 2022; Villanueva, 2021).

Reddit

Reddit is one of the largest social media platforms having around 52 million daily active users in 2020, while 303.4 million posts and 2 billion comments were posted in the same year (Reddit, 2020). Unlike Twitter, a follower-based network, Reddit's community is distributed into various *subreddits*, acting as separate forums for specific topics of interest (Proferes, et al., 2021; Treen, et al., 2022; Zapcic, et al., 2023). Hereby, there are over 100,000 active subreddits, indicating a variety of different topic areas (Reddit, 2024). As Reddit offers a significantly higher character limit per post when compared to Twitter, one may argue that this greater word limit may encourage deliberative discussions (Treen, et al., 2022). In general, Reddit users are more engaged in the site's content than other social media users (Kemp, 2019, as cited in Villanueva, 2021). Moreover, Reddit discourages the usage of one's real name as an account name as a measure to ensure the privacy of its users. Consequently, this enables anonymous users to engage in sensitive discussions more comfortably and openly without self-disclosure (Ma, et al., 2016). Reddit's anonymity contributes to increased self-esteem and decreased social anxiety and social desirability among its users (Joinson, 1999). Therefore, one could argue that anonymous users on Reddit are more inclined to share their honest individual opinions and experienced emotions regarding sensitive issues like climate change, while not being discouraged by the aforementioned social desirability bias. Many climate change discussions

are present on the platform (Hemsley, et al., 2021; Treen, et al., 2022). While deliberate discussions about climate change take place, some authors argue, that these discussions may further induce polarisation between the two opposing parties of deniers and believers, contributing to the formation of echo chambers (Falkenberg, et al., 2022; Walter, et al., 2018).

Polarisation and Echo Chambers

While online users can easily access a variety of information and opinions, their freedom of choice may also facilitate selective exposure, consequently, causing people to only seek information and discussions that are in line with their own opinions (Falkenberg, et al, 2022; Walter, et al, 2018). Furthermore, as climate change discussions usually involve politics to some extent, the authors argue that these political discussions often take place in polarised, "ideological silos" (Marchal, 2020). The concept of *echo chambers* "describes the way in which information can be amplified and repeated within enclosed networks of ideologically similar individuals, leading to further entrenchment of ideas and beliefs" (Farrell, 2015, p.719). Consequently, further exploration of novel information which may challenge the already established view is hindered. Moreover, echo chambers on social media platforms are characterised by high engagement, which may result in increased negativity and polarization (Del Vicario, et al., 2016).

When examining climate change Reddit posts on different subreddits, previous research identified no presence of echo chambers, however, the discourse between both parties of deniers and believers could be described as polarized (Treen, et al., 2022). The authors extended the preexisting research on echo chambers on the platform Twitter, on which greater formation of polarisation and echo chambers could be noted (Williams, et al., 2015, as cited in Treen, et al., 2022). The specific architecture of a social media platform may influence the formation of echo chambers. For example, both Twitter and Facebook enforce heavier recommender algorithms compared to Reddit, resulting in less presence of echo chambers on Reddit compared to the

other platforms (Cinelli, et al., 2021; Pariser, 2012, as cited in Efstratiou, et al., 2023). As messages in highly polarised climate change discussions contain more negative sentiments and emotions, the influence of emotions on engagement in these discussions is further explored (Del Vicario, et al., 2016; Williams, et al., 2015, as cited in Villanueva, 2021).

Influence of Discrete Emotions on Risk Perception

Discrete Emotions and Their Prevalence in Climate Change Discussions

Emotional responses play a crucial role when analysing climate change discussions occurring on social media platforms because they influence the individual's posting behaviour and the individual's engagement with social media posts. For example, anger stimulates an individual's tendency to engage in risk-seeking behaviour, (Heiss, 2020, as cited in Villanueva, 2021), and individuals tend to engage more with negative information, reacting strongly to it (Fiske, 1992; Taylor, 1991). Consequently, the researchers propose that the content of the shared message, not just the source of risk information, has a significant influence on the individual's evoked level of trust because of the bias towards negative information, regardless of the source's credibility (Siegrist & Cvetkovich, 2001). Understanding the emotions present in online discussions is crucial for gaining insights into the dynamics of online communities (Laaksonen & Rantasila, 2021).

The expressed emotions significantly varied across different types of subreddits (Villanueva, 2021). Overall, anger was the most prevalent emotion that was expressed in climate change discussions on ideologically neutral and denier subreddits, while the emotions of hope were expressed most frequently in subreddits of climate change believers (Villanueva, 2021). Contrastingly, according to Villanueva (2021), feelings of anxiety were expressed the least while respecting both their overall sample used and also noting the different subreddits examined. These findings align partially with the research of Hemsley, et al. (2021) on climate deniers' tweets, which uncovered that negative emotions were most prevalent, especially anger

and sarcasm. Overall, emotions were present in three-quarters of the deniers' tweets, while about half of the tweets of climate believers included emotions. Contrastingly to the deniers, however, the emotions primarily expressed by the believers were a combination of fear and anger, with low levels of expressed sarcasm, while not analysing the expressed hope.

Appraisal Tendency Framework and Affective intelligence theory

Discrete emotions can have diverse effects on an individual's risk perception and decision-making. According to the Appraisal Tendency Framework (ATF) theory by Lerner and Keltner (2000; 2001), discrete emotions, such as anger or fear, influence an individual's risk perception and decision-making by shaping cognitive biases in appraising future situations. The ATF theory distinguishes emotions based on various cognitive dimensions such as certainty and perceived control. Specifically, anger tends to lead individuals to perceive risks as predictable and controllable, whereas fear causes a sense of unpredictability and lack of control (Drače & Ric, 2012; Smith & Ellsworth, 1985, as cited in Drače & Ric, 2012). Additionally, the Affective Intelligence Theory (AIT) by Marcus, et al. (2000), indicates that emotions are crucial in motivating the behaviour of an individual (Hemsley, et al., 2021). Explicitly, the AIT explores distinct emotions, namely, anxiety, anger, and enthusiasm and, overall, focuses "on the association of emotional reactions to information-seeking and decision-making processes" (Marcus, et al., 2019, p. 110). The theory states that individuals possess two distinct emotional systems, which influence their cognitions, namely a dispositional system and a surveillance system (Marcus, et al., 2011). Namely, anger and enthusiasm activate the disposition system, which leads the individual to refrain from processing novel information. Contrastingly, when a person experiences anxiety, the surveillance system is activated, causing the individual to seek novel information which challenges their preexisting beliefs. Additionally, MacKuen et al. (2010) argued that experiences of hope also activate the surveillance system. In summary, both ATF and AIT highlight the significance of experienced

emotions in shaping an individual's decision-making processes and risk perception. Specifically, fear and hope induce compromise, as the individual seeks new information because the risk is assessed as beyond control. In contrast, anger causes the individual to rely on their preexisting risk evaluation because they perceive the risk as in control.

Text Mining

Integrating the ATF and AIT with the technology of text mining can provide meaningful insights in understanding large-scale emotional dynamics on social media platforms and their impact on the individual's decision-making and risk perception. Data mining can be described as a "strategy for discovering irregularities, examples and connection to predict results inside enormous arrangements of data" (Thange, et al., 2021, p.198). Furthermore, text mining is an information retrieval strategy that can be defined as a variation of data mining that is applied to textual datasets, to identify and extract knowledge (Agrawal & Batra, 2013; Tan, 1999). Subsequently, the extracted information can be further analysed for the discovery of insights or patterns within the large dataset. Hereby, text mining has proven to be a valuable tool for analysing social media climate change discussions (i.e., Hemsley, et al., 2021; Treen, et al., 2022; Villanueva, 2021; Williams, et al., 2015). While conducting sentiment analysis and emotion detection, one can gain valuable insights of the expressed opinions and experienced emotions of communities, which then can be used to predict the motivations of the public (Paltoglou, 2014).

Specifically, sentiment analysis is a technique which has been used in a variety of recent studies concerning social media communication to provide accurate predictions of opinions and emotional states shared by social media users (i.e., Hemsley, et al., 2021; Nemes & Kiss, 2021; Villanueva, 2021). Sentiment analysis is defined as "a technique to detect favourable and unfavourable opinions toward specific subjects within large numbers of documents" (Nasukawa & Yi, 2003, p.71), utilizing machine learning and natural language processing

(NLP). The technique aims to extract information from the data source that is subjective, capturing linguistic expressions of sentiments, opinions, emotions, and beliefs (Kumar & Sebastian, 2012; Wiebe, et al., 2004). A sentiment, thereby, represents a settled opinion, which reflects an individual's experienced emotions.

While sentiment analysis aims to determine the overall sentiment of the content, emotion detection provides a more thorough understanding of the emotions that were expressed in the content (Gaind, et al., 2019). Thereby, emotion detection aims to label the analysed content into various categories of discrete emotions, for example, happiness, fear, or anger. As this study is especially concerned with the discrete emotions of climate change believers and deniers, the technique of emotion detection can be used to gain further insights into the emotional content of climate change discussions. Overall, while implementing the technique of emotion detection, we can examine to which extent specific emotions were expressed in Reddit discussions. Furthermore, while further investigating the results of the analysis, we can better compare the emotional content and valence of the comments of climate change believers and deniers.

Lastly, in this study, a lexicon-based approach concerning sentiment analysis and emotion detection is chosen to classify the analysed content (Sham & Mohamed, 2022). Thereby, a lexicon is a coded dataset consisting of lexical features (e.g., words), which are labelled with scores indicating their semantic orientation or revealing the expressed emotions, in terms of their overall polarity and intensity (Hutto & Gilbert, 2014).

Aims and Hypotheses

This research study aims to further investigate climate change discussions occurring on Reddit while focussing on the discrete emotions of believers and deniers of climate change. To achieve further insights into the shared emotions of both groups, the technique of text mining is used. Specifically, sentiment analysis and emotion detection will be carried out in this study,

to answer the research question "How do the discrete emotions expressed in social media posts vary between climate change deniers and believers?". Hereby, Reddit content from subreddits created for both climate change believers and deniers were used in the data analysis. The study will particularly focus on the expressed emotions of both groups and compare them to, finally, gain further conclusions about both groups' dynamics, and their motivation for climate action, while further examining the formation of echo chambers and polarization. To answer the research question, hypotheses are formulated, which are informed by the previously mentioned theoretical framework.

H1: Content posted in climate change denier subreddits shows higher engagement than content posted in believer subreddits.

H2: Climate change denier messages contain significantly more negative emotions than believer messages.

H3: Messages of climate change deniers contain significantly higher levels of anger than messages of believers.

H4: Messages of climate change believers contain significantly higher levels of fear and hope than messages of deniers.

Methods

Overview

Our research methodology focused on techniques for extracting, pre-processing, and analysing data due to the complexity of the Reddit dataset used in this study. Several steps were taken to scrape the relevant data, prepare it for analysis, and, finally, analyse the sentiments and emotions expressed in the texts to answer the aforementioned research question and evaluate the hypotheses. Figure 1 illustrates the steps that were taken in this process. We used lexicon-based sentiment analysis to gain insight into the emotional polarity of each community; lexicon-based emotion detection is used to examine the emotional content expressed in the data.

Figure 1

Text Mining Pipeline.
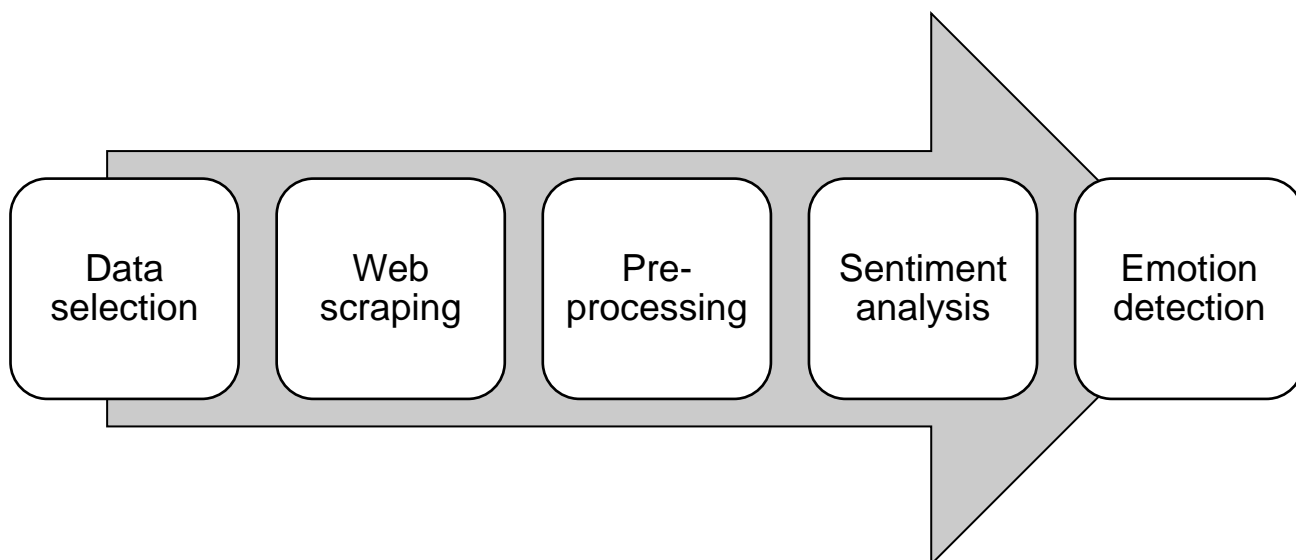
Data Collection

Selection of Subreddits

In this study, the data was collected from posts and comments across various subreddits explicitly designated as forums for climate change discussions, as indicated by their names and accompanying descriptions. For the sample of climate change believer content, we collected data from four subreddits previously used in the study by Parsa, et al. (2022), who examined climate change discussions on several subreddits. The subreddits the authors named "climate subreddits" were chosen, including the subreddits r/climate, r/environment, r/climatechange, and r/climateOffensive (Parsa, et al., 2022). Moreover, content from the subreddit r/climateskeptics was collected, specifically containing climate change denier content. Notably, there were significantly more subreddits created for those who believe in climate change than for those denying climate change, as the r/climateskeptics subreddit was the only active denier subreddit of noticeable size. Consequently, we decided to also collect content from other subreddits that deal with conspiracy theories, denial, or scepticism about science as a whole.

Namely, we collected content from the subreddits r/conspiracy, r/conspiracy_commons, and r/conspiracyNOPOL.

To ensure that the collected posts from all subreddits concerned climate change, a keyword search within these was carried out, to filter content which was not fitting for the inclusion criteria of this study. In respect to previous research carried out by Treen, et al., (2022), who also used a keyword approach while examining climate change discussions, we identified and collected posts that contained the words "global warming", "globalwarming", "climate change", "climatechange" or "climate" within these subreddits.

Subreddits of Climate Change Believers. Firstly, r/climatechange is the first subreddit we included, possessing 88.1 thousand members. As this subreddit is devoted to climate change, thereby, including climate change believers and deniers, it was of interest for this study. The subreddit describes itself as "A place for a rational discussion on a divisive topic.", while rules are present to explicitly promote a rational, deliberate discussion. The mentioned rules prohibit disparaging other users, mentioning politics in the discussion, and, especially, discouraging people from convincing other users that climate change matters. By requiring its users to adhere to these guidelines, the subreddit aims to foster an environment where users convinced of climate change are not offended by climate change deniers, but where deliberate discussions are present.

The r/climate subreddit was included in the data collection, which concerns "Information about the world's climate" as it is stated in the description accompanying the subreddit, thereby, also including information regarding activism and politics. The members of this subreddit amount to 183 thousand. Furthermore, the subreddit prohibits the spread of conspiracy theories and science denial in its discussions.

In addition, content from the r/environment subreddit was collected, which regarded itself as a place for discussion regarding "current news, information and issues related to the

environment", while prohibiting the spread of misinformation and non-informational posts, like satire or memes. Consequently, this subreddit can be identified as not primarily designed for climate change discussions, but as a forum concerning the environment.

Content from the r/climateOffensive subreddit was included, which concerns posts specifically relating to direct climate action, thereby prohibiting the spread of misinformation, propaganda, or science denial. This subreddit also prohibits the creation of news posts, as it describes itself as "a place to organize and to act". 69.7 thousand users are members of this subreddit.

Subreddits of Climate Change Deniers. Moreover, we also aimed to include content from a subreddit that was designed to focus on climate change denial. In respect to this, we included content gathered from the r/climateskeptics subreddit, which describes itself as "Climate Skeptics: Trying to see through the alarmism" and as "questioning climate related environmentalism.". Consequently, this subreddit is noticeably different from the previously addressed subreddits providing a distinct perspective of the community of climate change deniers. The community of the subreddit consists of 43 thousand members and prohibits its users from disparaging the subreddit and other members.

In addition, as earlier mentioned, we included content about climate change featured on three different conspiracy theory subreddits, to balance the amount of data collected for both parties. Namely, we included content from the r/conspiracy subreddit, which describes itself as "a forum for free thinking and for discussing issues which have captured your imagination", comprises 2.1 million members, making it the largest subreddit included in this research. Simultaneously, the r/conspiracy_commons subreddit, describing itself as a "provisional conspiracy sub", was created to accommodate Reddit users who were unable to participate in r/conspiracy due to account age restrictions. Despite its provisional nature, this subreddit is actively engaged with a membership of 192 thousand users and follows the same rules as its

parent subreddit. Lastly, the r/conspiracyNOPOL subreddit is similar to the previous conspiracy subreddits, however, it explicitly describes itself as "a place to discuss conspiracies – with the primary exception of domestic politics, especially US presidential politics". The last mentioned subreddit encourages their members to create posts about conspiracies, while replies to the posts should be "of a conspiratorial mindset".

Amount of Content Collected

Content from the mentioned climate change denier and believer subreddits was collected on the 8th of April 2024. Thereby, the data was comprised of content posted within the last year, beginning with the first entry posted on the 12th of April 2023. The content from the subreddits concerning conspiracy theories was collected within the same time range. Overall, 1547 posts and 41596 corresponding comments from all subreddits were collected. Thereby, 608 posts and 18304 comments were gathered from the climate change denier subreddits, while 939 posts and 23292 comments were collected from the climate change believer subreddits.

Sampling Strategy

The studies' sampling strategy could be described as purposive sampling because we intentionally selected the included subreddits and posts based on their relevance to the research question (Sibona & Walczak, 2012). The collected content specifically contains discussions concerning climate change; therefore, they were suitable to be included in this study. Moreover, as previously noted, we purposefully selected subreddits and the amount of content included from each subreddit to ensure that no bias towards content from a specific group is apparent, while maintaining a balanced representation of two homogeneous groups. Furthermore, the inclusion of subreddits from both groups allowed for a more nuanced analysis. Hereby, it enabled us to better understand the differences in opinions and emotions expressed while comparing both groups. Lastly, the selection of this sampling technique acknowledged the

limitation that there were more existing subreddits concerned about environmentalism and climate change believers than they were about climate change denial.

Scraping Reddit Content

After determining the relevant subreddits, the relevant content was extracted from Reddit and incorporated into a data set using the technique of *web scraping*. Hereby, web scraping describes "a technique to extract data from the World Wide Web (WWW) and save it to a file system or database for later retrieval or analysis" (Zhao, 2017, p.1). This data, in turn, can be integrated with natural language processing to examine the way a human user browses the web (Yi, et al., 2003, as cited in Zhao, 2017). In this study, the R package "RedditExtractoR", created by Rivera (2022), was used to scrape Reddit content. The used R package is freely accessible on both the comprehensive R archive network (CRAN) and on GitHub. In addition to the collected posts, their subsequent comments, upvotes and downvotes were collected. Additional metadata was also collected. This included the date the post or comment was created. Lastly, the scraped datasets were stored into a comma-separated values (CSV) file.

Reddit API Restrictions

In 2023, the access to Reddit's Application Programming Interface (API), which was previously freely available to the public and third-party applications, was restricted (Brown, et al., 2024; Juel, et al., 2024; Wright, 2024). The API enables researchers to collect and further analyse social media data; hence, the amount of data that can be collected has been considerably reduced (Brown, et al., 2024; Trezza, 2023).

Particularly, the RedditExtractoR package intentionally limits the number of collected posts and comments to comply with Reddit's API restrictions (Rivera, 2019, Yang, et al., 2024). The package allows for the collection of maximum 500 comments per Reddit post (Juel, et al., 2024). In addition, there are rate limits on the number of API requests that can be made, which

result in further constraints to the amount of data that can be collected in a single session. These restrictions can cause similar numbers of extracted posts and comments across the various subreddits, especially when dealing with popular or highly active subreddits.

Pre-processing of Data.

The accuracy of the emotion detection and sentiment analysis relies on the quality and preprocessing of the dataset used (Gard, et al., 2023). Social media content is frequently unstructured, consequently, if the data contains a significant amount of noise or errors, the analysis may, in turn, be less accurate (Nandwani & Verma, 2021). To ensure the accuracy of the analyses, the dataset was pre-processed and cleaned sufficiently and accordingly.

First, the documents in the dataset were be cleaned by converting the text to lowercase and removing stop words, accents, numbers, and URLs that are unnecessary to be included in the analysis (Garg & Sharma, 2022). By removing unnecessary text data, the performance of the analysis can be improved (Kannan & Gurusamy, 2014). Furthermore, word stemming will be applied to normalize the text and increase the accuracy and performance of the subsequent sentiment analysis and emotion detection (Arora, et al., 2023).

Sentiment Analysis

In this study, the technique of sentiment analysis was used to gain insights into the expressed emotions and opinions of users, who engaged in climate change discussions. Furthermore, with the use of the analysed data, we compared the expressed emotions and opinions of users identifying as either climate change believers or deniers. A lexicon-based sentiment analysis was carried out. The Valence Aware Dictionary for sEntiment Reasoning (VADER) (Hutto & Gilbert, 2014) was used, which is a sentiment analysis tool designed to be sensitive specifically to expressed sentiments in the social media context. For the application of VADER, the r package "vader" was used (Rochrick, 2020). The dictionary was extensively empirically evaluated. Furthermore, VADER utilizes a general sentiment lexicon and general

rules related to grammar and syntax. Consequently, it performs well in various applications and fields without requiring to be trained while using an extensive set of training data. The rules and the lexicon used by VADER are easily accessible, therefore, someone, even without a background in computer science, is able to easily understand the model and further modify it. Lastly, after examining its accuracy, the creators stated that it produces outputs of high quality while outperforming human raters.

Emotion detection

The technique of emotion detection was used to investigate the expressed emotional expressions of Reddit users and the intensity of those emotions. The NRC Word-Emotion Association Lexicon (EmoLex), created by Mohammad (2010; 2011; 2013) is chosen for emotion detection. The EmoLex is comprised of a lexicon of English words and associates them to eight different basic emotions and two different sentiments. This lexicon consists of commonly used English words and terms prevalent in the social media context (Mohammad, 2017). Consequently, we can compare to what extent the social media content of climate change believers and deniers differ in terms of emotions expressed. The association process is empirically evaluated and scores high on reproducibility.

Results

Descriptive Metadata

Number of Collected Content per Subreddit

Table 1 shows the frequency of collected content from climate change denier subreddits, while Table 2 displays the frequency of collected content posted on climate change believer subreddits. Overall, the community of climate change deniers exhibited a higher volume of both posts and comments compared to the community of climate change believers. Notably, the subreddit r/ClimateChange, which represents the community of climate change believers, can be highlighted as most actively engaging in discussions regarding climate change, with a

substantial number of posts and comments. Contrastingly, the content collected from the subreddit r/conspiracyNOPOL was comprised of 7 posts and 590 comments, highlighting it as the subreddit with the lowest frequency of climate change discussions.

Table 1

Frequency of Collected Climate Change Deniers' Content

Subreddit	Posts	Comments
r/ClimateSkeptics	250	5082
r/Conspiracy	246	8703
r/Conspiracy_commons	105	3929
r/ConspiracyNOPOL	7	590
Total	608	18304

Table 2

Frequency of Collected Climate Change Believers' Content

Subreddit	Posts	Comments
r/Climate	249	2587
r/ClimateChange	246	15068
r/ClimateOffensive	194	2099
r/Environment	250	3538
Total	939	23292

Moreover, Table 3 and Table 4 display the ratio of comments per post on the subreddits of both communities. It can be observed that, on average, subreddits of climate change deniers tend to have a higher ratio of comments per post compared to believer subreddits. Apart from the r/climatechange community ($M = 72.52$, $SD = 98.25$), the ratio of comments per post from each of the other believer subreddits was comparably lower. In addition, the subreddit, which had the highest ratio of comments per post was the r/conspiracyNOPOL subreddit ($M = 104.43$, $SD = 65.34$), which was also the subreddit with the lowest amount of collected content.

Table 3

Ratio of Comments per Post on Climate Change Denier Subreddits

Subreddit	Mean	SD	Max
r/ClimateSkeptics	21.10	27.05	155
r/ConspiracyNOPOL	104.43	65.34	208
r/Conspiracy_commons	42.12	54.97	300
r/Conspiracy	39.42	81.37	496

Table 4

Ratio of Comments per Post on Climate Change Believer Subreddits

Subreddit	Mean	SD	Max
r/Climate	72.52	98.25	654
r/ClimateChange	14.71	32.51	219
r/ClimateOffensive	12.27	30.66	227
r/Environment	11.10	20.03	110

Word clouds

Figure 2 displays word clouds which contain the most frequently mentioned words in the content posted by climate change deniers and believers, while Figure 3 displays the most unique mentioned words of both groups. Upon inspecting the word clouds of both groups, both similarities and differences can be observed.

Both groups mention "climate", "change", and "people" most frequently, emphasizing climate change as a societal issue. However, the language of climate change deniers is characterized by emphasizing scepticism, doubt and conspiracy theories, while frequently mentioning words such as "believe", "dont", "know", "think", as well as "hoax", "control", and "propaganda". This is further highlighted by a variety of the most unique words the group uses, which, for example, are "cabal", "mRNA", "illuminati", "Rockefeller", "masons". Notably, the most unique word mentioned by climate change deniers is archiveis, which is an online service for archiving web pages.

Contrastingly, climate change believers emphasize scientific terms and concepts such as "carbon", "warming", "science", "data", "temperature", and "°C" as well as "°F". Among

the most unique words mentioned by believers are "factcheckers", "factchecked", and "factchecks", highlighting the believers' handling and concern of misinformation.

All in all, the differences in the words used by both groups highlight the contrasting issues mentioned. While the language of the believers contains more scientific terms and concepts, deniers tend to express doubt and broader narratives of conspiracy and control.

Figure 2

Word clouds Displaying the Most Frequently Mentioned Words

Climate Change Denier

Climate Change Believer

Figure 3

Word clouds of the Most Unique Mentioned Words

Extracted Sentiments using VADER

Table 5 displays the sentiments extracted from comments posted in the climate change denier communities using the VADER sentiment analysis. Notably, the r/conspiracyNOPOL subreddit displayed the highest proportion of positive sentiments at 42.22% with the highest weighted average score of the deniers (0.10). The comments of r/climateskeptics demonstrated a nearly balanced distribution across sentiments, with a weighted average sentiment score close to neutral (0.02). Overall, none of the denier comments contained more negative sentiments compared to positive sentiments, which can be observed due to all weighted averages scores being positive.

Table 5

Extracted VADER Sentiments of Comments in Climate Change Denier Subreddits

Subreddit	Negative	Neutral	Positive	Weighted average (-1.0 – 1.0)
r/Climateskeptics	34.69%	27.63%	37.69%	0.02
r/Conspiracy	30.24%	30.35%	39.41%	0.06
r/Conspiracy_commons	32.55%	28.62%	38.83%	0.04
r/ConspiracyNOPOL	28.38%	29.40%	42.22%	0.10

In contrast, Table 6 shows the sentiment distribution detected within comments from climate change believer subreddits using the VADER sentiment analysis. Among these, the comments in r/climateoffensive emerged with the highest proportion of positive sentiments at 57.7% with the highest weighted average score (0.27). Notably, the comments of r/climate and r/environment exhibited a balanced distribution of sentiments close to the comments of r/climateskeptics, despite representing a climate change believer community.

Overall, when comparing the extracted sentiments of both communities, the believer subreddits stand out as indicating slightly more positive sentiments compared to the denier subreddits. Notably, the believer subreddits r/climatechange and r/climateoffensive contained

the highest proportions of positive sentiments, indicated by the positive weighted average scores. This suggests that comments in believer subreddits were generally slightly more positive in sentiment. However, all denier subreddits contained more comments that could be evaluated as possessing slightly more positive sentiments.

Table 6

Extracted VADER Sentiments of Comments in Climate Change Believer Subreddits

| Subreddit	Negative	Neutral	Positive	Weighted average (-1.0 – 1.0)
r/Climate	34.04%	24.64%	41.33%	0.05
r/Climatechange	28.97%	25.15%	45.88%	0.13
r/Climateoffensive	24.45%	17.85%	57.7%	0.27
r/Environment	33.88%	26.24%	39.87%	0.04

Sentiment Score by overall comment score

Figure 4 displays the relationship between the overall score of each comment extracted from the climate change denier subreddits with the comment's overall sentiment score. Hereby, the score of each comment resembles the number of upvotes and downvotes the comment received, while an upvote increases the score by 1 and a downvote decreases the score by 1. The hexbin plot of Figure 3 displays the density of comment score and sentiment score combinations, while the colour of the individual hexbin indicates the number of comments. The regression line suggests a negative relationship between these variables. The regression equation hereby is:

Sentiment Score = 0.0218-0.00156 x Comment Score

The adjusted r-squared value of less than 0.01 indicates that the comment score explains very little of the variance in the sentiment score.

Figure 4

Climate Denier Comment Score and its Relationship with the Comment's Sentiment
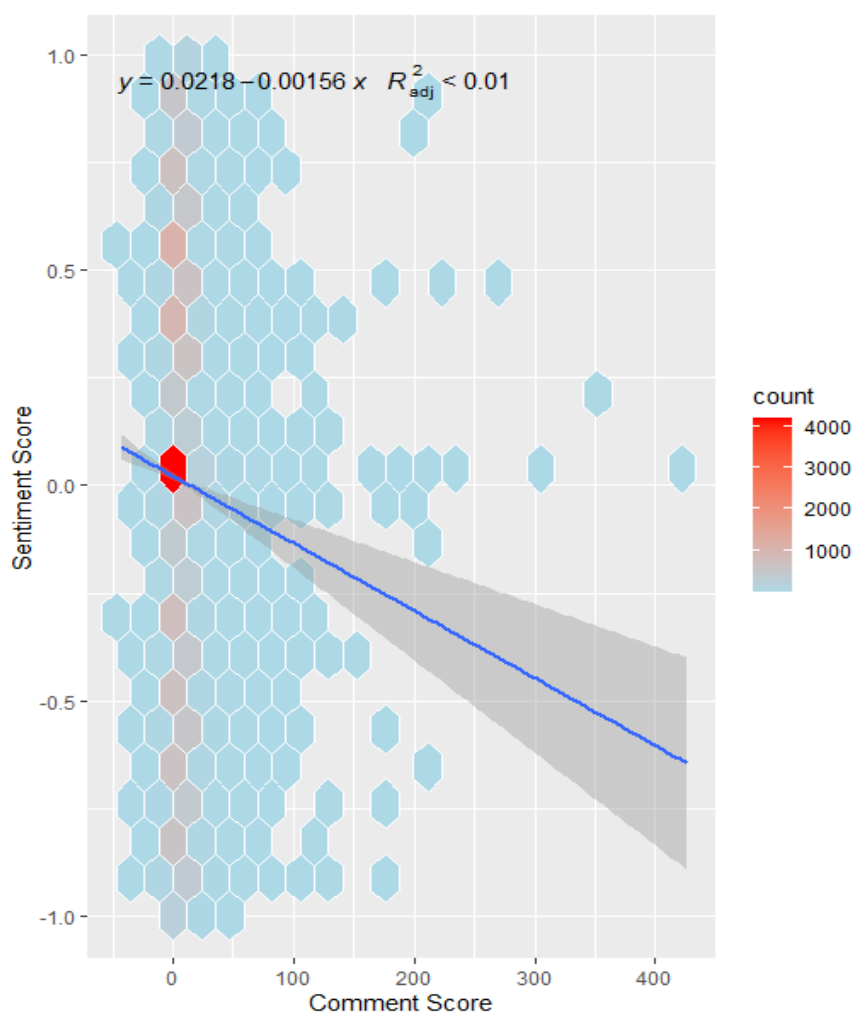
Note. The y-axis displays the sentiment score from negative (-1.0) to positive (1.0), which features the VADER compound sentiment. The x-axis features the comment score, which is made up of the number of likes and dislikes combined. The colour of each hexogram indicates the number of comments falling into that category from low (blue) to high (red).

Contrastingly, Figure 5 displays the relationship between the overall score of the climate change believers' comments with their sentiment score. The regression line indicates a negative relationship between these variables. The regression equation is:

Sentiment Score = 0.0756 – 0.000727 x Comment Score

The adjusted r-squared value of less than 0.01 suggests that the comment score explains very little of the variance in the sentiment score.

Figure 5

Climate Believer Comment Score and its Relationship with the Comment's Sentiment
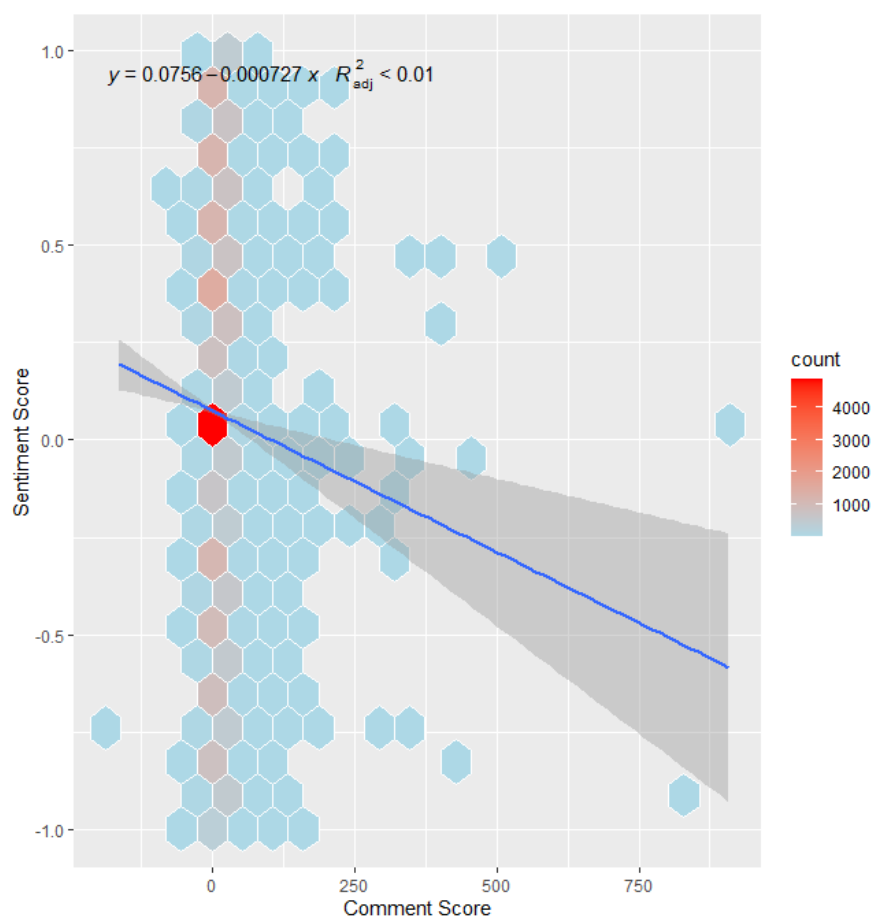
$y = 0.0756 - 0.000727 \, x \quad R^2_{adj} < 0.01$

Note. The y-axis displays the sentiment score from negative (-1.0) to positive (1.0), which features the VADER compound sentiment. The x-axis features the comment score, which is made up of the number of likes and dislikes combined. The colour of each hexogram indicates the number of comments falling into that category from low (blue) to high (red).

Extracted Sentiments Using the NRC Emotion Lexicon

Table 7 highlights the identified sentiments from comments posted in the climate change denier community while using the NRC emotion lexicon. Most of the denier subreddits show similar proportions of negative sentiments in relation to the identified positive sentiments. Namely, negative sentiments outweigh the positive sentiments leading to weighted average scores from -0.14 to -0.18 in those subreddits. However, in the r/conspiracyNOPOL subreddit, a nearly balanced distribution of sentiments can be observed, with a weighted average score of -0.04. Overall, in all subreddits a higher distribution of negative sentiments in relation to positive sentiments could be observed.

Table 7

Extracted NRC Sentiments of Comments in Climate Change Denier Subreddits

Subreddit	Negative	Positive	Weighted average (-1.0 – 1.0)
r/Climateskeptics	58.27%	41.73%	-0.16
r/Conspiracy	59.21%	40.79%	-0.18
r/Conspiracy_commons	57.16%	42.84%	-0.14
r/ConspiracyNOPOL	52.47%	47.53%	-0.04

Table 8 displays the sentiments identified in the comments posted in climate change believer subreddits while using the NRC emotion lexicon. Notably, all climate change denier subreddits show similar levels of identified sentiment proportions. In the comments of all subreddits, negative sentiments did outweigh positive sentiments, with weighted average scores from -0.12 to -0.18. The r/climatechange subreddit displayed the highest number of negative sentiments, with a weighted average score of -0.18.

Table 8

Extracted NRC Sentiments of Comments in Climate Change Believer Subreddits

Subreddit	Negative	Positive	Weighted average (-1.0 – 1.0)
r/Climate	58.11%	41.89%	-0.16
r/Climatechange	58.77%	41.23%	-0.18
r/Climateoffensive	55.93%	44.07%	-0.12
r/Environment	57.55%	42.45%	-0.16

Emotion Detection

Overall Emotions Detected in each Subreddit

The analysis of detected emotions in the individual subreddits revealed subtle differences in the expressed emotions between climate change deniers and believers, as shown

in Tables 7 and 8. Notably, r/conspiracy showed the highest level of fear (17.27%) among the deniers, with anger (15.09%) and trust (15.47%) also being prevalent emotions. Contrastingly, r/conspiracyNOPOL exhibited the highest levels of anticipation (14.18%) and trust (19.62%), while displaying the lowest levels of fear (15.06%) and anger (11.52%). Believers expressed similar levels of emotions detected, with r/ClimateChange showing the highest levels of fear (17.19%) and anger (15.15%).

These findings indicate the emotional landscape within each subreddit. Fear is consistently high in both groups, while anger is similarly expressed. Notably, the r/conspiracyNOPOL subreddit showed significant variations in different emotions. For example, trust was most prevalent in this subreddit (19.62%), while the highest trust levels among believers were expressed in the r/Climate subreddit (16.62%), highlighting a notable difference how trust is expressed across the two groups. Overall, while the levels of expressed emotions remained relatively consistent across both groups, each subreddit exhibited unique differences in the prevalence of certain emotions.

Table 7

Detected Emotions in the Corpus of Climate Change Denier Subreddits

Subreddit	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
r/ClimateSkeptics	15.25%	12.22%	11.65%	16.9%	8.51%	12.68%	6.91%	15.88%
r/Conspiracy	15.09%	11.63%	11.44%	17.27%	8.32%	13.38%	7.39%	15.47%
r/Conspiracy_commons	14.37%	12.23%	10.78%	17.14%	8.29%	12.65%	7.53%	17%
r/ConspiracyNOPOL	11.52%	14.18%	10.13%	15.06%	10.38%	11.14%	7.97%	19.62%

Table 8

Detected Emotions in the Corpus of Climate Change Believer Subreddits

Subreddit	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
r/Climate	14.66%	12.22%	11.1%	16.55%	8.73%	12.92%	7.19%	16.62%
r/Climatechange	15.15%	11.49%	12%	17.19%	8.29%	12.84%	7.28%	15.75%
r/ClimateOffensive	13.73%	13.45%	10.26%	16.57%	9.08%	13.04%	7.35%	16.5%
r/Environment	14.29%	11.76%	11.51%	17.33%	8.22%	13.41%	7.27%	16.19%

Measuring Fear and Anger

An analysis of variance (ANOVA) was carried out to examine the effect of group (deniers vs believers) on the normalized emotion scores for the two emotions "fear" and "anger". The emotion scores, hereby, represent the aggregated total count of words associated with each emotion. These scores are normalized per 100 comments to account for the varying amount of content collected from the individual subreddits. The main effects of the type of emotion and group on the emotion score were evaluated, in conjunction with their interaction.

The results indicated that there was no significant main effect of emotion $F(1,12) = 0.31$, $p = .59$. Consequently, the emotions "fear" and "anger" did not significantly differ in their scores, suggesting similar levels of expression. Similarly, no significant main effect of group could be observed $F(1,12) = 0.18$, $p = .68$. This indicates that there was no significant difference between both groups in terms of their overall normalized emotional scores for "fear" and "anger". In addition, the interaction effect between emotion and group on the score was not significant $F(1,12) = 0.01$, $p = .906$. This suggests that the difference in normalized emotional scores of "fear" and "anger" does not correspond to the group.

Measuring Hope

An analysis of variance (ANOVA) was conducted to examine the effect of group (deniers vs believers) on the normalized emotion scores of "hope". The emotion score, hereby, represent the aggregated total count of words associated with the emotion, normalized per 100 comments. The main effect of the groups on the emotion score was evaluated.

The results indicated that there was no significant main effect of group $F(1,6) = 0.21$, $p = .66$. Therefore, there was no significant difference between both groups in terms of their overall score of "hope".

Discussion

The purpose of this research was to further investigate climate change discussions occurring on Reddit while focussing on the discrete emotions of believers and deniers of climate change. Specifically, techniques of sentiment analysis and emotion detection were employed to achieve further insights into the differences in emotional expressions of both groups. Emotional expressions play a crucial role when analysing climate change discussions because they may influence the individual's behaviour, risk perception and decision-making (Heiss, 2020, as cited in Villanueva, 2021; Lerner & Keltner, 2000; 2001; Marcus, et al., 2000). Previous research on the emotional expressions in climate change discussions has mainly focussed on the social media platform Twitter (Hemsley, et al., 2021; Pearce, et al., 2018; Williams, et al., 2015), while Reddit possesses a different architecture and mechanisms that may encourage climate change discussions among its users (Pearce, et al., 2018; Treen, et al., 2022; Villanueva, 2021).

Our findings indicated that climate change discussions of believers and deniers were similar in the levels of emotional expressions. Namely, fear was the most expressed emotion of both groups, with trust and anger being emotions that were consistently identified as the second and third most. Furthermore, slight differences could be observed in terms of the proportions of the expressed sentiments of both groups. The comments of climate change believers and deniers contained similar levels of positive and negative sentiments, while slight differences could be observed among the individual subreddits. Lastly, the engagement levels of both groups significantly differed, as climate change deniers wrote more comments to each post than believers did.

Emotional Expressions

Our findings indicated that the most expressed emotion of both climate change deniers and believers was fear. These findings are in line with the research done by Amangeldi, et al, (2024), which indicated that the most expressed emotion in environmental posts on Reddit was fear. Consequently, since both groups express fear, they may be inclined to challenge their

preexisting views while reaching out for novel information because they mostly perceive the danger as not in control (Lerner & Keltner, 2000, 2001; Marcus, et al., 2000). Thus, the assumption that the deniers tend to evaluate the risk as in control while relying on their preexisting views can be refuted. However, our findings are surprising, because, in previous research, anger predominated in messages of deniers (Hemsley, et al., 2021; Villanueva, 2021), whereas believer messages most frequently contained the emotion hope next to fear (Villanueva, 2021). Contrary to popular belief, climate change deniers do experience feelings of worry and dread concerning environmental problems, such as dying species or pollution, as shown by Haltinner, et al. (2021). Consequently, they persist in engaging with conspiracy theories and climate denial to cope with these feelings (Haltinner & Sarathchandra, 2019). Hereby, they may tend to disconnect anthropogenic climate change from these environmental problems. However, to which aspects of climate change, natural catastrophes, and social issues the group of deniers of our sample expressed fear remains unclear. Furthermore, while climate change deniers are typically categorized as distrustful of scientific consensus (Beiser-McGrath & Bernauer, 2021; Maibach, et al., 2009; Poortinga, et al., 2019), our research identified high levels of expressed by the group of deniers. This notable finding can be explained by the dynamics of social identity and cognitive dissonance.

Climate change deniers may frequently engage with their community of deniers to reduce cognitive dissonance and strengthen their identity (Del Vicario, et al., 2016; Oswald & Bright, 2022). While interacting with subreddits centered around conspiracy theories and climate change denial they encounter arguments they assess as strong because these align with their views (Taber, et al., 2009). This communication allows them to establish a supportive network, in which they can distance themselves from opposing views and perceive climate denial as a consensus (Oswald & Bright, 2022). When being frequently confronted with dissonant content, their belief in their group's view may be strengthened (Karlsen, et al., 2017,

as cited in Oswald & Bright, 2022). This process may be facilitated by motivated social cognition, where individuals are more inclined to trust the opinions of their ingroup than the content of an outgroup (Jost & Amodio, 2012, as cited in Oswald & Bright, 2022). Consequently, frequent interaction within these networks could further strengthen the beliefs of climate deniers, as indicated by our findings.

Expressed Sentiments

The findings indicated that only minor differences in the proportions of sentiments could be observed between the subreddits of climate change deniers and believers. Discrepancies in the sentiment analysis results between the NRC and VADER lexicons could be observed. Namely, the results of the VADER sentiment analysis indicated balanced proportions of negative and positive sentiments, with a slight tendency towards positivity. Contrastingly, the NRC sentiment analysis highlighted a general trend towards negativity in the comments of both groups, while slight differences in the distribution of sentiments could be observed when inspecting the comments of each subreddit. These discrepancies in the results of the sentiment analyses highlight the importance of using multiple sentiment analysis tools to gain a comprehensive understanding of the data to be analysed.

Our findings revealed no significant differences in the overall negativity between denier and believer comments based on the NRC lexicon. Namely, based on the NRC sentiment analysis results, both groups expressed high amounts of negativity within their comments. The results of the VADER sentiment analysis underline that no substantial differences in the proportions of expressed sentiments between both groups could be observed, with a general slight tendency of the comments to contain positivity. Notably, the comments of the denier subreddit r/conspiracy_commons contained considerably high amounts of positive sentiments. Moreover, the believer subreddits r/climatechange and r/climateoffensive also contained increased levels of positive sentiments when compared to negative sentiments.

Amangeldi, et al. (2024) pointed out that environmental discussions on Reddit exhibited a rise in the levels of expressed positive sentiments, in contrast to the discussions present on Twitter and Youtube, which consistently contained a higher frequency of negative sentiments. In these environmental discussions, climate change could be identified as the predominant topic. These findings present a contrast to the findings of previous studies examining the sentiments expressed in climate change discussions on Twitter, which indicated that negative sentiments were predominant in messages of climate change deniers (Hemsley, et al., 2021).

Increased involvement in the discussions inside an echo chamber tends to result in increased negativity among the members of communities of both conspiracy theorists and believers of science (Del Vicario, et al., 2016). Thereby, group polarization, which is reinforced by interactions among like-minded individuals, may further influence the behaviour and emotions of the members of the community (Del Vicario, et al., 2016; Oswald & Bright, 2022). High interaction within these groups often corresponds to increased negativity and engagement, while the most active members of the community express more polarized views (Baumann, et al., 2020). However, according to the findings of the NRC sentiment analysis, the comments of deniers did not contain higher proportions of negative sentiments when compared to the believers, even though higher engagement could be observed in the community of deniers.

Engagement Levels

Our analysis revealed that posts published in climate change denier subreddits showed a higher ratio of comments per post compared to the posts in climate change believer subreddits. Consequently, higher engagement levels could be observed among deniers in comparison to believers. This finding may be explained by the research by Oswald & Bright (2022), who found that dissonant submissions to the r/ClimateSkeptics subreddit led to increased engagement. Cognitive dissonance, the formation of echo chambers and social identity may further contribute to the discovered differences in the rates of engagement.

Climate change deniers can be identified as a minority group (Beiser-McGrath & Bernauer, 2021; Maibach, et al., 2009; Poortinga, et al., 2019). Moreover, during recent times, the COVID-19 crisis emerged as a critical event, which has disrupted the media coverage of the climate crisis, which previously was at its peak (Stoddard, et al., 2023). Lastly, a recent growth of environmental discussions on Reddit could be observed, particularly in 2023 (Amangeldi, et al., 2024). The combination of these factors could have caused climate change deniers to be more frequently exposed to content created about the climate crisis, increasingly confronting them with content that questions or threatens their identity (Amengeldi, et al., 2024).

When confronted with opposing views, individuals tend to defend their identity and reduce their feelings of dissonance by posting identity-reinforcing content within their community (Oswald & Bright, 2022). This selective engagement with content that affirms their views can be named as confirmation bias (Falkenberg, et al., 2022; Hart, et al., 2009; Walter, et al., 2018). Furthermore, the heightened community engagement may be caused by the formation of echo chambers and social identity (Del Vicario, et al., 2016; Oswald & Bright, 2022).

On Reddit, users can choose to engage in subreddits that reinforce their existing beliefs, leading to more frequent interactions within their community. Consequently, contact with the ingroup may result in an increased desire to maintain one's identity, while engaging more with the ingroup and discrediting the outgroup (Oswald & Bright, 2022). Climate change deniers thus interact more with like-minded individuals to reduce their cognitive dissonance and strengthen their preexisting opinions and social identity. Hereby, further confrontation with opinions that may threaten their identity is hindered, while the formation of more extreme views within the echo chamber is fostered (Farrell, 2015).

In summary, while the hypothesis that denier content shows higher engagement is confirmed, the reasons behind this are complex. As Reddit users are free to choose what subreddit they want to engage with, the interplay of various factors such as the formation of echo chambers, cognitive dissonance, confirmation bias, subreddit selection, and broader social dynamics may contribute to the observed engagement differences.

Strengths and Limitations

Our data collection process was affected by the recent restrictions to Reddit's API, which have limited the volume of social media content available for research (Brown, et al., 2024; Rocha-Silva, et al., 2024; Trezza, 2024; Wright, 2024). This limitation of our study is reflected in the uniformity in the number of posts that were extracted across the various subreddits that were chosen. Specifically, the chosen RedditExtractoR package intentionally limits the number of collected posts and comments to comply with Reddit's API changes (Juel, et al., 2024; Rivera, 2019; Yang, et al., 2024). Moreover, Reddit limits the number of posts accessible by browsing individual subreddits to one thousand posts (Rocha-Silva, et al., 2024). This could potentially impact the representativeness of our sample, especially for highly active subreddits where not all content could be collected. Consequently, our analysis does not capture the whole picture of the discussions within climate change communities on Reddit. This could cause an incomplete understanding of the dynamics and sentiments within climate change discussions on Reddit. Additionally, the uniformity in data collection across the subreddits may obscure the observed differences in engagement levels.

The Pushshift Reddit dataset can be used in future research to overcome these limitations by providing extensive historical data archives and allowing for increased query limits (Baumgartner, et al., 2020). The dataset consists of the content extracted from the top 40 thousand subreddits from June 2006 up until December 2023 and is available online on the website Academic Torrents (Cohen, et al., 2024; Lo & Cohen, 2016). The pushshift dataset has

been used in various peer-reviewed publications and enables the researcher to quickly collect and analyse large amounts of data (Cummings & Lipworth 2023; Rocha-Silva, et al., 2024; Stillman & Kruspe, 2024).

However, it is important to acknowledge the limitations of the pushshift dataset. Firstly, the dataset may contain content that has been removed from Reddit, which raises ethical considerations (Rocha-Silva, et al., 2024). Furthermore, the use of a data repository can be characterized as a constraint, as the dataset may be incomplete (Gaffney & Matias, 2018, as cited in Rocha-Silva, et al., 2024). Despite these potential limitations, the Pushshift dataset remains a valuable resource for overcoming the constrains imposed by Reddit's API restrictions on the data collection process.

Furthermore, the selection of subreddits can be highlighted as a limitation. Namely, only one subreddit of substantial size could be identified which revolved around climate change denial, namely, r/ClimateSkeptics. By implementing a keyword search, we ensured that the content posted by the Reddit users contained impressions regarding climate change and global warming. However, while supplementing the dataset of climate change deniers with climate change content posted on conspiracy theories subreddits, our dataset contained a majority of individuals who identify themselves as conspiracy theorists. Climate change deniers have different motives other than conspiracy-based narratives that have caused them to deny anthropogenic climate change (Haltinner & Sarathchandra, 2021). Moreover, individuals are inclined to support conspiracy theory beliefs due to mechanisms of social identity, which may further cause increased engagement (Del Vicario, et al., 2016; Oswald & Bright, 2022; Robertson, et al., 2022). This difference in engagement levels when comparing the r/ClimateSkeptics subreddit with the conspiracy subreddits was identified in our findings. Consequently, future research is needed to evaluate the content created by climate change deniers. Specifically, we advise comparing the climate change content of climate denier

subreddits with conspiracy subreddits, to gain a better understanding of the different motives of individuals that cause them to deny the climate crisis and to gain further insights into the discovered differences in engagement patterns.

Moreover, the use of a lexicon-based approach to identify the expressed emotions and sentiments of climate change deniers and believers can be emphasized as a limitation. Hereby, hope is a complex emotion that is not included in a sentiment- or emotion lexicon (Guerra & Karakuş, 2023). Therefore, based on the findings by Guerra & Karakuş (2023), we used different subscales of the NRC-emotion lexicon to evaluate whether the content contains hope. Moreover, climate change discussions frequently contain sarcasm, an emotion that may be difficult to label for humans (Hemsley, et al., 2021; Verma, et al., 2021). Hereby, sarcasm can have effects on the accuracy of the sentiment analysis applied, as sarcasm frequently aims to imply negativity, while implying a positive sentiment on the surface (Amangeldi, et al., 2024; Joshi, et al., 2017; Maynard & Greenwood, 2014). In future research, deep learning models may be used to identify the contained sarcasm in climate change discussions and its effect on the detected sentiments and emotions (Razali, et al., 2021; Verma, et al., 2021). Machine learning and deep learning algorithms may provide higher accuracy and improved flexibility to complex concepts like the expression of hope and sarcasm (Hung & Alias, 2022).

Future Research

We recommend that future research into climate change discussions on Reddit should distinguish the opinions of conspiracy theorists, climate deniers, and believers to better understand the distinct motives of each group. Machine learning and deep learning algorithms could be employed to identify characteristics of the content created by these groups and to accurately detect complex emotions such as hope and sarcasm. This approach could offer a more nuanced understanding of the emotional dynamics within online climate change communities. Additionally, future research should explore how media coverage, political

events, and climate catastrophes influence sentiments and emotions expressed in climate change discussions. Examining trends in expressed emotional content over a broader timespan could provide valuable insights, and utilizing datasets like Pushshift could enable a more comprehensive analysis of the diverse range of climate change discussions online, without being restricted by API regulations.

Conclusion

This research aimed to compare the expressed discrete emotions and sentiments of climate change believers and deniers. The study found that both groups primarily expressed fear in their comments, with trust and anger being secondary emotions. Contrary to previous findings, slight differences were observed in the expressed sentiments of both groups, depending on the sentiment analysis tool used. The NRC analysis indicated that both groups expressed similar levels of negativity, whereas the VADER analysis suggested that believers' comments contained more positive sentiments than those of deniers. Additionally, deniers demonstrated higher engagement levels by posting more comments per post than believers. These findings suggest that cognitive dissonance, social identity theory, and emotional processes play a role in shaping the behaviour and decision-making of climate change deniers. Despite higher engagement levels among deniers, the expected formation of echo chambers and significantly increased negativity in denier content were not observed.

References

American Psychological Association (2017). *Mental Health and our Changing Climate: Impacts, Implications and Guidance*. APA, 2017. https://www.apa.org/news/press/releases/2017/03/mental-health-climate.pdf

Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., & Blackburn, J. (2020). The Pushshift Reddit Dataset. In *Proceedings of the International AAAI Conference on Web and Social Media, 14*(1), 830-839. https://doi.org/10.1609/icwsm.v14i1.7347

Beiser-McGrath, L. F., & Bernauer, T. (2021). Current surveys may underestimate climate change skepticism evidence from list experiments in Germany and the USA. *Plos One*, *16*(7), e0251034. https://doi.org/10.1371/journal.pone.0251034

Brown, M.A., Lukito, J., Greenfield, J., Chen, B., Graham, S., Shugars, S., Pruden, M.L. (2024). The state of digital media research, 2024. *Media & Democracy Data Cooperative*. https://repositories.lib.utexas.edu/server/api/core/bitstreams/b89a8cf5-10ff-457f-9cbd-6562ba4a5ff9/content

Capstick, S. B., & Pidgeon, N. F. (2014). What is climate change scepticism? Examination of the concept using a mixed methods study of the UK public. *Global Environmental Change*, *24*, 389-401. https://doi.org/10.1016/j.gloenvcha.2013.08.012

Chon, M. G., & Park, H. (2020). Social media activism in the digital age: Testing an integrative model of activism on contentious issues. *Journalism & Mass Communication Quarterly*, *97*(1), 72-97. https://doi.org/10.1177/1077699019835896

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. In *Proceedings of the National Academy of Sciences of the United States of America*, *118*(9), e2023301118. https://doi.org/10.1073/pnas.2023301118

Coen, D., Herman, K., Pegram, T. (2020) Private Sector and Climate Change. A Case Study of Carbon-Based Governance. *GLOBE,* 1-147. https://www.globe-project.eu/private-sector-and-climate-change-a-case-study-of-carbon-based-governance_11373.pdf

Cummings, J.A., Lipworth, A.D. (2023). Reddit and google activity to non-covid epidemic diseases surged at start of covid-19 pandemic: Retrospective study. *JMIR Formative Research, 7*, e44603. https://doi.org/10.2196/44603

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on Facebook. *Scientific Reports*, *6*(1), 37825. https://doi.org/10.1038/srep37825

Ding, K., Choo, W. C., Ng, K. Y., & Ng, S. I. (2020). Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation. *International Journal of Hospitality Management, 91*, 102676. https://doi.org/10.1016/j.ijhm.2020.102676

Dodds J. (2021). The psychology of climate anxiety. *BJPsych bulletin, 45*(4), 222–226. https://doi.org/10.1192/bjb.2021.18

Drače, S., & Ric, F. (2012). The effect of emotions on risk perception: Experimental evaluation of the affective tendencies framework. *Psihologija*, *45*(4), 409-416. DOI:10.2298/PSI1204409D

Efstratiou, F., Blackburn, J., Caulfield, T., Stringhini, G., Zannettou, S. & De Cristofaro, E. (2023). Non-polar opposites: Analyzing the relationship between echo chambers and hostile intergroup interactions on reddit. In *Proceedings of the Seventeenth Internation AAAI Conference on Web and Social Media (ICWSM 2023), 17,* 197-208. https://doi.org/10.48550/arXiv.2211.14388

Falkenberg, M., Galeazzi, A., Torricelli, M., Di Marco, N., Larosa, F., Sas, M., Mekacher, A., Pearce, W., Zollo, F., Quattrociocchi, W. & Baronchelli, A. (2022). Growing

polarization around climate change on social media. *Nature Climate Change*, *12*(12), 1114-1121. https://doi.org/10.1038/s41558-022-01527-x

Farrell, J. (2015). Echo chambers and false certainty. *Nature Climate Change*, *5,* 719–720. https://doi.org/10.1038/nclimate2732

Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from daguerrotype to laserphoto. *Journal of Personality and Social Psychology*,*63*(6), 877 – 889. doi: 10.1037//0022-3514.63.6.877

Forzieri, G., Girardello, M., Ceccherini, G. *et al.* Emergent vulnerability to climate-driven disturbances in European forests. *Nature Communications,* 12(1), 1081. https://doi.org/10.1038/s41467-021-21399-7

Gaffney, D., & Matias, J.N. (2018). Caveat emptor, computational social science: Large-scale missing data in a widely-published Reddit corpus. *PLoS ONE*, *13*(7), e0200162. https://doi.org/10.1371/journal.pone.0200162

Guerra, A., & Karakus, O. (2023). Sentiment analysis for measuring hope and fear from reddits posts during the 2022 russo-ukrainian conflict. *Frontiers in Artificial Intelligence, 6.* DOI: 10.3389/frai.2023.1163577

Haltinner, K., & Sarathchandra, D. (2018). Climate change skepticism as a psychological coping strategy. *Sociology Compass, 12*(6), e12586. https://doi.org/10.1111/soc4.12586

Haltinner, K., Sarathchandra, D. (2021) The nature and nuance of climate change skepticism in the united states. *Rural Sociology, 84*(4), 673-702. https://doi.org/10.1111/ruso.12371

Haltinner, K., Ladino, J., & Sarathchandra, D. (2021). Feeling skeptical: Worry, dread, and support for environmental policy among climate change skeptics. *Emotion, Space and Society*, *39*(12), 100790. DOI: 10.1016/j.emospa.2021.100790

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling valiyeard versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555. doi: 10.1037/a0015701

Heiss, R. (2020). The affective foundation of political opinion expression on social media: a panel analysis. *International Journal of Public Opinion Research*, *33*(1), 57-75, https://doi.org/10.1093/ijpor/edaa009

Hemsley, J., Lee, J. and Hopke, J. (2021). Climate Change Deniers Spread Negative Emotion Laden Tweets'. http://dx.doi.org/10.2139/ssrn.3837367

Hickman, C., Marks, E., Pihkala, P., Clayton, S., Lewandowski, R. E., Mayall, E. E., Wray, B., Mellor, C., van Susteren, L. (2021). Climate anxiety in children and young people and their beliefs about government responses to climate change: A global survey. *The Lancet Planetary Health*, *5*(12), 863-873. DOI: https://doi.org/10.1016/S2542-5196(21)00278-3

Huckelba, A. L., & Van Lange, P. A. (2020). The silent killer: Consequences of climate change and how to survive past the year 2050. *Sustainability*, *12*(9), 3757. doi:10.3390/su12093757

Hung, L.P., & Alias, S. (2023) Beyond sentiment analysis: A review of recent trends in text based sentiment analysis and emotion detection. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *27*(1), 84-95. doi: 10.20965/jaciii.2023.p0084

Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media, 8*(1), 216-225. https://doi.org/10.1609/icwsm.v8i1.14550

International Energy Agency (IEA) (2007), *Energy Use in the New Millennium:Trends in IEA Countries*, OECD Publishing. https://doi.org/10.1787/9789264034303-en.

Joinson, A. (1999), Social desirability, anonymity, and internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers*. *31*, 433–438. https://doi.org/10.3758/BF03200723

Jones, R., Cox, D., & Navarro-Rivera, J. (2014). Believers, sympathizers, & skeptics. *Public Religion Research Institute and American Academy of Religion*. https://www.prri.org/wp-content/uploads/2014/11/2014-Climate-Change-FINAL1.pdf

Joshi, A., Bhattacharyya, P., & Carman, M.J. (2017). Automatic sarcasm detection: A survey. *ACM Computing Surveys (CSUR), 50*(5), 1-22. https://doi.org/10.1145/3124420

Jost, J.T., & Amodio, D.M. (2012). Political ideology as motivated social cognition: Behavioural and neuroscientific evidence. *Motivation and Emotion*. *36*(1), 55-64. https://doi.org/10.1007/s11031-011-9260-7

Juel, E., Futch, W., Eiler, B. (2024). Connecting the pro-recovery eating disorder community: An analysis of the language on science, twitter, and reddit. *Psychology of Popular Media, 13*(2). 230-239. https://doi.org/10.1037/ppm0000461

Kannan, S., & Gurusamy, V. (2014). Preprocessing techniques for text mining. *Journal of Computer Science and Communications*, *5*, 7-16. https://www.researchgate.net/publication/273127322_Preprocessing_Techniques_for_Text_Mining

Karlsen, R., Steen-Johnsen, K., Wollebæk, D., & Enjolras, B. (2017). Echo chamber and trench warfare dynamics in online debates. *European Journal of Communication*, *32*(3), 257–273. https://doi.org/10.1177/0267323117695734

Kemp, S. (2019, March 4). *Reddit now has more active users than Twitter - and is more engaging than porn.* The New Web. https://thenextweb.com/contributors/2018/04/19/reddit-now-active-users-twitter-engaging-porn/.

Kumar, A., & Sebastian, T. M. (2012). Sentiment analysis: A perspective on its past, present and future. *International Journal of Intelligent Systems and Applications, 4*(10), 1-14. DOI: 10.5815/ijisa.2012.10.0

Laaksonen, S.L., Rantasila, A. (2021). Rocketing sheep: affective discipline in anonymous mobile social media jodel during the covid-19 pandemic. *AoIR Selected Papers of Internet Research.* doi: 10.5210/SPIR.V2021I0.12198

Lerner, J. S. & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition and Emotion, 14*(4), 473-493. https://doi.org/10.1080/026999300402763

Lerner, J. S. & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146-159. https://doi.org/10.1037/0022-3514.81.1.146

Ma, X., Hancock, J., & Naaman, M. (2016). Anonymity, intimacy and self-disclosure in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 3857-3869. https://doi.org/10.1145/2858036.2858414

Maibach, E., Roser-Renouf, C., & Leiserowitz, A. (2009). *Global warming's six Americas 2009: An audience segmentation analysis*. Center for Climate Change Communication, George Mason University. https://cdn.americanprogress.org/wp-content/uploads/issues/2009/05/pdf/6americas.pdf

MacKuen, M., Wolak, J., Keele, L., & Marcus, G. E. (2010). Civic engagements: Resolute partisanship or reflective deliberation. *American Journal of Political Science*, *54*(2), 440-458. https://doi.org/10.1111/j.1540-5907.2010.00440.x

Marcus, G. E., Neuman, W. R., & MacKuen, M. (2000). *Affective intelligence and political judgment*. University of Chicago Press.

Marcus, G. E., MacKuen, M., & Neuman, W. R. (2011). Parsimony and complexity: Developing and testing theories of affective intelligence. *Political Psychology*, *32*(2), 323-336. doi: 10.1111/j.1467-9221.2010.00806.x

Marcus, G. E., Valentino, N. A., Vasilopoulos, P., & Foucault, M. (2019). Applying the theory of affective intelligence to support for authoritarian policies and parties. *Political Psychology*, *40*, 109-139. doi: 10.1111/pops.12571

Maynard, D.G. & Greenwood, M.A. (2014). Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. In *Language Resources and Evaluation Conference (LREC)*, 4238-4243. http://www.lrec-conf.org/proceedings/lrec2014/pdf/67_Paper.pdf

McCright, A.M., & Dunlap, R.E. (2011). The politicization of climate change and polarization in the american public's views of global warming, 2001–2010. *The Sociological Quarterly, 52*, 155-194. DOI:10.1111/j.1533-8525.2011.01198.x

Marchal, N. (2020). The polarizing potential of intergroup affect in online political discussions: evidence from reddit r/politics. *Social Science Research Network*. doi: 10.2139/SSRN.3671497

Mohammad, S. M. & Turney, P. (2010). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, 26-34. https://aclanthology.org/W10-0204

Mohammad, S. M. (2011). *NRC Word-Emotion Association Lexicon*. https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

Mohammad, S. M. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence, 29* (3), 436-465. https://arxiv.org/abs/1308.6297

Mohammad, S. M. (2017). Word affect intensities. *arXiv preprint. arXiv:1704.08798*.

Nasukawa, T., & Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd international conference on Knowledge capture*, 70-77. http://dx.doi.org/10.1145/945645.945658

Nandwani, P., Verma, R. (2021) A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining. 11*(81), 1-19. doi:10.1007/s13278-021-00776-6

Nemes, L. & Kiss, A. (2021) Social media sentiment analysis based on COVID-19, *Journal of Information and Telecommunication*, *5*(1), 1-15, DOI: 10.1080/24751839.2020.1790793

Oswald, L., & Bright, J. (2022). How do climate change skeptics engage with opposing views online? Evidence from a major climate change skeptic forum on Reddit. *Environmental Communication*, *16*(6), 805–821. https://doi.org/10.48550/arXiv.2102.06516

Pariser, E. (2012). The filter bubble: What the internet is hiding from you. *InterActions: UCLA Journal of Education and Information Studies, 8*(2). 294. DOI: 10.5070/D482011835

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reddit.com. (2020). *Reddit's 2020 year in review*. https://www.redditinc.com/blog/reddits-2020-year-in-review/

Reddit.com. (2021). *Advertising—Audience—Reddit. Discover what makes Reddit ads unique*. https://web.archive.org/web/20210117184818/https://www.redditinc.com/advertising/audience

Reddit (2024). *Reddit Data API Wiki*. https://support.reddithelp.com/hc/en-us/articles/16160319875092-Reddit-Data-API-Wiki

Reddit.com. (2024). *What is Reddit?* https://www.redditforbusiness.com/learn/what-is-reddit

Rivera, I. (2019). RedditExtractoR: Reddit Data Extraction Toolkit. https://github.com/ivan-rivera/RedditExtractor/blob/master/README.md

Rivera, I. (2022). RedditExtractoR. https://github.com/ivan-rivera/RedditExtractor

Robertson CE, Pretus C, Rathje S, Harris E, Van Bavel JJ, How Social Identity Shapes Conspiratorial Belief. *Current Opinion in Psychology*, https://doi.org/10.1016/.copsyc.2022.101423.

Parsa, M. S., Shi, H., Xu, Y., Yim, A., Yin, Y., & Golab, L. (2022). Analyzing climate change discussions on reddit. In *2022 International Conference on Computational Science and Computational Intelligence (CSCI)*, 826-832. DOI: 10.1109/CSCI58124.2022.00093

Pearce, W., Niederer, S., Özkula, S. M., & Querubín, N. S. (2018). The social media life of climate change: Platforms, publics, and future imaginaries. *Wiley Interdisciplinary Reviews: Climate Change*, 10. https://doi.org/ 10.1002/wcc.569

Poortinga, W., Spence, A., Whitmarsh, L., Capstick, S., & Pidgeon, N. (2011). Uncertain climate: An investigation into public scepticism about anthropogenic climate change. *Global Environmental Change, 21*, 1015–1024. https://doi.org/10.1016/j.gloenvcha.2011.03.001

Poortinga, W., Whitmarsh, L., Steg, L., Böhm, G., & Fisher, S. (2019). Climate change perceptions and their individual determinants: A cross-European analysis. *Global Environmental Change-Human and Policy Dimensions, 55*, 25-23. DOI: 10.1016/j.gloenvcha.2019.01.007

Proferes, N., Jones, N., Gilbert, S., Fiesler, C., & Zimmer, M. (2021). Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media+ Society*, *7*(2), DOI: 10.1177/20563051211019004journals.sagepub.com/home/sms

Rocha-Silva, T., Nogueira, C., & Rodrigues, L. (2024). Passive data collection on Reddit: A practical approach. *Research Ethics*, *20*(3), 453-470. DOI: 10.1177/17470161231210542

Sharpe I, Davison CM. Climate change, climate-related disasters and mental disorder in low- and middle-income countries: a scoping review. *BMJ Open*, *11*(10), doi: 10.1136/bmjopen-2021-051908

Sibona, C., & Walczak, S. (2012). Purposive sampling on Twitter: A case study. In *45th Hawaii International Conference on System Sciences*, 3510-3519. doi: 10.1109/HICSS.2012.493

Siegrist, M., & Cvetkovich, G. (2001). Better negative than positive? Evidence of a bias for negative information about possible health dangers. *Risk Analysis, 21*(1), 199–206. https://doi.org/10.1111/0272-4332.211102

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, *48*, 813–838. DOI:10.1037//0022-3514.48.4.813

Sørheim, H. (2021). The relationship between household income, expenditure profiles, and C02 emissions in Norway. [Master's Thesis, University of Oslo] https://www.duo.uio.no/bitstream/handle/10852/87348/1/S-rheim-Hanna.pdf

Stillman, M., & Kruspe, A. (2024). Geolocation extraction from reddit text data. In *GeoExT 2024: Second International Workshop on Geographic Information Extraction from Texts at ECIR 2024*. https://ceur-ws.org/Vol-3683/paper2.pdf

Stoddart, M. C. J., Ramos, H., Foster, K., & Ylä-Anttila, T. (2023). Competing crises? Media coverage and framing of climate change during the COVID-19 pandemic. *Environmental Communication*, *17*(3), 276–292. https://doi.org/10.1080/17524032.2021.1969978

Taber, C. S., Cann, D., & Kucsova, S. (2009). The motivated processing of political arguments. *Political Behavior*, *31*(2), 137–155. https://doi.org/10.1007/s11109-008-9075-8

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychological Bulletin*, *110*(1), 67-85. https://doi.org/10.1037/0033-2909.110.1.67

Treen, K., Williams, H., O'Neill, S., & Coan, T. G. (2022). Discussion of climate change on reddit: Polarized discourse or deliberative debate? *Environmental Communication*, *16*(5), 680-698. https://doi.org/10.1080/17524032.2022.2050776

Trezza, D. (2023). To scrape or not to scrape, this is dilemma. The post-API scenario and implications on digital research. *Front Sociol. 8*. 1145038. doi: 10.3389/fsoc.2023.1145038

Verma, P., Shukla, N., & Shukla, A., P. (2021) Techniques of sarcasm detection: A review. In *International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 968-972. doi: 10.1109/ICACITE51222.2021.9404585.

Villanueva, I. I. (2021). *Climate Change Frames and Emotional Responses on Reddit.*

    [Master's Thesis, University of Arkansas, Fayetteville]

    https://scholarworks.uark.edu/etd/4076

Wiebe, J., Wilson, T., Bruce, R., Bell, M., & Martin, M. (2004). Learning subjective

    language. *Computational linguistics, 30*(3), 277-308.

    https://doi.org/10.1162/0891201041850885

Walter, S., Brüggemann, M., & Engesser, S. (2018). Echo chambers of denial: Explaining

    user comments on climate change. *Environmental Communication, 12*(2), 204-217.

    https://doi.org/10.1080/17524032.2017.1394893

Williams, H. T. P., McMurray, J. R., Kurz, T., & Lambert, F. H. (2015). Network analysis

    reveals open forums and echo chambers in social media discussions of climate change.

    Global Environmental Change, *32*, 126–138.

    https://doi.org/10.1016/j.gloenvcha.2015.03.006

Wright, J., (2024). Stakeholder management in change initiatives: reddit changes its api

    pricing. In *Sage Business Cases*. SAGE Publications, Ltd.,

    https://doi.org/10.4135/9781071941379

Yang, F.E., Pham, B., Zhang, S, & Kriss, L. (2024). Talking about pain plainly: Authentic

    expression of pain management on Reddit. *Health & New Media Research, 8*(1), 73-

    81. https://doi.org/10.22720/hnmr.2024.00038

Yi, J., Nasukawa, T., Bunescu, R., & Niblack, W. (2003*).* Sentiment analyzer: Extracting

    sentiments about a given topic using natural language processing techniques. In *Third*

    IEEE International Conference on Data Mining (ICDM), 427-434. doi:

    10.1109/ICDM.2003.1250949.

Zapcic, I., Fabbri, M., & Karandikar, S. (2023). Using Reddit as a source for recruiting

participants for in-depth and phenomenological research. *International Journal of

Qualitative Methods.* DOI:10.1177/16094069231162674

Zhao, B. (2017). Web Scraping. In L. A. Schintler & C. L. McNeely (Eds.), *Encyclopedia of

Big Data*, 1–3. Springer International Publishing. https://doi.org/10.1007/978-3-319-

32001-4_483-1

Appendix A

Wordclouds before pre-processing

Figure 6

Most frequently mentioned pre-processed words

Note. The left wordcloud displays the pre-processed words most frequently mentioned by climate change believers, while the right wordcloud displays the ones by the climate change deniers.

Figure 7

Most unique mentioned pre-processed words

Note. The left wordcloud displays the pre-processed words most frequently mentioned by climate change believers, while the right wordcloud displays the ones by the climate change deniers.

Appendix B

Topic Modelling

Topic Overview

While labelling the topics, their most prominently featured words were taken into consideration, hereby, featuring the words with the highest frequency-exclusivity (FREX) and the highest probability to be included in the topic (Ding, et al, 2020). Thereby, words with the highest FREX score were mainly considered to generate the label of a topic. However, when the FREX statistic was difficult to evaluate, the words with the highest probability score were considered. The generated topics, their label and the words that were used to create the labels are featured in Table 9.

Furthermore, topics that contained similar themes and words were grouped together and labelled. These topic groups can be seen as major discussion points that the climate change discussions of the collected comments contained. These topic groups and their corresponding topics are displayed in Table 10.

Table 9

Topic labels and Top Words

Topic Number	Topic Label	Distinctive words (FREX)	Highest probability words
1	Reddit Moderation and Interaction	rule, rconspiray, apply, meta, concerns, subreddit, bot	comment, please, keep, lol, rule, rconspiracy, apply
2	Climate Change Denial and Debate	wrong, deniers, change, deny, changing, denier, cult	climate, change, don't, point, wrong, evidence, doesnt
3	International Conflict and Politics	ukraine, deleted, Russia, war, israel, hamas, russian	one, that's, don't, youre, yeah, anything, maybe
4	Unidentifiable Topic	fucking, gonna, shit, really, thanks, heard, just	just, think, going, know, see, really, well

5	Government Policies and Societal Issues	children, capitalism, interests, socialism, protest, wef, legal	Government, world, theyre, state, public, got, pay
6	Electricity Production and Ecology	forests, forest, soil, wood, panels, burn, batteries	carbon, fossil, water, solar, fuels, fuel, use
7	Scientific Research	article, studies, chart, papers, ncei, reviewed, peer	science, data, article, scientific, research, study, read
8	Global Warming and its Consequences	cycles, activity, temperatures, warming, periods, correlation, graph	warming, global, temperature, human, past, temperatures, levels
9	Weather Events and Predictions	insurance, collapse, look, checkers, deaths, worse, scenario	will, time, every, look, weather, already, bad
10	Interaction with Media	religious, god, video, logical, religion, troll, narrative	believe, say, different, find, media, question, makes
11	Climate Change Effects (Change to "Seasonal Weather Patterns"?)	snow, polar, winter, summer, bears, days, florida	years, now, year, last, ice, ago, next
12	Energy Consumption and Sustainability	meat, eat, market, india, economy, buying, cows	energy, oil, power, emissions, money, stop, need
13	Increasing Atmospheric Pollution	radiation, atmosphere, vapor, gases, greenhouse, venus, stratosphere	earth, atmosphere, heat, effect, planet, greenhouse, water
14	Social Concerns and Public Opinion	care, people, kids, want, problem, app, let	people, will, can, make, want, get, problem
15	Unidentifiable Discussions	sounds, else, peoples, working, guys, everybody, truly	like, even, lot, much, everything, things, better

**Table 10.**

Topic groups and their corresponding topics

Group	Topic Number	Topic Label
Reddit Interactions	1	Reddit Moderation and Interaction
Interaction with Science and Media	2	Climate Change Denial and Debate
	7	Scientific Research
	10	Interaction with Media
Consequences of Climate Change	8	Global Warming and its Consequences
	9	Weather Events and Predictions
	11	Climate Change Effects
	13	Increasing Atmospheric Pollution
Government and Politics	3	International Conflict and Politics
	5	Government Policies and Societal Issues
	6	Electricity Production and Ecology
	12	Energy Consumption and Sustainability
Public Opinions and Personal Expressions	4	Unidentifiable Topic
	14	Social Concerns and Public Opinion
	15	Unidentifiable Discussions

Evaluation of the topic quality

After the creation and labelling of each topic, their semantic coherence and exclusivity were further evaluated. As can be seen in Figure 8, most of the topics score reasonably well when being compared with each other. Most notably, Topic 1, labelled as "Reddit Moderation and Interaction" scores the highest in both exclusivity and semantic coherence. In comparison, Topics 3, 5, 6, and 10 score lower than the majority of topics on exclusivity, while also scoring average to low on semantic coherence.

Figure 8

*Exclusivity and coherence of each topic*
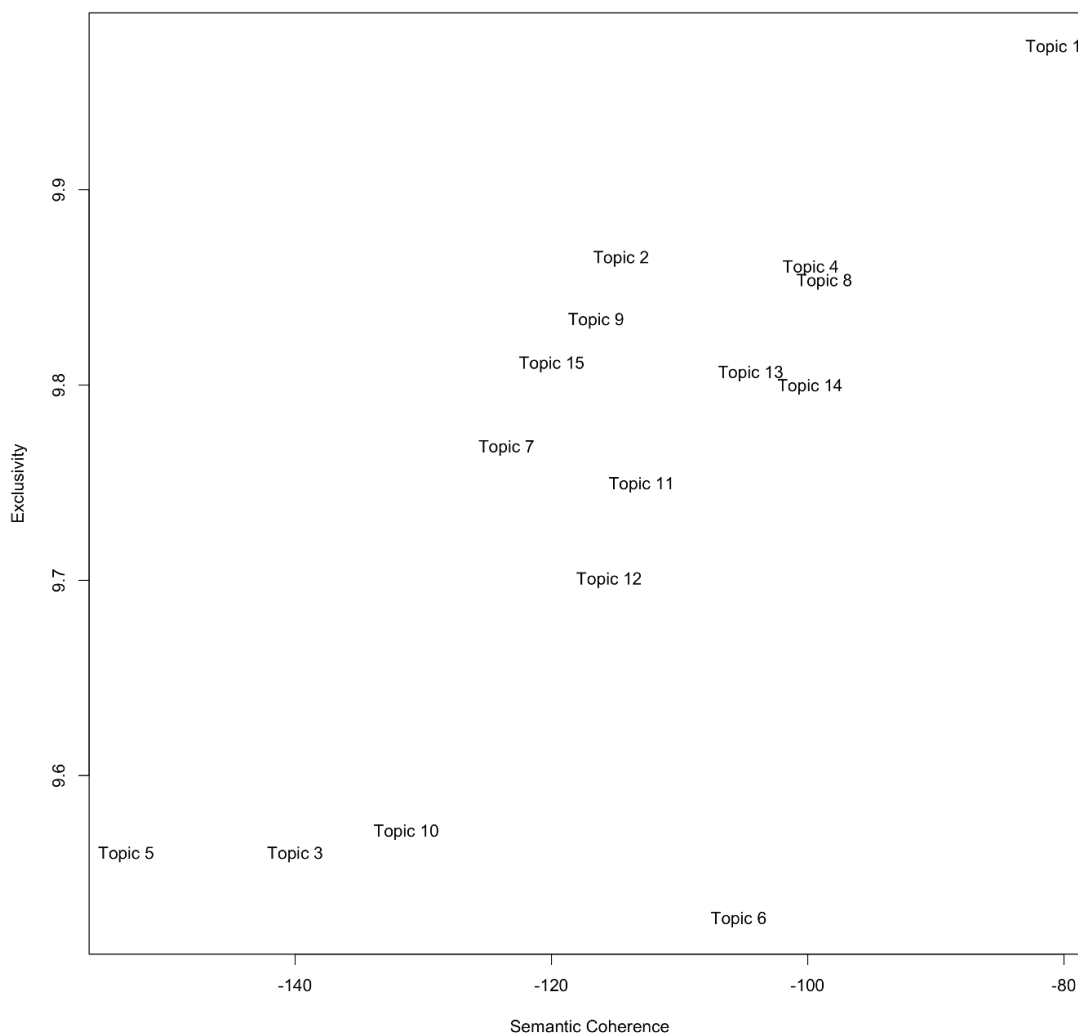
Topic proportions

  Figure 9 displays to what extent each topic is unique to either the group of climate change believers or deniers. Overall, the topic proportion is predominantly balanced between the content of believers and deniers. Notably, topics 7 and 12 are the most unique topics of climate change believers. Hereby, topic 7 is labelled as "Scientific Research" (FREX = article, studies, chart, papers, ncei; highest probability = science, data, article, scientific, research). Moreover, topic 12 resembles "Energy Consumption and Sustainability" (FREX = meat, eat, market, india, economy; highest probability = energy, oil, power, emissions, money, stop). Contrastingly, topic 14 is the most unique topic of climate change deniers and it resembles "Social Concerns and Public Opinion" (FREX = care, people, kids, want, problem; highest probability = people, will, can, make, want).

Figure 9

Topic proportions of deniers and believers

Topic Proportions of Deniers vs. Believers

Negative = Denier, Positive = Believer
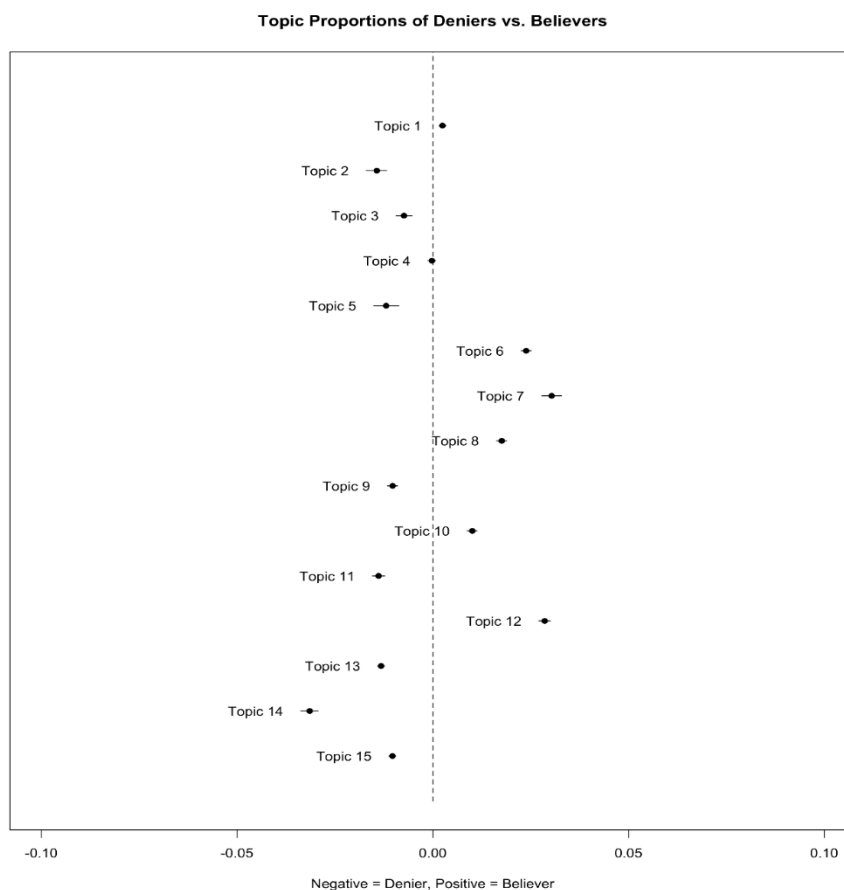
Topic prevalence over time

Figure 10 presents the prevalence of the topics throughout the time period from which the data was collected. Hereby, the x-axis resembles the number of days from the earliest data collected, which was on the 12th of April 2023, up until the 8th of April 2024, while the y-axis displays the topic prevalence in the corpus. The prevalence of the majority of topics over this time period remained stable. However, the prevalence of Topic 3 "International Conflict and Politics" spiked in between the around the 300 days mark. so around February to March 2024.

Figure 10

Topic prevalence over time