MAPPING DEMONSTRATIONS OF A SUSTAINABILITY POLICY CRISIS

ONLINE AND OFFLINE EVENTS DURING THE DUTCH NITROGEN CRISIS

Thesis submitted in fulfilment of the requirements for the degree of

MASTER OF SCIENCE

Delft University of Technology, Faculty of Technology, Policy and Management

by

Mila Hendrikse

4306146

To be defended in public on 18th of October 2022

FIRST SUPERVISOR	DR.IR. TRIVIK VERMA
SECOND SUPERVISOR	DR. JENNY LIEU
CHAIR	PROF.DR.MR.IR. NEELKE DOORN

Layout design by Ronny Lassooij

Acknowledgements

First of all, I would like to thank my supervisors: Trivik Verma, Jenny Lieu and Neelke Doorn. Jenny, we have been talking for over 1.5 years now, during which time you guided me through not only my second thesis but also my first thesis. Thank you for standing by my side. Not only did you supervise me, we had countless conversations about everything outside the thesis: our lives, thoughts, experiences, and work. I felt honoured that you felt you could open up to me and I learned from the stories you told. Also, it was extremely valuable to me to have someone to open up to and share experiences with during these tough times. I look forward to staying in touch after I leave the TU Delft.

Thank you Trivik, for making me feel seen and truly supervised from our very first meeting. I admire the space you created with CUSP to discuss societal problems, personal academic insecurities and the imperfections of the academic world. I think I do not only talk for myself when I say that CUSP feels like a space where academic taboos can be safely addressed and discussed. I met inspirational people through CUSP.

Even though we have not seen each other frequently, somehow every time I spoke to Neelke I ended up with an endless amount of energy and motivation by the end of the conversation. Trust me, energy and motivation can be extremely hard to come by during a thesis writing process that lasts over 1.5 years. Therefore, thank you, Neelke, for your passionate enthusiasm and for always asking spot-on questions that helped me sharpen my research process.

Although not an official supervisor of mine, I would like to thank Haiko van der Voort for always making time for a chat with EPA students, including myself, regardless of whether 'voor de gezelligheid'[1] or for discussing more serious matters. I find the way you open up to students inspiring. Also, thank you for the swiftness with which you helped me when I ran into administrative problems which could have taken a lot longer to solve without your help. That situation might have greatly impacted my study trajectory if not addressed properly.

I would like to thank my friends for their support. Thank you, Kim, for being my biggest inspiration when it comes to personal development and taking me on your journey with you. I want to thank the remainder of the 'thesis circle', Francien and Xander, for making countless days lively that would feel dull in their absence. Francien, I look forward to seeing your passion for gender equality take wing and to help change the world by your side. Thank you, Musaab, for your impressive ability to always provide a sympathetic ear. Thank you, also, for your contribution to this thesis in helping me navigate the confusing space where angry farmers and islamophobia somehow intersect.

Thank you, Rebecca and Francesca, my long-time friends, for making my life more colourful throughout the long and turbulent time we walked the campus of TU Delft as students together. I continue to learn from you about friendship, loyalty, warmth and acceptance. I consider myself lucky to call you my friends.

[1] Loose translation: 'for fun'

Last but absolutely not least, I would like to thank Monique and Gerard of the Service Desk at TPM for your warm treatment, your preparedness to help me overcome logistic problems and, most importantly, the nice chats that have brightened my time at TPM for 1.5 years.

Abstract

In May 2019, the Dutch Council of State rejected the national approach for reducing nitrogen emissions in Dutch nature. Farmers were targeted by the policy change: all licenses for agricultural expansion were revoked, affecting the financial livelihoods of farmers. Farmers did not take this well and, using social media platforms, started organising large-scale demonstrations.

When the demonstrations turned more aggressive over time, demonstrators started receiving increasing public criticism of their demonstration methods and saw a decrease in public support. The Nitrogen Crisis provoked many reactions and events in real life. However, the Nitrogen Crisis and the farmers' demonstrations were also extensively covered on Twitter and many demonstrative hashtags were born. As the farmers started a social movement during the Nitrogen Crisis and the crisis has sparked polarisation, in this research, I study the Nitrogen Crisis by comparing Twitter content to real-life events, answering the following research question: *How does the intensity of emotions and discussed topics on social media relate to the characteristics of real-life demonstrations during a sustainability policy crisis?*

To characterise online events, sentiment analysis and topic modelling are applied to analyse tweets about the nitrogen crisis to paint a picture of what subjects were discussed online, and how people are feeling about them. To characterise offline events, a codebook is developed with categories that reveal information on the scale, intensity and extremism during real-life demonstrations. Then, using newspaper articles as sources, this codebook is filled out. Comparing the insights from Twitter to the content of the codebook provides a picture of the similarities and differences between what discussions take place online and what demonstrations take place in real life. Also, the codebook provides context to the use of automated tools for social media data analysis and is used to evaluate these tools.

The codebook reveals that from the very beginning, the farmers' demonstrations were marked by intimidation, vandalism, and threats. It also shows that, although there are multiple farmers' organisations, the most radical one is mentioned most often as the organiser of demonstrations. In contrast, this organisation does not appear more than others in the calculated topic models. Additionally, the codebook shows demonstrations often take place in clusters and that news media coverage becomes less extensive after the first demonstrations.

The Twitter analysis shows that negativity is significantly higher than positivity during the whole crisis, throughout time and topics. The best performing and coherent topic model shows one topic, the topic depicting the farmers' demonstrations, that dominates when demonstrations from the codebook take place. Furthermore, the topic model shows no unexpected topics. However, when looking at less coherent topic models with significantly more topics, topics start appearing that show biases, political stances or describe smaller-scale events.

In terms of comparison, the results question what role sentiment analysis and topic modelling can play in studying policy crises and demonstrations. No patterns were found in

the outcomes of the automated methods that overlap with patterns found in the codebook. The most promising finding is that topic models with higher numbers of topics but lower coherence scores prove to be more interesting sources of information on underlying topics and trends on Twitter. Therefore, I propose to focus less on traditionally coherent topic models and to start exploring the hidden world uncovered by bigger and more messy topic models. The biggest challenge for this kind of research will be to create techniques for filtering out irrelevant topics and identifying the ones that exhibit bias or fine detail.

Before this study, studies on the Dutch nitrogen crisis were mostly empirical or based on interviews. The structural way of characterising demonstrations in a codebook will prove useful in expanding our understanding of protests during the nitrogen crisis and for future sustainability policy crises.

Table of Contents

List of Figures

List of Tables

List of Acronyms

BTM	Biterm Topic Modelling
CJEU	Court of Justice of the European Union
FDF	Farmers' Defence Force
LDA	Latent Dirichlet allocation
NT	Number of topics
PAS	Programma Aanpak Stikstof (Eng.: Integrated Approach to Nitrogen)
TU DELFT	Technical University of Delft

Preface

Writing this thesis has taken me on an interesting journey. Confronting bias was part of each step. I considered my own possible biases while searching literature on the topics and methods discussed and used in this thesis. When collecting newspaper articles as data sources, I considered the potential biases of news institutions. When interpreting the topics in my topic models, I was subjected to my own and my annotators' biases and background knowledge.

When I read about or saw videos of statements made by radical farmers, I was shocked by their words more than once. It became an interesting exercise to look at how these words were reported. For example, when an influential farmer compared the treatment of the Dutch farmers during the nitrogen crisis to the treatment of Jewish people during the Holocaust, there was understandably a national outcry and apologies were demanded by multiple parties. However, when that same farmer called politicians "Climate Salafists" in the same speech[2], it was not explicitly reported on by news media.

Only one source mentions the use of this term: in his paper, Van der Ploeg comments on a remark made by the protester and farmer Daniëlle Hekman that "Climate Salafists" are destroying the farmers (Van der Ploeg, 2020, p. 2). The use of the word 'Salafist' is remarkable for multiple reasons. First of all, it shows a deeply Islamophobic stance, comparing environmentalists to an Islamic religious branch. Second, it follows a current trend of throwing the Salafist Islamic movement and Salafi jihadism in one pot, even though the first is a religious branch while the second is a sub-current of Salafism that elicits an image of violence and certain hostility towards the western world. Ironically, the organisation the two mentioned farmers are members of, the Farmers' Defence Force (FDF), was later listed as a terrorist threat by the Dutch National Coordinator for Security and Counterterrorism (Bouwmeester, 2020). I have found no other sources criticising the use of this language even though some mention it (Hakkenes, 2019; Janssen, 2019; Pluimveebedrijf, 2019). One of these news articles describes the language used as 'harsh'[3], which I consider an understatement (Hakkenes, 2019).

While performing research, many choices need to be made. What to include, what not to include, what to use as a data source, what not, etc. While I did look at opinions in tweets, I did not have the time to further pursue analysing biases and framing in newspaper articles. I hope someone will one day have a critical look at how events are framed during a sustainability policy crisis and what it tells us about underlying cultural developments. In the meantime, I encourage the reader to think about their own biases while reading this research and, frankly, everything they will ever read.

This thesis was not written with a specific audience or client in mind. However, as with all research, the findings may be used by parties for their agenda.

[2] A video with the speech is available here:
https://www.youtube.com/watch?v=enaA_6wfS74&ab_channel=OmroepBrabant

[3] Original words: 'harde toon'. Translated by me.

1 | Introduction

1.1 The Dutch Nitrogen Crisis

Following a rejection by the Court of Justice of the European Union (CJEU) in 2018, the Dutch Council of State rejected the national nitrogen approach to lowering nitrogen emissions in Dutch nature (Natuurmonumenten, 2018; NOS, 2019). This approach was called Programma Aanpak Stikstof, or PAS, in Dutch. The PAS allowed expanding firms to get permits even if the expansions would result in more nitrogen being emitted, as long as there were preparations in place to compensate for these nitrogen emissions in the future. Despite warnings from previous cabinets about the policy's faults, the Dutch government proved to be entirely unprepared for the likelihood of EU rejection. The initial response was to cancel all emission rights permits, which halted roughly 18.000 construction and infrastructure projects. Farmers were heavily affected by the policy shift, as all agricultural expansion licenses were withdrawn. In quest of alternatives, policymakers in the Netherlands began discussing reducing cattle and large-scale farmer buyouts.

1.1.1 Initial response from farmers

Farmers were not pleased with this and began organising demonstrations (Leeuwarder Courant, 2019). The Hague, where the Dutch House of Representatives is located, saw the most demonstrations, although there were also rallies in front of several Province Houses throughout the Netherlands. Dozens of tractors blocked the busiest highways on their way to the Hague to occupy Malieveld[4]. The farmers' demonstrations appeared to be successful in a few areas at first. For starters, posts about the protests flooded social and mainstream media platforms, and

[4] A park in the centre of The Hague known for being the location for festivals and big demonstrations.

photos and videos of the event went viral, garnering widespread attention. Second, following the first demonstrations at their provincial houses, several provinces cancelled their local Nitrogen policies. (Van der Boon, 2019). Finally, officials appeared to take the farmers' displeasure seriously, showing an interest in speaking with and representing the protestors (Bosma & Peeren, 2021). Protesting farmers allowed some politicians to speak to the crowd on stage during rallies, and Prime Minister Rutte and Agriculture, Nature and Food Quality Minister Schouten invited leaders of farmers' organisations to meet and debate the Nitrogen Policy (Schelfaut, 2019b; Dagblad, 2021).

1.1.2 Policy proposals and reactions

A report with a long-term Nitrogen policy strategy was presented to the Dutch House of Representatives in 2020 by the Dutch advisory board for Nitrogen, tellingly titled "Not everything is possible everywhere"[5] (Remkes et al., 2020). Among other things, this proposal constituted the buyout of farms close to Natura2000[6] areas. Many farmers were unhappy about it and felt like they were not taken seriously. The months that followed were filled with debates in the Dutch House of Representatives and more farmers' demonstrations. Protests took place in October and December 2019, February and June 2020, and July 2021. Although the farmers' demonstrations seemed successful in attracting much public and political attention, dissatisfaction with each new proposal and how the Cabinet handles the Nitrogen Crisis kept resurfacing. Demonstrations got increasingly aggressive: heavy agriculture machinery was used to block roads, and break through fences and some politicians started receiving threats (Wijnants, 2020; Kos, 2020; Winterman, 2019). Criticism about these actions started to arise, and politicians expressed they felt less willing to talk to, and negotiate with, farmers (Hermenet, 2020).

1.1.3 FDF: Polarisation and radicalisation

Farmers Defense Force (FDF) was founded shortly before the start of the Nitrogen Crisis in reaction to climate and animal rights activists' occupations of farms. FDF members formed private WhatsApp and Facebook groups to assist and 'protect' fellow farmers when demonstrators arrived at their farms (Kalkhoven, 2021). During the Nitrogen Crisis, it became the most radical farmer organisation and seems to believe it needs to organise increasingly radical protests to keep pressuring the government to listen to their requests (Kalkhoven, 2021). FDF is known for making statements that have been interpreted as threats. Examples of such statements include:

1. *"The well-oiled machine that FDF has become is a threat to the status quo in The Hague and to the advocates who are at the mercy of the government. The Hague now knows*

[5] Original text: "Niet alles kan overal". Translated by the author.
[6] Protected natural areas in the Netherlands

that when the FDF fighters say action will be taken, action will be taken"[7] (Van Rooijen, 2020)

2. *"Now the Netherlands can see what the Ministry's position is. That there is no desire to change anything. If they wanted to talk to us, we would have been willing to extend the deadline of our ultimatum by a week, if necessary. Now we are going to take things further."*[8] (Van Rooijen, 2020)
3. *"We will accelerate if they do not start listening to us."*[9] (Driessen, 2019).

Although FDF gained many new members at the beginning of the Nitrogen Crisis, with the demonstrations turned more aggressive over time, demonstrators started receiving increasing public criticism of their demonstration methods and saw a decrease in public support (Cornelisse, 2020). Many opinions were expressed about whether the demonstrations are useful or not in attracting public support, and whether the way demonstrators behaved was fair and democratic. There have been incidences of vandalism, use of threatening language, and, as the above quote showed, an ultimatum was given to the government by FDF (AD, 2019; NRC, 2020; Winterman, 2019). In 2020, the Nationaal Coördinator Terrorismebestrijding en Veiligheid, the main Dutch counter-terrorism unit, reported that the excesses in the actions of FDF contribute to social polarisation around the theme of climate (Bouwmeester, 2020).

The Nitrogen Crisis provoked many reactions and events in real life. However, the Nitrogen Crisis and the farmers' demonstrations were also extensively covered on Twitter and many demonstrative hashtags were born, e.g. #boerenprotesten (Eng: farmers' demonstrations), #stikstofcrisis (Eng: Nitrogen Crisis). Also, demonstrative hashtags like #trotsopdeboeren (Eng: proud of the farmers) and #NoFarmersNoFood have been widely used.

1.2 Twitter as a platform for public opinion

Twitter can serve as an excellent platform for monitoring social dissatisfaction and unrest. Twitter is one of the most wide-used social media platforms worldwide. While Facebook traditionally focuses on connecting people and sharing updates on users' lives, Twitter is mostly used as a

[7] Original text: *"De geoliede machine die FDF geworden is, is een bedreiging voor de status quo in Den Haag én voor de belangenbehartigers die aan de tiet van de overheid liggen. Den Haag weet nu dat wanneer de FDF-strijders zeggen dat er actie komt, er actie komt".* Translated by Google Translate.

[8] Original text: *"Nu ziet heel Nederland hoe het ministerie erin staat. Dat er geen wil is om iets te veranderen. Als ze wel met ons hadden willen praten, hadden we het ultimatum nog wel willen verlengen met een week, als dat nodig is. Nu gaan we optoeren.".* Tranlated by the author. The word "optoeren" is not officially a Dutch word. As it consists of 'op' (Eng: 'up') and 'toeren' (refers to the English rpm of a motor), here it is translated as 'taking things further', interpreting the word as 'taking things up a notch' or 'accelerate'.

[9] Original text: *"Als ze niet naar ons luisteren, gaan wij meer gas geven".* Translated by the author.

platform for sharing opinions. Users do this through short messages, called Tweets, which have a limit of 280 characters each and can include media, like photos, videos and URLs as well.

There are advantages to analysing tweets to monitor public sentiment. First, in contrast to traditional public opinion polling methods like surveys and interviews, posting on and interacting with Twitter takes its users relatively little time. Since all messages are shorter than 280 characters, a Tweet with a thought, opinion or fact is written and sent out into the world easily and in no time. Second, extracting and analysing Twitter data requires no engagement with Twitter users. Third, Twitter is widely used as a platform for discussing and arguing about important topics on a large scale. These topics can include controversial policies or responses to global crises, like the Covid-19 pandemic. Last, because of the Covid-19 crisis, and its restrictions on social life and daily interactions, Twitter has recently seen a noteworthy increase in use (Miao, Last, & Litvak, 2020). The Covid-19 crisis and the accompanying increase in the use of Twitter starts a few months after the start of the ongoing Nitrogen Crisis. Last but not least, huge volumes of data are posted on Twitter on a daily, even hourly basis. Where polling methods take much effort and time to provide limited information, Tweets can be easily downloaded and used for large-scale data analysis.

The number of tweets that are posted to Twitter each day is too high to analyse manually. With the rise of the internet came the development of computational text analysis techniques, often referred to as Natural Language Processing (NLP) (Jurafsky & Martin, 2020). The two most common NLP tasks are sentiment analysis and topic modelling.

Sentiment analysis is the task of automatically calculating sentiment in a piece of text, and is one of the most rapidly growing research areas (Mäntylä, Graziotin, & Kuutila, 2018). *Topic modelling* is the procedure for automatically determining a set of topics from a dataset consisting of text documents. Sentiment analysis can provide information on how people feel, but not on *what* they are discussing. Topic modelling can provide information on what people are discussing on social media. Therefore, this study combines sentiment analysis and topic modelling, aiming to provide Twitter insights that can be compared to real-life demonstrations.

1.3 A codebook for demonstrations

As seen above, there have been instances of radicalisation and polarisation during the nitrogen crisis, but little research has been performed on the characteristics of these demonstrations and how they changed over time. In this research, I develop a codebook, where I document the characteristics of each demonstration. The combination of these characteristics is intended to provide an understanding of the scale, intensity, and extreme nature of the demonstrations. The content of the codebook will then enable comparison with Twitter analysis.

1.4 Research questions

Previous studies have shown that social media can provide an interesting data source for studying social movements (Burns & Eltham, 2009). Also, research has shown that social media platforms can contribute to polarisation (Grover, Kar, Dwivedi, & Janssen, 2019; Hong & Kim, 2016; C. Lee, Shin, & Hong, 2018; F. Lee, 2016). As the farmers started a social movement during the Nitrogen Crisis and the crisis has sparked polarisation, in this research, I study the Nitrogen Crisis by comparing Twitter content to real-life events. The Nitrogen Crisis is an example of a policy crisis that might occur when an EU country fails to implement timely sustainability policies and forces itself into crisis management instead of long term policy planning. This is why it was chosen as a case study. In the future, similar sustainability policy problems are likely to occur, as climate change begins to exert its influence and is increasingly recognised as a major issue (Pörtner et al., 2022). In answering the following research question, this research aims to look at the dynamics of online social media activity and real-life political and demonstrative events:

MQ

How does the intensity of emotions and discussed topics on social media relate to the characteristics of real-life demonstrations during a sustainability policy crisis?

The following sub-research question will be answered in tackling the main research question:

SQs

SQ 1 How can sentiment analysis be applied for studying online social media activity?

SQ 2 How can topic modelling be applied for studying online social media activity?

SQ 3 How can demonstrations be characterised?

This research applies a mixed-methods approach. The first two sub research questions examine how sentiment analysis and topic modelling, respectively, can be used to automatically generate information from the tweet dataset. To answer the third research question, information on the demonstrations during the nitrogen crisis is collected and characterised manually. In answering the third research question, I create information that is analogized to the sentiment analysis and topic modelling results. I look for patterns in the results of the automated methods and use the characterisation of the demonstrations to try and explain them. In doing this, I aim to evaluate sentiment analysis and topic modelling as tools for examining social media data during sustainable policy crisis.

Hypothesis

This research compares data acquired by hand with data collected using automated methods. I expect that the successful implementation of NLP approaches such as sentiment analysis and topic modelling will disclose information that will be beneficial to anyone trying to understand the protests and public discourse that took place throughout the crisis. For example, looking at when there were shifts in the general sentiment expressed in tweets could help pinpoint major events in the past and serve as an indication of public sentiment towards the nitrogen crisis and the farmers' demonstrations. I anticipate an initial positive sentiment in the dataset based on reports and news stories analysed during my prior research on the nitrogen crisis, around August and September 2021. Then, I expect negativity to increase over time as FDF becomes increasingly radicalised, public support for the demonstrations decreases and public discontent with the still-unresolved problem grows.

I am hoping that the generated topic models will highlight not just the major themes, but also the more niche ones. For example, I anticipate observing indications of political affiliation in topic model subjects: Since hashtags are a popular way to indicate support for a movement (think of #StemZeWeg, translation: VoteThemAway), I expect to find words that are commonly used with hashtags in the calculated subjects. I expect there will be an increasing number of topics that can be identified as expressing political affiliation in topic models with a higher number of topics. This is because I anticipate it to only uncover general overarching topics when the number of topics in generating a topic model is low. I expect that as the number of subjects increases, more unidentifiable topics will emerge, but that these topic models will also include several niche topics. These findings may be useful in revealing hidden rhetoric.

The sentiment analysis and topic modelling findings can be contrasted to the manually recognised and categorised information on the demonstrations. By doing this, the aforementioned hypothesis will be tested. Various outcomes could result from this. One possibility is that sentiment analysis could show patterns that are also found in the characterisation of demonstrations based on newspaper articles. This would hint that there is overlap between online and offline activity during a policy crisis, and incentivise the application of this automated method for future research on public reaction to policy crises. I am hoping that these new connections between the manually collected data and the results of the Twitter analysis will be discovered during this research. However, if no connection is found, I will discuss the potential of automated methods and what improvements can be made. In addition, if no connections are found I will pose questions about the relevance of both automated and non-automated methods for understanding public opinion during policy crises.

1.5 Structure of report

In Chapter 2: Related work, key concepts and methods are introduced and related research areas are discussed. Chapter 3: Methods describes the research approach of this study, how data is collected and analysed and how the steps in this research are evaluated. In 4: Results the main findings of this study are presented and reflected on. In Chapter 5: Discussion, limitations and future research the findings are interpreted and discussed in a larger context, after which the limitations are elaborated and recommendations for future work are presented. Chapter 6: Conclusion the research is summarised and the research questions are answered. The report ends with Chapter 7: Appendix and, finally, the list of references.

Even though the research questions and analysis are original, there are elements of this study that are either copied or strongly based on sections of my previous master thesis (Hendrikse, 2021), as this thesis expands on that work. The headings of these sections are marked with a footnote.

Key Findings of Chapter 1: Introduction

1. **Case**: Dutch Nitrogen Crisis
2. **Data**: Tweets and newspaper articles
3. **Methods**: Mixed methods approach: sentiment analysis, topic modelling and qualitative analysis
4. **Aim**: Compare real life demonstrations with sentiment and topics discussed online
5. **Main research question**: How does the intensity of emotions and discussed topics on social media relate to the characteristics of real-life demonstrations during a sustainability policy crisis?

2 Related work

This chapter first introduces the Dutch nitrogen crisis and research that evolves around it. Then, current knowledge on the interplay between social media platforms and politics is summarised. After, as the contribution of this research comprises the combination of sentiment analysis and topic modelling, both methods are introduced and their limitations are discussed. Finally, literature on the definition of demonstrations and their characteristics is discussed.

2.1 Case: The Dutch Nitrogen Crisis

When the European Court of Justice overturned the Dutch nitrogen emission policy, it drew widespread attention to the Dutch nitrogen problem for the first time. The Court found that the policy allows too much nitrogen emission into the environment. The following section begins by outlining the negative impacts of nitrogen emissions. Following that, it reviews existing literature on the Dutch nitrogen crisis, some of its findings, and research gaps that have yet to be filled.

2.1.1 The problem with nitrogen emissions

Nitrogen is the most important fertilizer for plant growth. Together with the other important fertilizers phosphorus and potassium, it plays an essential role in agriculture. However, the fertilizer that is used in farming land can spill into surface water and ground water. When spilled into marine ecosystems, it gives an advantage to algal blooms that will grow and suffocate other organisms (Stokstad, 2019). This deregulation of ecosystems and resulting death is called eutrophication.

Air is another route for nitrogen to leak into the ecosystem and cause harm. Power plants and car engines both emit nitrogen oxides. Even more nitrogen oxides vapour as ammonia from cattle manure. Both types of nitrogen leaks produce airborne particles that cause smog, vegetation damage, and soil acidification (Stokstad, 2019). These are direct threats to both non-human and human health and contribute to biodiversity loss and global warming.

The nitrogen problem cannot be considered new: the Ministry of Agriculture first published a policy paper on the topic in 1974. However, there has been resistance from both the government itself and farmers' unions to address it (Frouws, 1993; Hendrikse, 2021; Van der Ploeg, 2020).

2.1.2 Earlier studies on the Dutch nitrogen crisis

The Dutch nitrogen crisis started fairly recently, and only a few published academic studies, several unpublished studies and some non-academic reports on the topic are available. Van der Ploeg analyses the underlying reasons for the eruption of farmer rage during the crisis (Van der Ploeg, 2020). He argues that the farmers' movement ignores the political and environmental complexity of the crisis and the impacts of nitrogen emissions. In doing so, according to Van der Ploeg, the farmers start a "regressive populist" movement that fights for the preservation of the current state of affairs. Van der Ploeg also notes that while there is very little trust amongst farmers in governmental institutions and the organisations that represent agricultural interests, the farming community is deeply divided on how to address the sector's current issues (Van der Ploeg, 2020).

Interestingly, Van der Ploeg also comments on a remark made by FDF member Daniëlle Hekman that "Climate Salafists" are destroying the farmers (Van der Ploeg, 2020, p. 2). The use of the word 'Salafist' is remarkable for multiple reasons. First of all, it shows a deeply Islamophobic stance, comparing environmentalists to an Islamic religious branch. Second, it seems to follow a current trend of throwing the Salafist Islamic movement and Salafi jihadism in one pot, even though the first is a religious branch while the second is a sub-current of Salafism that elicits an image of violence and certain hostility towards the West (Drevon, 2016).

Another paper explores framings used during the crisis by farmers and representatives of the farmers (Bosma & Peeren, 2021). It describes xenophobic "us" and "them" thinking and the sexist and homophobic frame of the anger of the farmer being "righteous" since the farmer is a "real man" (Bosma & Peeren, 2021, pp. 9–10). Bosma and Peeren warn that "As long as the engrained idealised perception of farming and the rural that underlies these fantasies is not definitively dispelled, important rural policy discussions are easily derailed and populist narratives will continue to seem commonsensical to many" (Bosma & Peeren, 2021, p. 13).

Several Bachelor's and Master's theses have been written on the Dutch nitrogen crisis. In her bachelor thesis, Tessa de Weerd analysed how the feeling of societal unappreciation affected the demonstration participation amongst farmers through interviews (Weerd, 2021). Although she tries to find out whether farmers feel underappreciated, she notes she does not address the question as to whether this feeling is *legitimate*. She proposes that future studies examine both public opinion and sentiment toward the farmers as well as how the government responds to protests in policymaking (Weerd, 2021). Similarly, Rick Fuhler interviewed agricultural stakeholders for his Bachelor's thesis and identified that stakeholders in the agricultural sector do not feel properly represented in policymaking. Marin Visscher manually analysed 160 Dutch newspaper articles in search of frames and narratives applied to the nitrogen crisis and its stakeholders. It showed that the representation of actors differs significantly amongst newspapers.

Finally, last year I found the most extensive media analysis on the nitrogen crisis so far, which is a report by Lieuwe Kalkhoven commissioned by the Ministry of the Interior and Kingdom Relations (Kalkhoven, 2021). Apart from extensively discussing news media coverage of the nitrogen crisis, the most frequent framings and public opinion on the demonstrations, this is the only documented social media analysis of the crisis. Although the report is a rich source of information, it is an unofficial document that lacks a rigorous methodology[10]. Also, its social media analysis chapter is minimal: it includes only a plot of the volume of tweets per month and a word cloud.

One important aspect that all these papers and reports have in common is that they are *qualitative* and, in the cases of Van der Ploeg and Kalkhoven, empirical research. To start addressing this gap, I wrote a thesis about the application of sentiment analysis and topic modelling to Dutch tweets on the nitrogen crisis. The research aimed to provide insights for policymakers by finding what analysis outcomes could flag major events, like demonstrations, during the nitrogen crisis. Using sentiment analysis, the mean positivity and negativity of tweets were plotted over time. Complementarily, by calculating the difference between topic models of subsequent weeks, the difference between what topics were discussed every week was plotted. Surprisingly, it was not the increase or decrease in tweet sentiment nor in the topic difference that was the most significant indicator for major events. It was simply an increase in the volume of tweets. The most substantial recommendations of the research concerned the application of more state-of-the-art sentiment analysis and topic modelling techniques to the dataset. Additionally, I advised analysing the content of topic models in future research. This current research builds forth on the previous thesis by implementing the recommended improvements and by extending the research with a mixed methods approach. As far as I can find, no scientific quantitative research on the nitrogen crisis and its media coverage exists that predates my previous Master's theses on the topic.

[10] I wanted to ask Kalkhoven about his sources and methods, so I called the Ministry of the Interior and Kingdom Relations and managed to get in touch with him. He was surprised at my call and said that his report was an internal document and that the analysis was not based on academic methodology but on his scanning through many news articles. The report was not meant to be available to anyone outside the department. I wanted to see a list of the news articles he based his findings on to help define the events that took place during the crisis in this study. He, unfortunately, did not keep a list but sent me the notes he took during his analysis.

2.2 Analysing social media

2.2.1 Social media and polarisation[11]

Social media platforms are new and alternative communication platforms. They allow people all over the world to connect and see other users' content. With millions of users on these platforms attending them daily, social media have gained a prominent position in society. However, the exact societal relevance of social media is debated.

There are two prominent, yet contradicting, views on the effect of social media on its users (Hong & Kim, 2016). First, it is argued that social media platforms remove barriers between groups of people with differing political views. Based on sociological theory, exposure to differing political views and their supporting arguments can lead to higher political tolerance amongst users as it can increase understanding of underlying rationales (Mutz & Mondak, 2006). Following this rationale, social media could play a role in uniting people and contribute to peaceful societies.

On the other hand, other theories argue that social media form breeding grounds for isolated networks of users with similar, sometimes extremist, views. These so-called 'echo chambers' originate from users finding other users with similar political opinions. These connections are subsequently algorithmically enforced until users end up in a filter bubble which enforces their biases, cutting them off from users with opposing views. This bubble accommodates the radicalisation of political views and political polarisation (Benkler, Faris, & Roberts, 2018; Pariser, 2011; Sunstein, 2018). In the past years, more attention has been paid to this second theory.

<u>Studies on social media, online political activity and polarisation</u>

The political engagement on social media and its polarising potential have been widely studied. Already in 2010 a study showed that, in Germany, Twitter was widely used for discussing politics and political preferences (Tumasjan, Sprenger, Sandner, & Welpe, 2010). A study on the polarisation of social media has seen a significant increase after 'fake news' on social media was found to play a part in the 2016 US presidential election, (Allcott & Gentzkow, 2017; Looijenga, 2018; Waikhom & Goswami, 2019). An overwhelming number of these studies show there is indeed a correlation between the uprise of social media and societal ideological polarization, pointing towards the second echo chamber theory (Hong & Kim, 2016; Spohr, 2017). A study based on data from Hong Kong asserts that the polarizing effect of social media is intensified during periods of amplified political tensions (F. Lee, 2016). Examining the connection between extreme outings of politicians on Twitter and their Twitter readership, a study by Hong and Kim

[11] Partially adapted from Hendrikse (2021)

shows a strong increase in readership when Tweets are increasingly politically extreme (Hong & Kim, 2016).

It is clear there are connections between the global rise of social media and political engagement, yet the dynamics between what is said on social media and societal developments are only partially known. The following section discussed research on the relation between social media activity and offline political activity.

2.2.2 Hashtag activism: social media and offline political activity

There are various studies which investigate the connection between social media usage and offline participation in demonstrative events. The use of social media for activism is commonly referred to as 'hashtag activism'.

There are two theories in relation to hashtag activism on social media that sometimes lead to disagreements among researchers (Brantly, 2019; Howard, Agarwal, & Hussain, 2011; Shirky, 2011). The first argues that social media engagement stimulates offline political engagement. The second argues that social media engagement has a passivizing effect on offline political engagement. The following paragraphs will discuss both dynamics one by one.

Social media could play a promising role in political activism: "(…) the internet empowers traditionally excluded people with the tools to create their own spaces for self-expression, movement-building, and grassroots-organizing. For young people, particularly young people of color, the promise of digital technology for social justice work proves irresistible."(Fang, 2016, p. 139). Influence on offline political spaces can be exerted purely through online social media engagement: the earliest example of hashtag activism in the US was an online outcry against an Abercrombie & Fitch t-shirt line depicting stereotyping and racist cartoons of Asian caricatures (Fang, 2016). The online campaign prompted the pulling of the line by Abercrombie & Fitch (Fang, 2016).

Encountering injustices through social media can not only grow a user's dedication to a social cause but, interestingly, also to its accompanying offline demonstrative events (Smith, McGarty, & Thomas, 2018). A study from 2018 on social media usage in South Korea points out that the use of social media is correlated with an increased political engagement, both online and offline. It seems exposure to political content online might lead to an increase in offline political engagement, like having political discussions and voting (C. Lee et al., 2018). Not only can social media engagement influence the views of its users, but social media platforms themselves can play a critical logistic role in the organisation of political movements and demonstrative events. A study covering the Egyptian revolution identified multiple functionalities of twitter, like how the expression of solidarity online can stimulate protesters and how twitter can be used for sharing live information from "on the ground" during demonstrations (Starbird & Palen, 2012).

However, although there are various examples of successes of political social media engagement and hashtag activism for political activism, the possible downsides should not be ignored. One study suggests that hashtag activism can persuade social media users to think they are advancing a movement more than they are (Harlow & Guo, 2014). Because these users believe they have already participated through their online engagements, some may become more inactive in the physical space than they would otherwise be, opting out of offline political engagements and demonstrations. Another study points out the danger of viewing twitter as a platform for creating your "own spaces for self-expression, movement-building, and grassroots-organizing", as cited above (Fang, 2016; Ofori-Parku & Moscato, 2018). This study shows that activists may perceive Twitter as a platform that is independent of political dynamics offline (Ofori-Parku & Moscato, 2018). However, institutional norms, local politics, and contextual realities can significantly influence the framing of societal discourse online. In their coverage of social movements, traditional media will use framing that may not align with the discourse intended by activists. Activists need to overcome this discrepancy between the perception of agency over their social movement discourse and the power of institutions to shape this discourse in order to realise the full potential of hashtag activism.

In this research, I will look at both online engagement (through tweets) and offline engagement (through demonstrations) with the nitrogen crisis. By looking at the patterns that might arise from this, I will shed light on the two dynamics and how they relate to each other throughout the nitrogen crisis.

2.3 Sentiment analysis[12]

2.3.1 Sentiment analysis and Twitter

In the early years of sentiment analysis, the most researched datasets comprised online reviews, for example, for movies or, more interestingly for commercial purposes, of products. Nowadays, though, the vast majority of sentiment analysis studies focus on social media platforms like Twitter and Facebook (Mäntylä et al., 2018). Sentiment analysis can reveal information on positive or negative attitudes towards, for example, support of social movements, political parties or events. In 2018, one of the three most cited papers in Scopus and Google Scholar on sentiment analysis examine whether sentiment in tweets can help predict election results in Germany (Mäntylä et al., 2018; Tumasjan et al., 2010). Another such study, looking at the popularity of Italian political leaders in 2011 and the voting intention of French Internet users in the 2012 presidential ballot, found the results of their analysis showed a high correlation with the data of more official mass surveys. Additionally, they found their analysis showed predictive ability for the election outcomes (Ceron, Curini, Iacus, & Porro, 2014). This shows the potential of applying sentiment analysis to social media for characterizing the attitudes of its users.

[12] Adapted from (Hendrikse, 2021)

2.3.2 Challenges of applying sentiment analysis on social media

Although the application of sentiment analysis on social media data, like Twitter, is very promising, it can be difficult and there are important challenges to consider (Poria, Hazarika, Majumder, & Mihalcea, 2020; Zhang, Xu, & Jiang, 2018). Tweets often contain slang, misspelt words and overflow with emojis, and are, of course, very short because of character limits. This can make it difficult for the sentiment analysis tool to correctly score the sentiment of the Tweet (Giachanou & Crestani, 2016; Kim, Weber, Wei, & Oh, 2014; Zhang et al., 2018). Other great challenges to sentiment analysis which may lead to misclassification are sarcasm and lack of context (González-Ibáñez, Muresan, & Wacholder, 2011). To date, no perfect solutions have been found, which complicates the interpretation of insights from applying sentiment analysis to social media data. Finally, there is no standardized evaluation protocol, which makes it hard to determine what the most advanced implementations of sentiment analysis are. To apply sentiment analysis, researchers need to either train their own sentiment analysis models, which is time consuming, or carefully review models available and decide what model is best fitted for the task at hand.

Although various sentiment analysis approaches give different results on different scales, most tools return the sentiment scores on a linear scale from negative to positive (Mäntylä et al., 2018). However, there are newer approaches to sentiment analysis which consider, for example, that a message can contain both positive and negative sentiment at the same time and return two scores, or go further than classifying emotions as positive and negative, but try to distinguish between various specific emotions (Cambria, Gastaldo, & Bisio, 2015; Thelwall, Buckley, Paltoglou, & Kappas, 2010).

2.4 Topic modelling[13]

Topic modelling is a procedure for automatically calculating a set of topics from a dataset comprising text. To understand topic modelling it is essential to understand the following terms (Vayansky & Kumar, 2020):

- *Corpus*: the dataset, consisting of documents
- *Word*: each unique word in a corpus is indexed
- *Document*: a set of words, in bag-of-words representation, representing
- *Topic*: distribution of a pre-set vocabulary

The idea behind topic modelling is to calculate what topics are present in the corpus. The words and their probabilities of belonging to a certain topic are calculated based on the

[13] Adapted from (Hendrikse, 2021)

cooccurrences of words in the documents. This way, it is possible that a word belongs to multiple topics with varying probabilities. A corpus can consist of, for example, movie reviews, journal articles or, in the case of this study, tweets (Vayansky & Kumar, 2020).

2.4.1 LDA

The most popular algorithm for topic modelling is Latent Dirichlet Allocation (LDA). Originally developed by Blei et al. in 2003, LDA is a topic modelling algorithm that takes as input a corpus, a vocabulary matrix β (beta) and a parameter for the number of topics that it should find (Blei, Ng, & Jordan, 2003). It returns the word prevalence distributions per topic (words that belong to the topic and the likelihood they belong to that topic, see Figure 2.1 for an example) and the topic prevalence distribution per document (topics that belong to the document and the likelihood they belong to that document) (Blei, Ng, & Jordan, Latent Dirichlet allocation, 2003). This is done through an iterative process where initially each word in the vocabulary matrix is arbitrarily appointed to one of the topics, and the distributions are randomly assigned. Then, for each word per document, both the chance that the word's assigned topic belongs in that document and the chance of that word actually belonging to its assigned topic is calculated. Based on these outcomes, the distributions are updated and the process continues until the algorithm converges (Liu, Tang, Dong, & Yao, 2016). Unlike sentiment analysis, LDA, being a statistical approach to NLP, is language-independent: as long as the corpus comprises documents in the same language, topic modelling can be applied to corpuses of any language. With topic modelling, you can calculate what the distribution is of certain topics over the whole corpus, or look at each document individually and identify the distribution of topics present in it

For a more extensive explanation of LDA and its details, the following publications are recommended: (Blei, Ng, & Jordan, Latent Dirichlet allocation, 2003; Roberts, Stewart, & Airoldi, A Model of Text for Experimentation in the Social Sciences, 2016; Liu, Tang, Dong, & Yao, 2016).

```
------ Topic 1 ------
Stikstofcrisis(0.486) boer(0.404) snelheid(0.290) kabinet(0.183) verminderen(0.067)...

------ Topic 2 ------
boer (0.45) landbouw(0.394) subsidie(0.276) maatregel(0.151) negeren (0.090)...

------ Topic 3 ------
boerenprotest(0.427) crisis (0.357) denhaag(0.237) pfa(0.138) groot(0.012)...
```

Figure 2.1 | Example of three topics and their word prevalence distributions.

2.4.2 Biterm Topic Modelling

Although LDA often provides good results, there are challenges to applying it to short texts, like tweets (Steinskog, Therkelsen, & Gambäck, 2017). A study comparing the content of tweets to a traditional news medium (the New York Times) found that the standard LDA model performs poorly on tweets (Zhao et al., 2011). Multiple studies have focussed on improving LDA results, for example by clustering tweets into bigger pieces of text (tweet pooling) and using hashtags for automatic labelling (Luyi & Wei Song, 2016; Mehrotra, Sanner, Buntine, & Xie, 2013; Zhao, et al., 2011). These extensions to the classic LDA algorithm often lead to better results. Additionally, LDA-inspired topic modelling algorithms have been developed. One such method is Biterm Topic Modelling (BTM): a topic modelling algorithm specifically designed for short texts (Xiaohui Yan, Jiafeng Guo, Yanyan Lan, 2013). There are various comparative analyses where BTM outperforms LDA and other topic modelling algorithms when applied to short texts (Guzman, Ibrahim, & Glinz, 2017; Jonsson & Stolee, 2016; Xiaohui Yan, Jiafeng Guo, Yanyan Lan, 2013). Although BTM does not appear to have been applied to tweets about a policy crisis or demonstrations yet, a study from 2016 applied BTM to tweets about typhoons in the Philippines to unravel typhoon-related tweets.

2.4.3 Challenges of topic modelling

One of the biggest downsides of BTM and classic LDA is that the generated topics are not labelled, hence human interpreters are needed to label topics. There are various studies on automatic labelling of topics (Allahyari, Pouriyeh, Kochut, & Arabnia, 2017; Mehrotra, Sanner, Buntine, & Xie, 2013). However, they have mixed results and some require external data. Because of this, I rely on manual labelling in this research.

2.5 Demonstrations

2.5.1 Why do people attend protests?

Demonstrations are a tool for expressing political dissatisfaction. While voting and putting oneself up for election are seen as conventional forms of political participation, demonstrating can be conceived as unconventional political participation (Klandermans, van Stekelenburg, & van der Toorn, 2008). While it is not easy to define what a demonstration is and consists of, many definitions exist (Hutter, 2014). Jesus Casquete describes demonstrations as: "collective gatherings in a public space whose aim it is to exert political, social and/or cultural influence on authorities, public opinion and participants through the disciplined and peaceful expression of an opinion or demand" (Casquete, 2006, p. 47). Because not every demonstration that took place during the Nitrogen was fully peaceful, I use Casquete's definition but without the word 'peaceful'.

There have been many theories over the years on why people attend demonstrations. Early research focussed on grievances as the main driver for participation (Grasso & Giugni, 2016), Here, a grievance is defined as “a sense of indignation about the way authorities are treating a social or political problem” (Klandermans et al., 2008, p. 993). Building on this, relative deprivation theory focuses on how people are inclined to demonstrate when they feel they are treated unfairly compared to what they expect to receive or compared to what others get (Klandermans et al., 2008). Claims of relative deprivation are made during the Dutch nitrogen crisis, for example when farmers claim they receive ‘unequal treatment’ and that they deserve more ‘respect’ (Van der Ploeg, 2020, p. 590).

However, as relative deprivation is a feeling generally felt by demonstrators, more recent research on the incentives of demonstrators focuses on other components as predictors for demonstration attendance, asking not whether demonstrators feel aggrieved, but *why* some aggrieved citizens mobilise to demonstrate while others do not. Klandermans claimed in 1997 that people participate in demonstrations when they feel the demonstration can change the situation, policy or perceived injustice they object (Klandermans et al., 1997). Resource mobilisation theory argues in line with this reasoning, claiming that elements such as educational background, wealth and type of employment are predictors for demonstration participation (Verba, Schlozman, & Brady, 1995). If material resources, like tractors, are included in this definition of resources, this theory could explain why the farmers received much more attention and continued demonstrating longer than the construction workers who demonstrated the nitrogen policy alongside the farmers during the first demonstrations, but whose engagement, in contrast to the farmers, faded thereafter. Some sources claim the use of tractors by the farmers is a message of the power and influence of the agricultural sector on the Dutch government and citizens (Kalkhoven, 2021; NOS, 2019b). Finally, from a completely different angle, other studies claim that people are most likely to participate in demonstrations simply when they hear of other attendees in their social circles (Schussman & Soule, 2005; Tufekci & Wilson, 2012). While most studies on demonstrations test the above mobilisation factors individually for their influence on demonstration participation, Grasso and Giugni argue that they are likely interrelated and should be studied as such (Grasso & Giugni, 2016).

On the one hand, the farmers' demonstrations seem to have a lot in common with the social justice movements discussed. The farmers protest a perceived unjust treatment by the government by organising demonstrations bottom up. On the other hand, the polarization amongst the farmers, the rise of a radical farmer organisation, and the increase in the extremity of demonstrations and threats by farmers are events that show some overlap with patterns found in the literature on radical group behaviour, violence and terrorism (Tausch et al., 2011). To be clear, no claim is made here that farmers are a social justice movement, nor that the most radical farmers are terrorists. In all likeliness, the farmers' movement is something in between.

It is hard, if not impossible, to find literature where the differences and similarities of different types of political movement and engagement are examined. The next section looks at what definitions for demonstrations exist in literature and will form the basis for the characterisation framework for demonstrative events in this thesis.

2.5.2 Characterising demonstrations

In this study, I compare various demonstrative events which took place during the nitrogen crisis. To do this, I create a codebook to extract information from news articles and define categories that, when combined, characterise a demonstration. As mentioned above, although many definitions of demonstrations exist, there is surprisingly little literature that defines the characterisation of demonstrations: how can you characterise a demonstration? How big is a collective gathering from Casquete´s definition, and what defines a 'public space'? In search for characteristics of demonstrations, I found only one paper with such a list, by Olivier Fillieule (Fillieule, 2012). According to him, a street demonstration comprises four elements:

1. the temporary occupation of open physical spaces
2. expressivity
3. number of participants
4. political nature of the demonstration

Here, a *physical space* can be public as well as private. The term '*expressivity*' refers to the crowd's uniformity: whether there is a 'unifying principle' (Fillieule, 2012, p. 235). Although the *number of participants* is mentioned as an element of what defines a street demonstration, Fillieule himself notes that as there is "no means of sociologically determining the minimum number of individuals likely to act collectively, it is useless to set an arbitrary threshold" (Fillieule, 2012, p. 235). Lastly, the *political nature of the demonstration* refers to "the expression of demands of a political or social nature" (Fillieule, 2012, p. 235). In section 3.3.3: 'Characterising demonstrations' I explain how I based the list of characteristics of demonstrations in this research on Fillieule elements.

Key Findings of Chapter 2: Related Work:

Gap 1: Lack of research on public opinion and sentiment toward the Dutch nitrogen crisis

Gap 2: Lack of quantitative research on the nitrogen crisis

Gap 3: The dynamics between social media engagement and societal developments are only partially known

Gap 4: Lack of studies where various demonstrations are characterised and compared

3 Methodology

This research consists of two parallel parts. On the one hand, tweets about the nitrogen crisis are analysed to find out how emotional the tweets were and what topics were discussed. On the other hand, news articles are collected and their information synthesised into a codebook with an overview of the demonstrations that took place during the nitrogen crisis. These outcomes are compared to each other to evaluate how useful these automated NLP methods can be during a sustainability crisis.

This chapter discusses every step taken in this research. Figure 3.1 below shows an overview of the methodology of this research. On the left, we start with data, described in section 3.1: Data collection. The preparation of data is discussed in section 3.2: Pre-processing: preparing tweets for analysis. In the middle, Figure 3.1 shows the methods used, described in section 3.3: Content analysis 3.2: Pre-processing: preparing tweets for analysis. Finally, the evaluation of the applied methods is discussed in section 3.4: Evaluation.

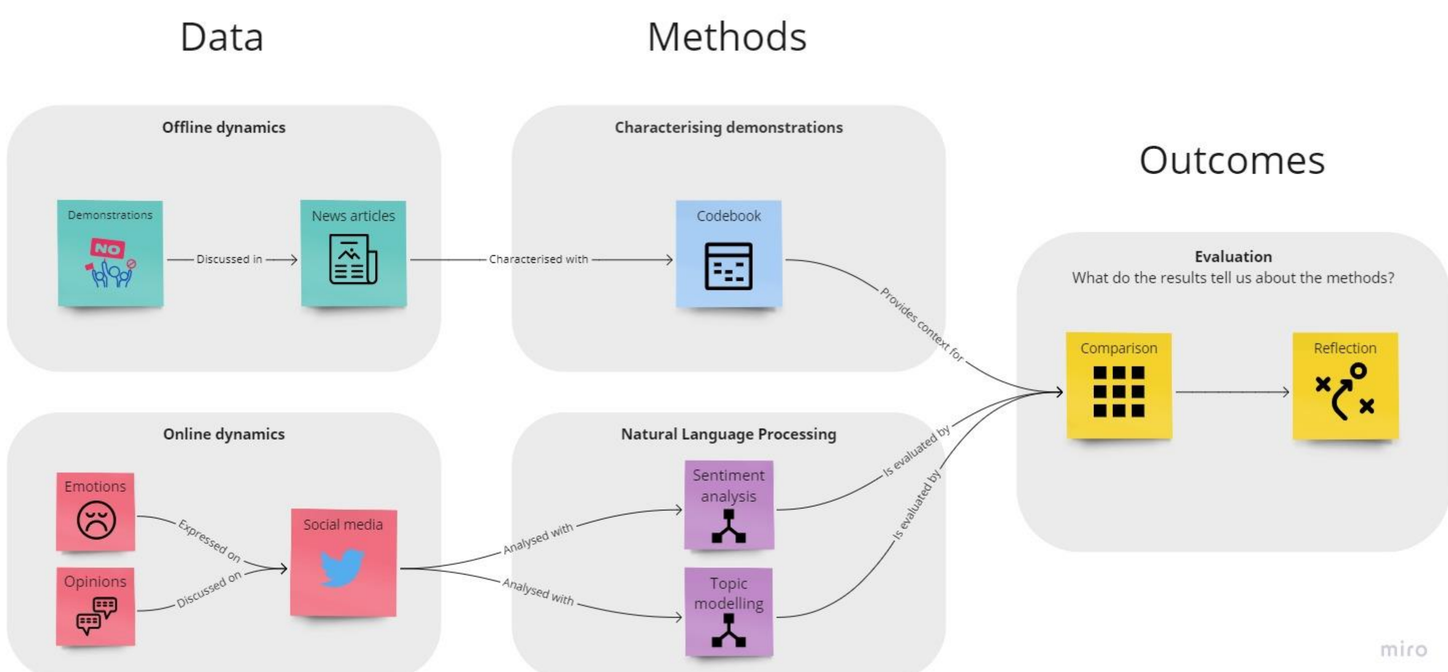

Figure 3.1 : Overview of methods

Figure 3.2 below shows a research flow diagram of the analysis of tweets in this research. From top to bottom, it shows how the data are treated. This process consists of the collection of tweets (top layer), pre-processing (purple layer), content analysis (pink layer), evaluation and quality improvement of the applied models (orange layer) to plotting the results (yellow layer). Each individual technical step is described in detail in this chapter.

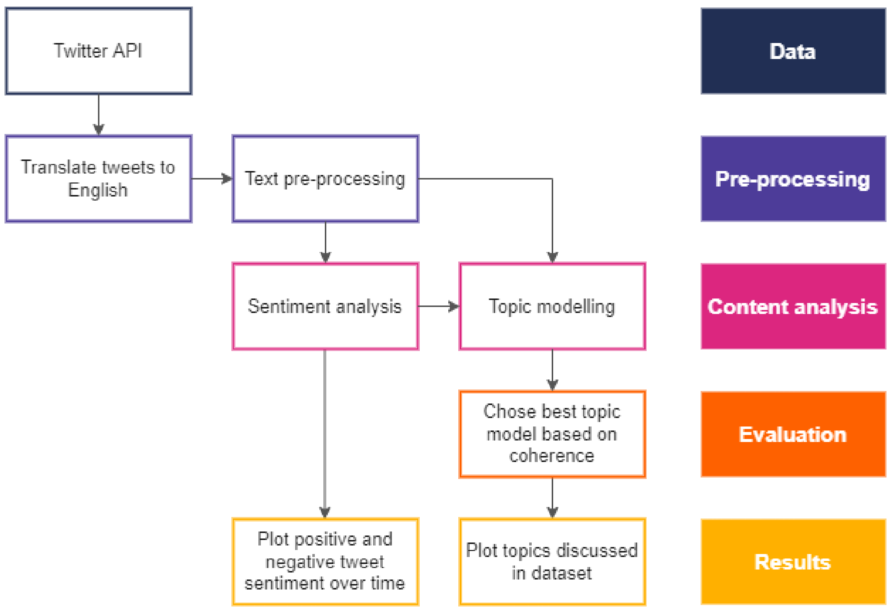

Figure 3.2 | Research flow diagram, showing how information flows through the steps in this research.

3.1 Data collection

For this research, two types of data are collected: tweets, for analysing online events, and newspaper articles, for analysing offline events. The next sections describe how these are collected.

3.1.1 Desk research: Identifying demonstrative events

Desk research was performed to identify the events during the Nitrogen crisis that received a lot of attention from the news media. News articles were the primary source of information. However, (news) sources with an overview of all major events throughout the whole Nitrogen Crisis were scarce. Some news sources had a special page on their website with a collection of their news articles on the Nitrogen Crisis, but these would only contain news articles from 2019, while the Nitrogen Crisis is still ongoing at the time of writing. After a long search, it turned out that Wikipedia had the most complete, well-sourced and comprehensive overview of the Nitrogen Crisis and the demonstrations that took place (Wikipedia, n.d.). Therefore, this page was taken as the main source for identifying important events. However, when looking up news articles about these events, they sometimes mentioned other events that were not documented on the Wikipedia page. If I could find multiple articles about these events, they would then also be included in this list of events. In reading news articles, I found out about more demonstrations through snowballing.

It is important to note that the list of events is by no means extensive. Due to time constraints, at some point I had to stop adding events I would come across in other articles.

3.1.2 Twitter API[14]

After some experimentation with various data collection methods, the Twitter Developer Portal was used for the collection of Tweets. In order to request data from the Portal for non-commercial purposes, a researcher needs to fill in an application and explain what the Portal will be used for and what the research is about. If this request gets approved, one gets access to the 'Standard product track', one of various types of Portal access with each different rate limits.

Upon being granted access, a Bearer token was generated and a Python script with a query was written to retrieve tweets from the past using the Twitter API. In order not to exceed Twitter API rate limits, the script only does 299 requests per 15 minutes[15]. The parameters inserted define the timeframe, type of tweets and the key words used to gather tweets. Keyword parameters were found through a process of trial and error: starting with the words

[14] Adapted from (Hendrikse, 2021)
[15] https://developer.twitter.com/en/docs/rate-limits

"stikstof" (Eng.: Nitrogen) and "Stikstof Crisis" (Eng.: Nitrogen Crisis) all tweets were collected that included either of these terms. Looking at the resulting tweets with the aim of extracting more terms associated with the Dutch Nitrogen Crisis, a list of terms was created with keywords, as seen in Table 3.1. The `OR` operator is used to collect tweets that contain either one or more of the keywords in the list. When keywords are put between quotation marks, only tweets containing the keywords in that exact same order are returned. When keywords are between brackets, Tweets are returned when all the keywords between the brackets are present in the tweet, regardless of the sequence of the words. As is visible in Table 3.1, many keywords are in the query twice: once as a word and once as a hashtag. This did not apply to all keywords though, as for example the word "pas", apart from being an abbreviation for the Dutch Nitrogen policy, is also a very common Dutch word with many meanings (e.g., "step", "just now", "pass", "only", "just" etc.). This is why the word "pas" is only included in the query when it is preceded by a hashtag: when "#PAS" is present in a tweet it will very likely be about the Nitrogen Crisis. Similarly, the word "stikstof" (Eng.: Nitrogen) on its own would result in too many tweets about, amongst others, biological processes that have nothing to do with the Crisis.

In order to create this query, many decisions had to be made. Only tweets that are not retweets are collected. This is because the inclusion of retweets could lead to many duplicates. Next, there is the time limit. Following the removal of the first few months with few tweets, the dataset is extended until July 2021. More recent tweets were purposefully not added to the dataset, because this research expands on my previous Master's thesis (Hendrikse, 2021). The datasets used in this and the preceding Master's thesis are identical. As a result, outcomes of both theses can be compared and data collection time is saved. This time is spent on a more thorough qualitative analysis of the demonstrations that took place during the nitrogen crisis, described in section 3.1.1. Finally, the actuality of the tweets is not relevant to the outcome of this study.

Table 3.1 | Query parameters for the retrieval of the dataset.

PARAMETER	VALUE INSERTED		
Start_time	2019-01-01T00:00:00Z		
End_time	2021-07-31T00:00:00Z		
Language	nl		
type	all but retweets		
Exclude	Mercosur OR #mercosur		
Include	stikstofcrisis OR PFAS OR #PFAS OR boerenprotest OR boerenprotesten OR #boerenprotest OR #boerenprotesten OR bouwprotest OR #bouwprotest OR grondinverzet OR grondverzet OR #stikstof OR #stikstofcrisis OR #PAS OR	"programma aanpak stikstof" OR "programma aanpak #stikstof" OR #stikstofbudget OR stikstofbeleid OR #stikstofbeleid OR (stikstof uitstoot) OR (stikstof crisis) OR (stikstof uitspraak) OR (stikstof probleem) OR stikstofdepositie OR #stikstofdepositie OR #grondinverzet OR #grondverzet OR #boerenterreur OR	(boeren terreur) OR #stikstofdebat OR stikstofdebat OR stikstofprobleem OR #stikstofprobleem OR stikstofgedoe OR #stikstofgedoe OR (stikstof gedoe) OR farmersdefenceforce OR #farmersdefenceforce "Stikstof-probleem" OR stikstofbudget OR

Table 3.2 shows the number of nitrogen tweets per month (see Figure 7.1 in Appendix A for a plot). Because up until June 2019 the number of tweets per month is lower than 1000 tweets per month these are removed preventatively (as indicated by the striped background), so these low numbers will not skew the sentiment analysis and topic modelling results later. In total, 1636 tweets are deleted.

Table 3.2 | Number of tweets per month in the dataset.

YEAR	MONTH	NUMBER OF TWEETS
2019	January	175
	February	336
	March	217
	April	205
	May	703
	June	2034
	July	1331
	August	1526
	September	15224
	October	85449
	November	31824
	December	29036
2020	January	8079
	February	15422
	March	6005
	April	4415
	May	2669
	June	4604
	July	23810
	August	6398
	September	4202
	October	5719
	November	11594
	December	6213
2021	January	2555
	February	2129
	March	3975
	April	2055
	May	3141
	June	4933
	July	7903

3.2 Pre-processing: preparing tweets for analysis[16]

To use text as input for machine learning algorithms multiple cleaning and pre-processing steps are required. Figure 3.3 shows an overview of my approach. The next sections describe these steps to prepare tweets for topic modelling.

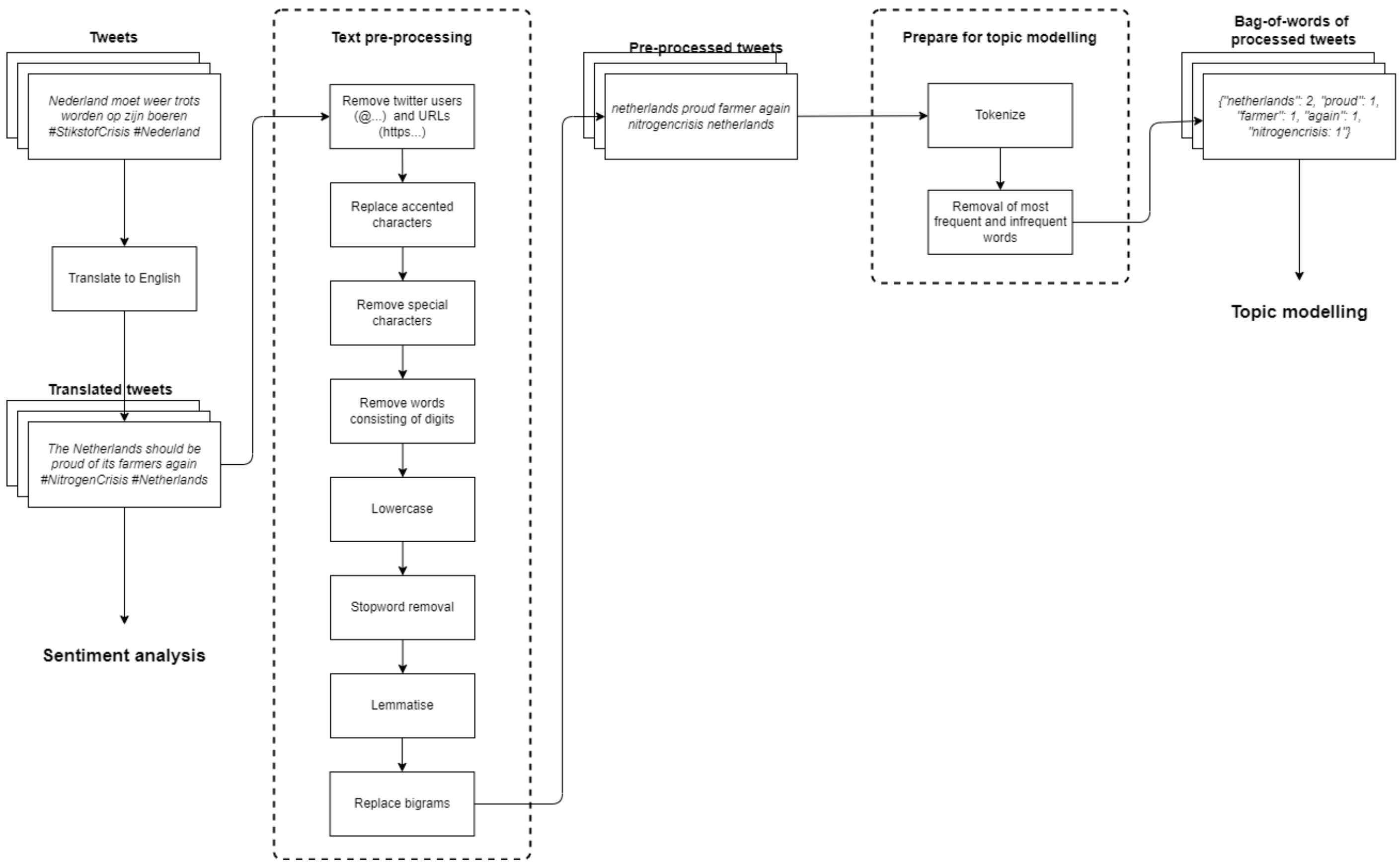

Figure 3.3 | Steps in pre-processing tweets.

3.2.1 Translation

The first column on the left in Figure 3.3 above shows that before being processed for sentiment analysis and topic modelling, tweets are first translated to English. The most important reason for this is that sentiment analysis tools in English are the most advanced sentiment analysis tools compared to other languages, including Dutch (Araujo, Pereira, Reis, & Benevenuto, 2016; Balahur & Turchi, 2012; Hendrikse, 2021). Even though some information is 'lost in translation' when (automatically) translating tweets, there are various studies where the increase in

[16] Adapted from (Hendrikse, 2021)

sentiment analysis quality is more significant than this loss of information (Jusoh & Alfawareh, 2011). For the translation, the Google Translate API is used via the deep-translator Python package developed by Nidhal Baccouri.

The decision of whether or not to translate the tweets before applying topic modelling is determined separately from sentiment analysis. Biterm topic modelling is a language-independent topic modelling algorithm (Xiaohui Yan, Jiafeng Guo, Yanyan Lan, 2013). Topic modelling could be applied to Dutch tweets, and the resulting topic models would be in Dutch. This study, however, strives to be comprehensible to supervisors and researchers who do not speak Dutch. As a result, the list of individual words that make up a topic would need to be translated into English before being included in this report. Yet, sentence translation is more accurate than the translation of contextless individual words (Jusoh & Alfawareh, 2011). Consequently, the translated topics may have lost more information than if the tweets had been translated in their entirety prior to topic modelling. To preserve the interpretability of topic models in this research for non-Dutch speakers, topic models are trained on tweets that are translated to English.

Table 3.3 shows nine examples of tweets and their English translation. Overall, it performs well and deals well with the challenging language used in tweets. For example, tweet 3 contains "naaaaaaaar" which is meant as a 'screamy' version of 'naar' (translation: 'to'). GoogleTranslator seems to pick up on the meaning of the word, translating it to 'to' but not ignoring the way the word was spelled, adding 'aaaaaaar'.

Statements with a hashtag are harder for GoogleTranslator because often several words are glued together to make it a hashtagged statement. Sometimes GoogleTranslator translated the individual words, and sometimes it does not. Table 3.3 shows examples of this: in tweet 3 '#boerenprotest' is not translated, but in tweet 9 it is translated to '#farmers protest', which is a correct translation. However, it adds a space between the two words, which can influence the results of topic modelling. Partially to mitigate this effect, bigrams are introduced to the pre-processing, which are explained in the next section. Finally, I think it is interesting that GoogleTranslator even translated numbers: it translated the Dutch decimal separator, a dot, to the US and UK decimal separator, a comma (see tweets 8 and 9).

Table 3.3 | Examples of original Dutch tweets and their English translation by GoogleTranslator

	Original Dutch text	English translation
1	Pijnlijk vooral voor het stervende CDA. Ooit de partij voor de boeren en platteland, gaat nu mee met Groen Links en het stikstofprobleem dat nog geen 2 jaar geleden is uitgevonden.	Painful especially for the dying CDA. Once the party for the farmers and the countryside, it now goes along with the Green Left and the nitrogen problem that was invented less than 2 years ago.
2	Liveblog: Kabinet sluit akkoord met boeren. https://t.co/yBfaWlnWNa https://t.co/WBXlKEsJwo	Live blog: Cabinet reaches agreement with farmers. https://t.co/yBfaWlnWNa https://t.co/WBXlKEsJwo
3	De eerste prijs voor domme doos van het jaar gaat naaaaaaaar........merellllllll #boerenprotest https://t.co/SLaoyaDcmY	First prize for stupid box of the year goes toaaaaaaar........merelllllllll #boerenprotest https://t.co/SLaoyaDcmY
4	@68evz PFAS en stikstof zijn onzinptoblemen, zelf gecreÃ«erd op basis van zelfverzonnen regelgeving. Aan CO2 moet gewerkt worden maar wel met beleid en niet op basis van paniekzaaierij. Nederland h...	@68evz PFAS and nitrogen are nonsense problems, created on the basis of self-made regulations. CO2 needs to be worked on, but with policy and not on the basis of scaremongering. The Netherlands do...
5	Deze slag om stikstofbeleid is voor boeren https://t.co/3Jk9orUUCO De kabinetsmaatregel om de bouw te helpen door met regels voor veevoer stikstofneerslag terug te dringen is van de baan.	This battle for nitrogen policy is for farmers https://t.co/3Jk9orUUCO The cabinet measure to help construction by reducing nitrogen precipitation with animal feed rules has been discontinued.
6	Wat mij bezig houdt is dat keer op keer de boeren voor nieuwe maatregelen komen te staan en dus nieuwe investeringen moeten doen.	What concerns me is that time and again farmers are faced with new measures and therefore have to make new investments.
7	Fotoreportage boerenprotest gemeentehuis Wijhe https://t.co/CyF3GBnrV8	Photo report farmers protest town hall Wijhe https://t.co/CyF3GBnrV8
8	Bijna 3.500 aanmeldingen voor boerenprotest op 1 oktober https://t.co/3ktydDvhZQ	Nearly 3,500 registrations for farmers' protest on October 1 https://t.co/3ktydDvhZQ
9	#boerenprotest Europa verliest 1.000 boerderijen per dag en klimaatveranderingen kunnen leiden tot uithongering https://t.co/eyxc51aYsF	#farmers protest Europe is losing 1,000 farms a day and climate change could lead to starvation https://t.co/eyxc51aYsF

After translation, tweets are ready for sentiment analysis. However, for topic modelling, they require pre-processing. The next section describes these steps.

3.2.2 Text pre-processing

The second column in Figure 3.3 above shows the steps taken for the pre-processing of text in preparation for topic modelling. From top to bottom, the first step in 'Text pre-processing' is the removal of URLs and references to twitter users. Then, the accents on characters are removed. For example, the word "Naïve" is turned into "Naive". This is done because, on Twitter, while some users are very strict with traditional spelling, many users are more careless and will, among other things, ignore the accents in words. Subsequently, special characters, like "&" and "#", are removed. Then, numbers are removed. However, when words consist of both letters and digits, the words are kept. This is because words like ‘natura2000’ and ‘CO2’ are relevant to the nitrogen crisis debate and should therefore not be removed.

After this, all characters are converted to lowercase, so, for example, the word "Dutch" and "dutch" will not be treated as two distinct words. The next step is to remove stop words. In NLP, these are words so common and frequently used that they add little meaning to a sentence. Examples of stop words are 'a', 'the', 'of' and 'it'. There are various ready-made lists of stop words. Here, the Dutch stop words list from NLTK, a rich Python library for NLP, is used (Bird, Klein, & Loper, 2009). If in the process of analysing the results more words are identified that have little added value, these words are added to the list of words that are filtered out. The full extended list of stop words is listed in Appendix B: List of stop words.

After this, all words are lemmatised. Lemmatisation converts words into their root form, known as lemma (Nandathilaka, Ahangama, & Weerasuriya, 2018). This means all plural forms of words will be converted to their singular form (e.g., "trees" → "tree"), verbs are converted to the first person regular (e.g., "demonstrating" → "demonstrate"), etc. Lemmatisation tries to take the context of a word into account. For example, the word "running" can lemmatised differently depending on context. In the sentence "I am running to the train station" the lemma is "run", while in the sentence "The actor is in the running for an Oscar" the lemma would be "running". For this step, the lemmatizer of Python library NLTK is applied.

Finally, collocations, or bigrams, are made. This is the combination of two words that occur frequently one after the other in the dataset. For example, if the words “nitrogen” and “crisis“ will occur often one after another, a bigram is created that is “nitrogen_crisis”: the words are combined with an underscore to form one word, so topic modelling will treat them as such. Figure 3.4 shows the list of most frequent bigrams in the dataset. Many bigrams consist of names, like ‘von_martels’ (House of Representative member) and ‘yvon_jaspers’ (farmer and activist). Some bigrams are organisations, like 'extinction_rebellion' and 'groen_links' (left-wing political party).

['biltse_rading',	'vnom_nieuwegein',
'von_martels',	'qrazy_shirts',
'forumvoordemocracy_reddit',	'landlord_levy',
'poly_perfluoroalkyl',	'polar_bear',
'grote_markt',	'fell_swoop',
'yvon_jaspers',	'divide_conquer',
'harry_mens',	'esther_ouwehand',
'trekzevantheir_pedestal',	'proudonzeboeren_boomersinresistance',
'arno_wellens',	'dark_waters',
'pim_fortuyn',	'marc_calon',
'pedestal_boerenzijnzo20th',	'shirts_qrazy',
'fairy_tale',	'groen_links',
'parental_alienation',	'pulse_fishing',
'abn_amro',	'bart_kemp',
'john_deere',	'ultrafine_particle',
'jaap_majoor',	'showed_overestimation',
'arjen_schuiling',	'albert_heijn',
'blah_blah',	'bnr_ikhoorbijbnr',
'maxime_verhagen',	'roos_canal',
'prof._dr.',	'pieter_omtzigt',
'uwv_jurisdiction',	'theo_hiddema',
'crocodile_tear',	'tata_steel',
'extinction_rebellion',	'persecution_jews',
'vno_ncw',	'omrop_fryslan',
'cbr_uwv',	'precautionary_principle']

Figure 3.4 | List of the 50 bigrams created from the dataset

3.2.3 Prepare for topic modelling

For topic modelling, two more steps are required, which are shown under 'Prepare for topic modelling' in Figure 3.3. First, the tweets are tokenized, meaning that instead of being a long string, the sentences are split on every space into single words. Then, a Dictionary is created with each unique word and how often it occurs in the corpus. To reduce the size of the corpus, shorten the processing time for topic modelling and increase the chance of coherent models, both the most occurring words and the least occurring words are filtered out: words that occur a lot will probably not have a specific relation to a topic, and words that are barely used are of little value. Here, words which occur in less than 2 documents and words which occur in over 50% of all documents are deleted. Then, on the right of Figure 3.3, the bag-of-words representation of the corpus remains, which is ready to be used for topic modelling.

3.3 Content analysis

3.3.1 Sentiment analysis

During my last thesis, I wanted to prevent translating tweets and risking their meaning getting 'lost in translation'. Therefore, I applied two sentiment analysis models that were developed especially for the Dutch language. I evaluated these and found they did not perform well. From literature I learned that NLP tools for less-spoken languages (compared to English) are often of lesser quality than tools developed for English (Dashtipour et al., 2016). Additionally, I learned that translating tweets to English and applying a pretrained English sentiment analysis method can lead to better results than applying sentiment analysis tools developed for a language that is not English (Mohammad, Salameh, & Kiritchenko, 2016). Also, various papers assessing the quality of automatic translators have found that the quality of translation suffices for applying sentiment analysis and retrieve reliable results, sometimes even better results than the sentiment tools in the original languages (Araujo et al., 2016; Balahur & Turchi, 2012). That is way for this research, I looked for high-quality translation and sentiment analysis tools.

HuggingFace is an open-source online platform for state-of-the-art NLP and AI models ("The AI Community Building the Future," n.d.). I selected a sentiment analysis model that is trained and evaluated on tweets and performs well: twitter-roberta-base-2021-124m[17] (Barbieri, Camacho-Collados, Neves, & Espinosa-Anke, 2020). The model is trained using approximately 124 million English tweets from 2018 to 2021, which roughly corresponds to the timeline of the nitrogen crisis dataset. This makes the tool compatible with the dataset of this research both in terms of data type (tweets) and timeline. The model returns three sentiment scores: positive, neutral, and negative, which together sum to 1. Unfortunately, there is no documentation on what the model considers 'neutral'. Therefore, I only use the positive and negative scores.

3.3.2 Topic modelling

As described in section 2.4.2, I use BTM for topic modelling (Xiaohui Yan, Jiafeng Guo, Yanyan Lan, 2013). I use the Python package 'bitermplus' by Maksim Terpilowski (Terpilowski, 2022). BTM is similar to classic LDA but differs in a few ways. For example, it models the likelihood of word pairs co-occurring in documents instead of individual words.

During my last thesis, I chose the split the dataset into week-sized chunks. I did this because the aim was to look at *differences* in topics discussed over time (Hendrikse, 2021). I looked only at whether the topic models differed over time, I did not look at the actual content of the topic models: the topics. For this study, I wanted to look at the *contents* of topic modelling. I aim to find out what topics were discussed and compare the prevalence of these topics over time, to see whether certain topics were discussed more than others at specific points in time. This

[17] Code available at: https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest

cannot be done if I create topic models for each week, because I would have to manually label all topics for each topic model for each week (over 100 topic models), which would be extremely time-consuming. Furthermore, it would be impossible to compare different topics. For example, imagine that from the 20 top words of one of this week's topic models, 15 are also present in a topic from next week's topic model. Can I then consider this the same topic and say that in both weeks the same topic was discussed? It would be very hard to define when two topics can be considered the same. Because I want to look at the prevalence of topics over time and compare their frequency. Additionally, during my last thesis I developed a processing timeline that could inspire future work on the real-time processing of social media content. However, in this study actuality of data and real-time monitoring of tweets is not relevant. For these two reasons, I chose to calculate just one topic model over the full dataset.

3.3.3 Characterising demonstrations: codebook

As discussed in Chapter 2 section 2.5.2, the definitions of categories for drawing up a demonstration codebook are loosely based on the Fillieule's list (Fillieule, 2012):

1. the temporary occupation of open physical spaces
2. expressivity
3. number of participants
4. political nature of the demonstration

I evaluated these elements on quantifiability and how feasible it is to collect data on them from news articles as data sources. Number 1, 'the temporary occupation of open physical spaces' contains the elements of time, occupation and location. The time a demonstration lasted is, however, rarely reported on, especially for the smaller, lesser-known, demonstrations. Since an 'occupation' can take various forms and my list of demonstrations contains several types of demonstrations, I made 'Type' of demonstration one of the categories. I defined four types of demonstrations based on my list: a 'field demonstration', 'home visit', 'occupation' and 'other'. A 'field demonstration' is defined here as an occupation of a park, in this case: the Malieveld. A 'home visit' is defined as a visit by multiple farmers to the private home of a politician or other well-known stakeholder. Finally, an 'occupation' is defined as an occupation of any other kind. For the last element of number 1, location, I created two categories: 'Location nr' and 'Type of location'. The first, Location nr, is the number of locations where demonstrations took place on the same date. I included this information because farmers' organisations at times planned demonstrations throughout the Netherlands on the same date, to demonstrate in front of Province houses in different cities for example. I do not count different locations within the same city, though, so if demonstrations took place in the Hague both at the Malieveld and later at another location in the Hague, I count this as one location. This category provides insight into the sparsity of demonstrations and, depending on the type of demonstration, into the degree of organisation by the farmers' organisations. The second, 'Type of location' is defined as one or more of the following: 'Public space', 'Highway', 'Governmental institution', 'Commercial private property' or 'Personal private property'. 'Public space' is a demonstration on a location that allows demonstrations, like the Malieveld in The Hague. 'Highway' is an intentional blockade on the highway, which is illegal. 'Governmental institution' is a location that is governmental, like province houses or the RIVM (the National Institute for Public Health and the Environment). A 'Commercial private property' is private property that is owned by a company, for example, supermarket distribution centres. Finally, 'Personal private property' refers to a residence belonging to, for example, a politician and their family.

Number 2 on Fillieule's list is, 'expressivity', refers to the crowd's uniformity. It is hard to come up with categories to characterise expressivity. The category 'Organiser/initiator' aims to cover an aspect of uniformity: it documents the organiser of a demonstrative event. However, unfortunately, this information is often not reported on in news articles. This leaves some missing entries in this category. What is sometimes mentioned is which organisation the protestors were members of.

Number 3 is the 'number of participants', which is the number of attendees of a demonstration called 'Nr of attendees' in the codebook. If the 'Location nr' is higher than 1 the attendees of the various demonstrations are summed up to form the Nr of attendees.

Last on Fillieule's list is the 'political nature of the demonstration'. The political nature of the demonstrations is rather consistent throughout the crisis: to demonstrate the reduction of the livelihood of farmers that the nitrogen policy imposes. Therefore, no categorisation is for this item is included in the codebook.

What Fillieule's list does not capture is the intensity, or extremity, of a demonstration. The demonstrations have grown more aggressive and demanding over time (Kalkhoven, 2021), which could indicate polarisation. Therefore, several categories are defined to capture extremism and the aggressive nature of demonstrations. First, 'Heavy machinery', indicates whether heavy machinery, like tractors, was present and/or used during an event. This is a 'yes' or 'no' binary category. When news articles mention heavy machinery or contain pictures of the demonstration which show tractors, this value is set to 'yes'. In one case, I even looked at Youtube videos to see whether they showed tractors, and referred to this video as a source Accompanying this category is 'Nr of tractors': the number of tractors present at an event. When a newspaper mentions (an estimate of) the number of tractors that were present, this is the Nr of tractors.

Furthermore, the category 'Threats' indicated whether threats were made during or surrounding a demonstration. This is one of the harder categories to define, as the opinion on whether something is a 'threat' can differ per person. For example, the person making threats can perceive their comments as 'unthreatening' while the receiver might feel threatened. There have been examples of these types of disagreements during the nitrogen crisis. There are multiple types of threats considered here. If the receiver of a threat *felt* threatened, this category will be set to 'yes'. There is an example where about 140 tractors went to Vollenbroek's house to 'hand him a letter', where Vollenbroek reported not feeling threatened. Even though many other stakeholders and media sources reported this as a threat, I do not consider the action a threat because Vollenbroek did not consider it threatening. When it is not reported on how the receiver felt about the action but multiple sources mention an event was a form of intimidation, it is assumed here that there were threats as well. Additionally, statements that consist of a structure like: "if [stakeholder] does not do [action] within [period of time], we will [announcement of more hostile counter action]" are considered threats. This is based on the definition of a 'threat' in the Cambridge Dictionary[18]. Furthermore, spreading confidential private information, like someone's phone number or house address, is considered a threat because it can lead to unwanted and potentially dangerous outcomes, like threatening phone calls or unwanted house visits. When tractors are parked in a way that suggests they might be used to cause damage, this is considered threatening. An example of this is when farmers parked a tractor in front of the door of a Province House (which was eventually driven down by the tractor). Finally, when physical objects are thrown at other people, I consider this threatening behaviour as well. This is because even though the object being thrown can be harmless, like hay, I argue this action violates someone´s personal space and is therefore considered threatening, especially in the context of a demonstration.

What the 'Threat' category does not consider is threatening statements on protest boards. This is not because these statements cannot be considered threatening, but because most of the protest boards at these events contain possibly threatening sentences (like "This is war"[19]) (Omroep West, 2020). However, as the slogans from protest boards are not systematically documented anywhere, the only way to find what protest boards were present is by scanning the collection of pictures of an event. This is incredibly time-consuming, and it will result in a bias toward demonstrations that gained more media coverage and which, as a result, have more photographs online. As a result of this, threatening language on protest boards does not contribute to this category. However, when I find articles about a visual statement during a demonstration discussing how threatening it was, I do consider that a threat. An example of this is the use of a coffin with a politicians name on it during a demonstration. This event was picked up by several newspapers that considered it shocking. When 'Threat' is set to 'yes', an explanation of what happened and why it is considered threatening is added to the codebook.

Complementary to the 'Threat' category, are 'Vandalism', 'Police intervention' and 'Arrests'. 'Vandalism' is considered present (and set to 'yes') when something was broken intentionally or unintentionally during a demonstration. 'Police intervention' is set to 'yes' if police officers were present. If they were, a description of what they did during their presence is included (sometimes they are just there to guard, other times they engage and make arrests). Finally, 'Arrests' documents whether arrests were made, and if 'yes', explains why and how many arrests were made, if known.

[18] https://dictionary.cambridge.org/dictionary/english/threat
[19] Original text: "Het is oorlog". Translated by the author.

Additional to these categories, the codebook includes 'Date' with the date of the event, 'Event' with the title of the event, 'Description' with a description of the demonstration and finally 'Sources' with references to the news articles that were the sources of information per event.

Table 3.4 below shows an overview of all categories in the codebook with descriptions and what occurrences are included and excluded when filling in each category.

For all categories, if nothing is mentioned in news articles describing the phenomenon of the category, I insert 'Unknown'. For the yes/no categories, only if there is a source explicitly stating that there was either the presence or absence of the phenomenon, I insert 'yes' or 'no'. This is because when one of these phenomena is not mentioned in a news article, for example when there is no mention of vandalism, I cannot prove that indeed there was no vandalism. It could simply not have been reported on. However, when news sources describe a demonstration as, for example, peaceful or 'free of incidents', I will insert 'no' in the 'Vandalism' category.

Table 3.4 : Overview of the categories in the codebook

CATEGORY	DESCRIPTION	TYPE	INCLUDES	EXCLUDES
Date	Date of the event	Date	NA	NA
Event	Name of the event	Text	NA	NA
Description	A description of the demonstration	Text		
Organiser/initiator	Organisation responsible for organising the demonstration	Text	NA	NA
Type	Type of demonstration	'Field demonstration', 'Occupation', 'Home visit' or 'Other'	NA	NA
Location nr	Number of locations where demonstration took place	Number	Cities, roads	Demonstrations within the same city
Type of location	Type of location	One or more from: 'Public space', 'Highway', 'Governmental institution', 'Personal private property' or 'Commercial private property'	NA	NA
Nr of attendees	Number of attendees	Number	Sum of attendees is Location nr > 1	
Heavy machinery	Whether heavy machinery was present	'yes', 'no' or 'Unknown'	Tractors	
Nr of tractors	Number of tractors	Number		
Threats	Whether threats were made	'yes', 'no' or 'Unknown', incudes description if 'yes'	Threatening language, when the victim felt threatened or when objects are thrown	Threatening statements on protest board. Furthermore, when the victim nor other stakeholders report an action as threatening this is not considered threatening
Vandalism	Whether there was vandalism	'yes', 'no' or 'Unknown', incudes description if 'yes'	When objects were intentionally or unintentionally broken	NA
Police interaction	Whether the police were present and took action	'yes', 'no' or 'Unknown', incudes description if 'yes'	Fines, police presence, traffic regulation by officers, deployment of the army	NA
Arrests	Whether arrests were made	'yes', 'no' or 'Unknown', incudes description and number if 'yes'	Arrests	NA
Sources	The news articles that were the source of information for filling in categories per event	References	NA	NA

3.4 Evaluation

3.4.1 Sentiment analysis: compare to annotators

During my previous thesis research, I asked two people to annotate 100 tweets by sentiment. Annotating the same tweets myself, I created a small set of tweets with sentiment scores from three annotators that I then compared to the automatically generated sentiment. By calculating the overlap between the annotations and the sentiment scores of two different sentiment analysis tools, I chose the best-performing one. Because, in this research, I selected a state-of-the-art sentiment analysis model, I am more interested in the distribution of sentiment scores. These are plotted in Figure 3.5 below.

Figure 3.5, on top, shows the sentiment scores per annotator for 100 tweets. The annotators were asked to score the tweets on a scale from -10 to 10, where -10 is the most negative, 0 is neutral, and 10 is the most positive score. This figure shows that most tweets are annotated with a score between -5 and 2.5. Roughly an equal amount of tweets are labelled an extremely positive or negative sentiment.

The figure on the bottom shows the distribution of sentiment scores of the *whole dataset*, labelled by the sentiment analysis model. It is important to note how the scale differs from the top figure. In the lower figure tweets are labelled with *three different sentiment scores*. One for positivity, one for neutrality and one for negativity. These scores are given on a scale from 0 to 1.0. So, the closer to 0, the lower the positivity/neutrality/negativity. The lower figure shows that all tweets have very low positivity scores. Neutrality scores are low to medium, ranging from 0 to 0.5. Negativity, however, ranges from medium (0.4) to very high (1).

Both figures show a high number of tweets with a low positivity score. However, the most apparent difference between the two graphs is the complete lack of highly positive scores from the sentiment analysis, where all three annotators manually assigned multiple tweets as highly positive. This poses the question of whether the automated sentiment analysis is inadequate for the detection of positivity in the nitrogen crisis tweets. The implications of this are discussed in section 5.5.

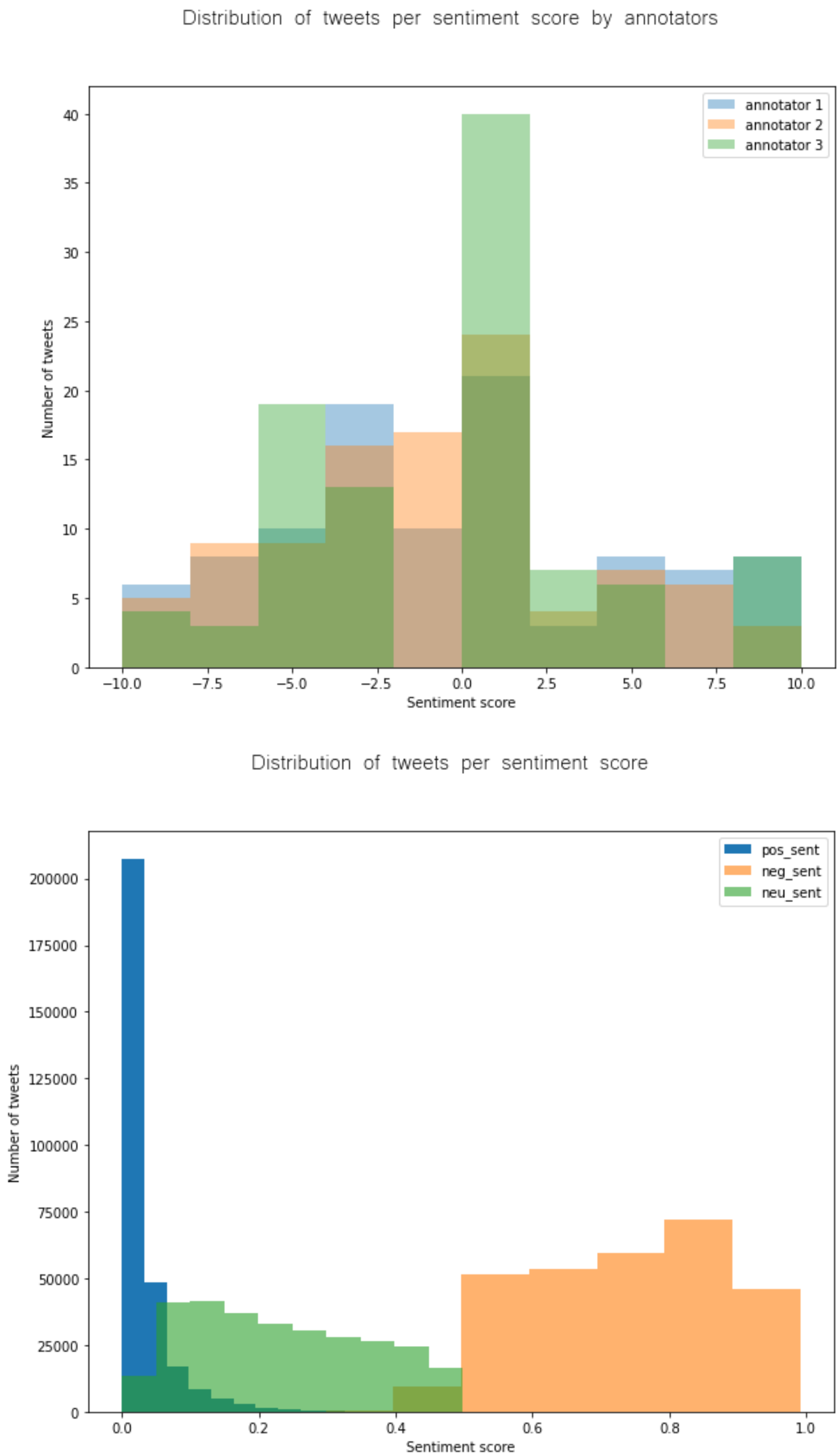

Figure 3.5 | Top: Distribution of tweet per sentiment by annotators. Bottom: Distribution of tweets per sentiment score. Note the difference in labelling and scale. In the top figure tweet scores are plotted once per annotator. On the bottom figure, tweets get three scores: one for negativity, one for neutrality and one for positivity.

3.4.2 Topic modelling: selecting the best topic model

Tweaking the topic model

Tweaking is important during topic modelling to get good results. One important hyperpar ameter for the quality of a topic model is the number of iterations. I defined this as high as I could: 75 iterations. My computer would crash if I tried more iterations. To evaluate whether this number of iterations is sufficient, I compared it with benchmarks from the bitermplus documentation, see Figure 3.6 below (Terpilowski, 2022). The figure shows the results of sixteen BTM models trained with iteration numbers from 10 to 2000 and the topic number set to 8. These models are evaluated with two metrics: perplexity (the lower the better) and semantic coherence (the higher the better. This metric is further explained in section 'Coherence' below). Both the perplexity and the coherence of the topic models stabilize around 200 iterations and higher. At 75 iterations the perplexity is a little higher, and the coherence is roughly the same. I conclude that 75 iterations are sufficient for training and interpreting topic models, but I suggest future research use a higher number of iterations to guarantee more stable and higher quality topic models.

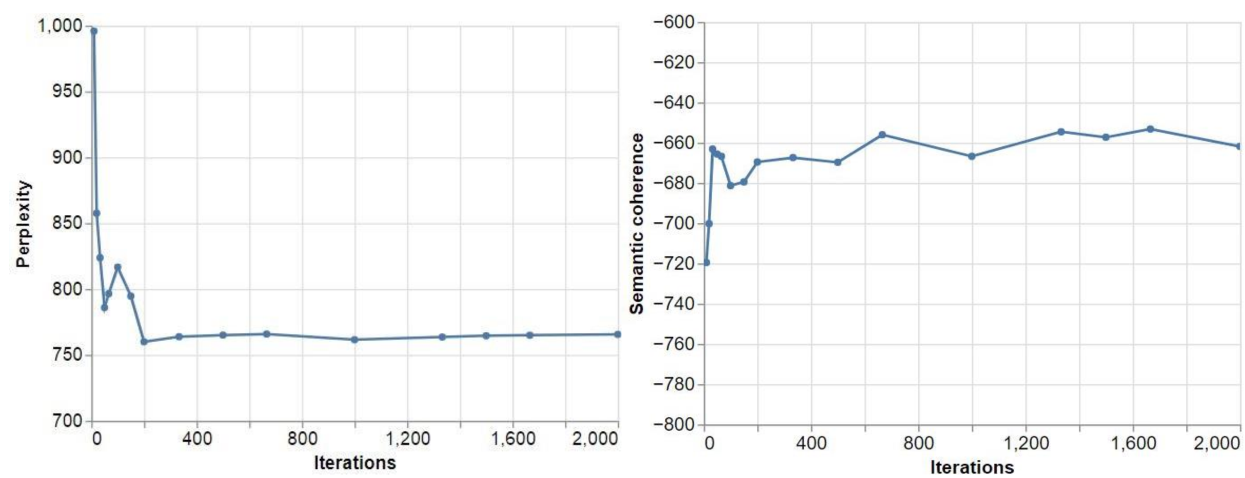

Figure 3.6 | Results of a series of benchmarks done on SearchSnippets dataset. Sixteen models were trained with different iterations number (from 10 to 2000) and default model parameters. Topics number was set to 8. Perplexity (left) and semantic topic coherence (right) were calculated for each model (Terpilowski, 2022).

For the number of topics hyperparameter, I tried 10, 20 and 30. After computing topic models with these varying amounts of words per topic, I decided to use 20. Also, while computing various topic models with different hyperparameter inputs, I would add words to the stopword list if they were meaningless or occurred in multiple topics (See Appendix B: List of stop words for the extended list). For the most important hyperparameter choice, of how many topics the topic model should consist, I used an evaluation metric described in the next section.

<u>Selecting a topic model</u>

The choice of a number of topics for topic modelling is made based on three steps:

1. Comparing Coherence metric scores per topic model
2. Running the built-in function for finding 'stable topics'
3. Manual analysis of topics per topic model

The next sections describe each of these steps.

Coherence

The algorithms BTM and LDA do not guarantee 'coherent' results: that each topic returned comprises words that are contextually related and properly represent a topic that was described in the corpus that served as the training set. The state-of-the-art approach for evaluating the coherence of topics in topic models is human evaluation, a method that is costly, sensitive to biases and, in the case where many topic models need evaluation, extremely time-consuming (Röder, Both, & Hinneburg, 2015). Newman et al. developed a metric that approximates human annotation, called topic model coherence (Lau, Baldwin, & Newman, 2013). A high coherence score stands for a more coherent and interpretable topic model. By selecting topic models based on topic coherence, topic models are picked with a high chance of being interpretable by humans, and thus decision-makers. Topic coherence is a general metric that is widely used for estimating the quality and coherence of topic models and it is implemented in the bitermplus package, hence easily applicable.

Figure 3.7 shows the mean coherence scores for all calculated topic models with between 1 and 25 topics. The figure shows that a lower number of topics leads to a higher coherence score. However, although choosing a number of topics of 1 or 2 would lead to the highest coherence, I am interested in more than the overarching topics of the dataset, which are the nitrogen crisis and farmers' demonstrations, namely what additional underlying topics are being discussed on social media during this policy crisis. Therefore, based on Figure 3.7. a number of topics between 4 and 7 is preferred, because coherence scores are still on the higher side but there are several topics to analyse.

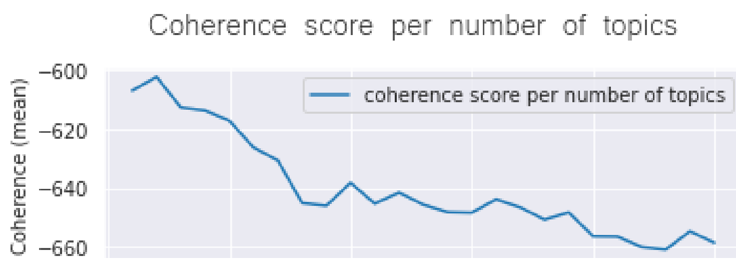

Figure 3.7 | Coherence scores (mean) of topic models with between 1 and 25 topics

Stable topics

The bitermplus package comes with a built-in function called 'stable topics'. This function aims to check whether a topic occurs throughout multiple topic models. To do this, stable topics analysis compares the topics from several topic models. If it finds topics that are 'stable', this indicates a stable topic and therefore a higher probability of being a 'real' topic. I used this function to analyse topic models of the dataset with a number of topics between 4 and 7.

The function first calculates the distance to the topics in other topic models. The smaller the distance, the higher the probability that they are the same topic. There are a few distance metrics available, for which the Kullback-Leibler and Jaccard index were tried, as these are the ones mentioned in the docs for the next step. The final parameter in this distance calculation step is how many top_words of a topic to include in the distance calculation. This is set to 20, in because this is the number of words chosen to represent a topic in section 3.4.2. All these distances are calculated with respect to one of the topic models: the reference model. The outputted topic number IDs correspond to the topic numbers in this reference model. The reference model chosen here is the one with the lowest number of topics: the topic model with 4 topics.

The distances between topics are then used to determine which topics occur across multiple models. The first parameter in determining stable topics is the distance threshold, describing which distance value would mean two topics from different models are considered to be the same topic. I experimented with this value, but ended up using the default 0.9 threshold value because it made little difference. The second parameter is the minimum topic recurrence frequency across all models. Because my models have an overarching topic, the nitrogen crisis, and therefore are quite similar, this value is set high to 80%.

Despite experimenting with various values, all topics seem to be stable. When choosing extreme values, only topic 1 can get filtered out. This is probably due to the overarching main topic, the nitrogen crisis and its demonstrations, meaning that a topic will always show some overlap with one of the topics from another topic model. This method of detecting stable topics is likely more suitable for clearly separable topics. For now, it shows that there is consistency throughout the topic models with varying numbers of topics, which shows that the choice of number of topics will not heavenly impact the topics shown in the selected topic model.

Manual analysis

Finally, to choose a determinate topic model, I looked at the top 20 words for each topic in the topic models with between 4 and 7 topics. I looked at how well I could label these topics, and how distinguishable the topics were. If a topic model with a higher number of topics contains two or more topics that to me look like they describe the same topic, I looked at the topic model with a topic less until I saw no more apparent duplicates. Additionally, if I could not label a topic because it appeared incoherent to me, I would move to a topic model with a topic less. Based on this third step in the topic model selection process, I chose the topic model with 4 topics. This topic model has a high coherence score and is interpretable to me.

Labelling topics

When the best-performing topic model is selected with the methods described above, I will label the topics. This is done by looking at the words in each topic, paying special attention to the most relevant topics, and connecting a topic or phenomenon from the nitrogen crisis to it. Additionally, when topics appear similar, I look at which words are present in one that are absent in others, which could be key in defining a topic. The labelling is heavily based on interpretation and prior knowledge of the nitrogen crisis and its events.

3.5 Data and code

The data and scripts used in this study are available at my repository 'ThesisNLP 'on GitHub, under the account of username milahendrikse[20].

[20] https://github.com/milahendrikse/ThesisNLP.git

Key Findings of Chapter 3: Methods

1. A **query** is set up to collect tweets about the Nitrogen crisis
2. Tweets are **translated** and **pre-processed** before sentiment analysis and topic modelling
3. A **codebook** is defined for categorising demonstrations
4. **Coherence**, **stable topics** and **manual inspection** decides what topic model is chosen as the best performing model

4 Results

This chapter describes the filled-in codebook with characteristics of the demonstrations of the nitrogen crisis and the results and interpretation of topic modelling and sentiment analysis. First, the demonstrations are arranged and characterised in section 4.1. Then, the best-performing topic model is shown and labelled in section 4.2.1. In section 4.2.2 first the mean sentiment of each topic from the best performing topic model is plotted. This is followed by a figure containing the weekly volume of tweets, the weekly mean sentiment of the dataset, the volume of each topic in the best performing topic model over time and the weekly ratio of each topic. Lastly, section 4.2.3 discusses interesting findings in topic models that consist of more topics but a lower coherence.

4.1 Defining demonstrations

Table 4.1 shows the filled-in codebook of the major demonstrations during the nitrogen crisis. Entries with missing data are coloured grey. The categories with the most missing data are 'Nr of attendees' (9 missing entries), 'Nr of tractors' (9 missing entries) and 'Vandalism' (10 missing entries). It turns out that newspapers report on protests in a very unsystematic manner. For multiple demonstrations, it is impossible to find information about the number of attendees or tractors that were present. Additionally, when looking at the missing entries, it shows that the first four demonstrations have fewer missing entries compared to the later demonstrations. This indicates news media coverage of demonstrations decreased over time. Some categories, like 'Vandalism' or 'Threats', are bound to have more missing entries because I only filled them with 'yes' or 'no' when the presence or absence of vandalism or threats are explicitly stated. However, a journalist's interest will be served more by mentioning the appearance of one of these phenomena than by reporting on their absence.

In the next sections, I will describe the demonstrative events per year.

4.1.1 2019: The nitrogen crisis kicks off with large demonstrations

When we look at the 'Location nr' category, we notice demonstrations taking place at numerous sites at first. The multiple locations of early demonstrations can be explained by the national outcry and the anger farmers felt and expressed for the first time during these demonstrations. More unrest can mean more demonstrations. Another reason could be that the farmers were not as well organised and united in the beginning as they would later become. It is unfortunate there is so little information on the number of people who attend demonstrations; otherwise, this data could be compared to the number of locations to gain a sense of the scope of the protests. However, at least three of the demonstrations in 2019 are huge: several thousand people attend.

In 2019, the first and biggest demonstration day is mostly a field demonstration in The Hague. Some farmers block a train station in Groningen, but the enormous demonstration in The Hague receives the most attention. Even though there are already instances of vandalism and farmers are arrested for breaking through fences and driving dangerously in The Hague, there is public empathy for the farmers in the beginning. It is noteworthy that FDF founder Mark van den Oever already uses threatening language. He mentions that FDF will sabotage the national food supply, air traffic at Schiphol and the popular Formula1 if the government does not act the way FDF wants. We will see how a few of these threats are (partially) realised in the future. During the remainder of 2019, all demonstrations are occupations (see category 'Type'). FDF (maybe in cooperation with other organisations) organises several demonstrations at Province Houses in the Netherlands. These have quite a threatening atmosphere: tractors are intimidatingly parked in front of the doors of Province Houses, farmers throw hay at police and in an altercation a police horse is injured. In one city, the monumental door of a Province House is broken into. Other demonstrations in 2019 are similar. They are occupations of governmental buildings, supermarket distribution centres and highways. Threats are still being made, and we see the start of personal threats in 2019 as well. FDF organise or co-organise all demonstrations in 2019. Intimidation, vandalism and threats are part of the farmers´ demonstrations from the very beginning.

Overall, 2019 shows the passionate beginning of fervent farmers' demonstrations. It depicts the usage of a large number of tractors, as well as threatening language and behaviour, and also police intervention. Arrests are made solely for extreme behaviour, such as throwing fireworks, and not for the fact that tractors are not permitted to be used in demonstrations. Following the success of the large-scale demonstration in a public location, more demonstrations on private property and highways are organised throughout 2019.

4.1.2 2020: Continuation of both large- and small-scale demonstrations with more police intervention

In 2020, we see more single-location demonstrations at first, and only in December do demonstrations take place at multiple locations again. There are several possible explanations for this. Farmers have very demanding jobs and leaving the responsibility of a farm out of hands for a day must be inconvenient for a farmer. This could explain why demonstrations are less scattered after the first year of demonstrations: maybe farmers are waiting out the results of their first demonstrations and cannot afford to leave their farms for a day. On a few occasions, there may have been multiple demonstrations per day, but these protests might have been so small-scale that only a few were reported. Another explanation is that, after I found news items about a July demonstration, it was easy to find news articles about other July demonstrations because news articles frequently refer to other recent demonstrations. Note that in July there are multiple demonstrations just a day apart. In December, the nr of locations goes up again.

In 2020, the personal nature of demonstrations develops further. Home visits occur for the first time in 2020: farmers going to the house of key stakeholders in the crisis to 'hand them a letter' or 'bring them food'. Although the largest scale of these gestures, visiting Johan Vollenbroek, is not considered threatening by Vollenbroek, it seems the farmers want to intimidate and put their former threats to action. The awaiting of Minister Schouten at the location of her official visit to Zeeland, as well as the dissemination of personal contact information of high-ranking Friesland Campina personnel, are in line with this behaviour.

Blockades of distribution centres become a frequent form of protest for the farmers. It seems FDF sticks to its promise to sabotage food supply in the Netherlands if the government does not act as FDF wants. However, FDF distances itself from the last distribution blockade that takes place in 2020, advising its members not to act on their own accord and demonstrate illegally. It is possible that FDF was indeed not in charge of organising the event. However, another possibility is that FDF was alarmed by the recent public opposition towards its latest demonstrations and did what it said never to do: make a strategic decision to try and reassure the public they were not involved to save some public support.

The police step up and enforce the prohibition of demonstrating with tractors[21] for the first time in 2020. This could be because public support for the farmers' demonstrations is on a decline and the police, less afraid of public backlash, can react with more decisiveness now.

4.1.3 2021: Potential polarization amongst the farmers

The table contains only one demonstration date in 2021. On this date several demonstrations took place. What's striking is that both Agractie and FDF, two farmer organisations, demonstrate, but they do so completely independently from each other. They were both present during the Malieveld demonstrations in 2019. However, this time Agractie organises a peaceful demonstration in The Hague while FDF spreads out to demonstrate at four different province houses. The fact that Agractie and FDF separately organise demonstrations on the same date indicates the possibility they do not want to cooperate anymore. FDF has developed a militant and rule-breaking reputation by now, and it is probable that Agractie, which is less radical and more cooperative, no longer wants to be associated with FDF. This can be seen as a form of

[21] Demonstrating with heavy machinery was already illegal before the nitrogen crisis but was tolerated during the first demonstrations.

polarisation of farmers between more radical (FDF) and engagement in a more traditional way (Agractie).

4.1.4 An overview of 2.5 years of crisis

Overall, Table 4.1 depicts an aggressive side to the farmers' demonstrations, showing that intimidation, vandalism and threats were evident from the very beginning of the farmers' demonstrations. Tractors, though forbidden in demonstrations, were used during nearly all demonstrations, if not all. The vast majority of the demonstrations were organised by FDF. While the first response may be that this shows that most demonstrations were in fact organized by FDF, another possibility is that FDF protests just attract more media attention due to their disruptive character.

Another pattern shown in Table 4.1 is that demonstrations often take place in clusters. We see multiple demonstrations in a row in October 2019, July 2020 and December 2020. One way to explain this is that when unrest is sparked, farmers are inclined to express their dissatisfaction and try to get as much media attention as possible before the issue becomes old news. It is also easier to find demonstrations that were not far apart because news articles often refer to recent demonstrations. Although not likely, it is possible that entire clusters of demonstrations are absent from this table if I missed an article about one of them.

Table 4.1 | Demonstrations of the nitrogen crisis and their characteristics

	Date	Event	Description	Organiser/initiator	Type	Location nr	Type of location	Nr of attendees	Heavy machinery	Nr of tractors	Threats	Vandalism	Police interaction	Arrests	Sources
2019	1th October 2019	Demonstration on Malieveld	Large demonstration on Malieveld in the Hague. 1136 km traffic jam on highways caused by protesters on their way to The Hague. There is a smaller demonstration in Groningen where farmers park their tractors in front of a train station.	FDF, Agractie and potentially more	Field demonstration, Occupation	2	Public space	thousands	Yes	2200	Yes, by FDF, "there needs to be change in the coming 4 weeks, otherwise there will be 'hard' actions from FDF". Mark van den Oever, founder of FDF mentions they will sabotage national food supply, Schiphol and stop Formula1 from happening if necessary.	Yes, farmers breaking through fences and the grass of the Malieveld is damaged because of the tractors	Yes, the police helped navigate the demonstrations and regulated traffic. They also made sure the farmers could not get to the binnenhof[22]. In Groningen, the police navigate the traffic around the train station.	3 arrests were made for breaking through fences with tractors and driving over median strips	(NOS, 2019c)
	14th October 2019	Nationwide farmer protests	Farmers raise demonstrations nationwide. Farmers travel across the country to protest in front of various Provincial Government Buildings.	FDF and potentially more	Occupation	6	Governmental institution	Unknown	Yes	Unknown	Yes, tractor in front of province house door, farmers throw with straw at policemen	Yes, the door of a Province house in Groningen is broken in with a tractor. Also, farmers broke through fences. Also, a police horse was wounded	Yes, there was a fight with the police. A police horse was injured	5 people were arrested and sentenced to 100 hours of community service	(AD, 2021; Klumpenaar & Van Laarhoven, 2019; NOS, 2019a; NU.nl, 2019)
	16th October 2019	Protests at RIVM in the Hague	Farmers are protesting near the RIVM in the Hague. The Dutch army is deployed to keep the farmers from entering the Binnenhof (where the House of Representatives is). Tractors break through fences to reach Malieveld.	FDF and potentially more	Occupation	3	Governmental institution	Thousands	Yes	Thousands	Yes, a coffin on a tractor with politician Jesse Klaver's name on it	Yes, farmers breaking through fences	Yes, the army is deployed to limit where farmers can demonstrate	4 people were arrested, 2 for throwing fireworks, 1 for offending a police officer and the last for not obeying an order from an officer.	(NOS, 2019d)
	25th October 2019	Farmer protests at province house and dairy firm	Farmers protest the Nitrogen policy and, this time, also the planned policies of the Dutch dairy sector. They block the entrance to the head office of Friesland Campina, the biggest Dutch dairy firm. Also, personal telephone numbers of high-ranking employees of dairy companies are made public by farmers. Finally, 2000 farmers head to a province house in North Brabant to demonstrate.	FDF and other	Occupation	2	Commercial private property, Governmental institution	2000 farmers at the province house	Yes	20	Yes, personal information was spread	Unknown	Police were guarding the Province House	Unknown	(Omroep Brabant, 2019; Schelfaut, 2019)
	18th December 2019	Occupation of highways and supermarket distribution centers	Farmers blocked highways and supermarket distribution centers, even though a Dutch judge had forbidden it.	FDF	Occupation	Multiple	Highway, Commercial private property	Unknown	Yes	Unknown	Unknown	Unknown	Unknown	Unknown	(Nieuws, 2019; Youtube, 2019)

[22] The most important pariliamant building in The Hague.

2020

18th and 19th of February 2020	Farmer protests in the Hague	More farmer protests, many fines are handed out to farmers entering the highways with their tractor despite instructions not to do so. In the Hague a few hundred people demonstrate.	Unknown	Field demonstration	1	Public space	Unknown	Yes	Unknown	Unknown	Unknown	Yes, During the two days, 49 farmers are fined on two different locations for driving on the highway despite strict instructions not to do so. One farmer refuses to cooperate with the police, after which he is peppersprayed and the door of his tractor is forced open to remove the tractor from the highway.	1, for driving dangerously with a tractor on the highway	(RTLNieuws, 2020)
6th July 2020	Farmers visit house of influential environmentalist to hand him a letter	Farmers visit the house of Johan Vollenbroek, the man who sued the Dutch government for not following EU nitrogen deposition rules. Around 140 tractors and cars tried to gather around his house, but the police redirected them to a public space. The police also took mister Vollenbroek there, where he had conversations with the farmers. The farmers gave mister Vollenbroek a letter where they explained they were in this crisis because of him.	FDF	Home visit	1	Personal private property	Over 100	Yes	140	No, even though the event has been described by politician Rob Jetten and many others as intimidating, Johan Vollenbroek has not made this comment and simply went to talk to the farmers	Unknown	Police were present to protect Johan Vollenbroek	Unknown	(Gelderlander, 2020c, 2021)
7th July 2020	Occupation of supermarket distribution centre	Around 150 tractors blook a AH distrubtion center and the roads that lead up to it. The occupation takes place during the evening.	Unknown	Occupation	1	Commercial private property	Unknown	Yes	150	Yes, considered intimidating	Unknown	No, the guards and police are too late and do not intervene	No	(Gelderlander, 2020b)
8th July 2020	Various protests across 4 different provinces. A waste disposal site is occupied. Minister Schouten cancels her visit to Zeeland due to security risks.	Carola Schouten, Minister of Agriculture, Nature and Food Quality, cancels her visit to the province Zeeland on the advice of the police of Zeeland. The police reported that farmers had tracked the location of her destination and had formed a large group on the location, awaiting her. At the same time, farmer protests take place in 3 other provinces. A waste disposal side is occupied, which turns into a very chaotic event where many arrests are made.	FDF	Occupation	4	Commercial private property, Public space	Unknown	Yes	Unknown	Yes. 3 ministers receive death threats. Also, FDF wants to show people what happens when there is no food ("Hen erop wijzen wat het belang van goed, gezond én bereikbaar voedsel is. Want wat als er een kinkje in die kabel komt? We gaan het beleven. Binnenkort…"[23])	Unknown	Yes, police tried to stop farmers from reaching the waste disposal site. Many arrets were made.	63 arrests at the waste disposal site for not adhering to the prohibition of demonstration by tractor	(BNNVARA, 2020; RTV Drenthe, 2020; Wijnants, 2020)
11th July 2020	Farmers blockade at Centraal Bureau Levensmiddelenhandel in Leidschendam.	Farmers, among whom FDF members, occupy a supermarket trade association in Leidschendam. They are not only protesting nitrogen policy but also the low prices that supermarkets pay for the farmers' products. It is a peaceful demonstration. However, after the demonstation a FDF member is arrested for threatening to drive into a police officer.	FDF and potentially more	Occupation	1	Commercial private property	Unknown	Yes	50-60	Yes, threat of driving into police officer	Unknown	Police officers are present, but do not intervene	1, FDF foreman Thijs Wiegers for threatening to drive into police officer	(Gelderlander, 2020a; Leidsch Dagblad, 2020)

[23] Translation: "*Point out to them the importance of good, healthy and accessible food. Because what if there is a kink in that cable? We're going to experience it. Shortly…*". By Google Translate.

	28th October 2020	Farmer show up at home of politician	Farmers show up late at night at the house of Rob Jetten, a D66 politician who wants to reduce the livestock in the Netherlands. Jetten is in quarantaine because of COVID, and the farmers bring him a package of food with, among others, meat, even though Jetten eats vegetarian as much as he can. Jetten reports he thinks this gesture goes too far: intimidation wrapped in a nice gesture.	FDF members	Home visit	1	Personal private property	5	Unknown	Unknown	Yes, visiting a politicians house	No	No	No	(NRC, 2020)
	17th November 2020	Farmer protests in the Hague	Farmers go to the Malieveld with their tractors to protest the Nitrogen policy. The police try to stop them from reaching Malieveld with their tractors, but some protesters find ways to pass the police and reach Malieveld.	FDF	Field demonstration	1	Public space	Unknown	Yes	Unknown	Unknown	Yes. The year before, the Malieveld was badly damaged by the tractors. This is why they were not allowed on the Malieveld this time. Since tractors managed to reach Malieveld anyway, it is assumed there were some damages again.	Yes, police intervene to try and stop farmers from reaching the city center. Military defense trucks are used for this. Also, with sirens on, police chase a tractor that is driving backwards against traffic in the city center of the Hague.	2, unknown for what	(HartvanNederland, 2020; West, 2020)
	First half of December, 2020	Highly criticised farmer protests and blokades	Farmers protest both the Nitrogen policy and the low prices Dutch supermarkets ask for their products. Various supermarket distribution centres were barricaded, and the protests received a lot of criticism for going too far and the protesters were criticised for not following COVID regulations. FDF starts asking its followers not to demonstrate at distribution centers.	Other	Occupation	5	Commercial private property	Unknown	Yes	Unknown	Unknown	Unknown	Yes, police were present	Unknown	(NOS, 2020a, 2020b)
	14th December 2020	Occupation of supermarket distribution centre and visiting of directors house	De distibution centre of Jumbo is blocked, farmers make sure trucks can not reach it. They demand a conversation with the manager. Also, some farmers are send away by police at another distribution centre, and decide to visit the house of Jumbo top executive Frits van Eerd. They have a conversation with him and then leave.	Unknown	Occupation	2	Commercial private property, Personal private property	Over 100	Yes	Unknown	Yes, 10 farmers visiting the family home of Jumbo-director Frits van Eerd	Unknown	Yes, police officers send farmers away at a closed distribution center	Unknown	(NOS, 2020b; Trouw, 2020)
2021	7th July 2021	Farmer protests by various farmer organisations	Members of Agractie came to the Hague once again to protest on the Malieveld, but this time the atmosphere was more friendly. FDF protested on multiple other locations in the Netherlands. The atmosphere in the Hague seems peaceful.	Agractie and FDF	Field demonstration	5	Public space, Governmental institution	thousands	Yes	hundreds, 300 in The Hague and 500 in Zwolle	Unknown	Yes. More tractors enter Malieveld than allowed and cause damage to the grass.	Yes, police officers remove some tractors from the highway	1, a tractor drives into a cyclist and does not stop. This farmer is later arrested, the cyclist is fortunately unharmed.	(DeStentor, 2021; NieuweOogst, 2021; NOS, 2021)

4.2 Analysis of tweets

4.2.1 Best performing topic model

Table 4.2 shows the selected topic model for the complete dataset. The higher a word appears in a topic, the more significant it is to the subject. Three people labelled the topics: the two annotators and myself. I included all these topics in the table, and will first discuss the topics I identified and why. Then, I will reflect on the labels the annotators came up with.

The labels I gave each topic are the following:

1. Farmer protests in the Hague
2. Construction crisis
3. Relation between emissions, agriculture and nature
4. Role of the cabinet in the nitrogen crisis

The first topic contains the words 'Malieveld' and 'Hague', pointing at the Hague. Furthermore, it has various words which refer to demonstrations. There are four words containing 'protest', and the words 'action' and 'support'. Finally, it has the word 'tractor', a defining aspect of the farmers' demonstrations. Therefore, I labelled it 'Farmer protests in the Hague'.

The top words for the second topic are 'crisis', 'construction' and 'pfas', hinting toward the construction crisis. In line with this, the words 'housing' and 'home' are also present lower. The construction is broader than the nitrogen crisis because it is about more emissions than just nitrogen. The construction sector was already having trouble before the nitrogen crisis when restrictions were implemented for pfas, a toxic chemical that is emitted during construction work, and CO_2 ('pfas', 'co2' and 'emission' are present in the topic). However, some less relevant words in this topic, 'traffic', 'speed', 'home' and 'car', hint toward an alternative nitrogen reduction policy: the lowering of the highway speed limit.

Among the most relevant words of the third topic are 'emission', 'pfas', 'co2' and 'deposition', referring to different types of emissions. Then, the words 'nature', 'area'[24], 'climate', 'soil' and 'air' describe the Dutch environment. Finally, 'livestock', 'animal', 'agriculture', 'farming' and possibly 'soil' point toward Dutch agriculture. Therefore, this topic is labelled 'Relation between emissions, agriculture and nature', describing the broader context of the nitrogen crisis.

The fourth topic is more political, including the words 'policy', 'minister', 'government', 'cabinet', 'pas', 'vvd'[25], 'law' and 'debate'. Interestingly, while some of the main words in the other topics also occur in other topics (like 'nature'), topic 4 is the only topic containing the aforementioned

---

[24] 'area' likely refers to protected natural areas in the Netherlands, like natura2000 areas.
[25] A Dutch conservative political party.

political words. Besides, it contains several of the words more generally describing the nitrogen crisis, like ‘crisis’, ‘agriculture’ and climate. Therefore, it is interpreted as the role of the cabinet in the nitrogen crisis.

Even though I was able to label the four topics with distinct labels, it is noteworthy that there is overlap between the topics. Words like ‘crisis’, ‘protest’ and ‘netherlands’ occur multiple times in the topic model. This is likely because the dataset has an overarching topic: the nitrogen crisis.

Additionally, Table 4.2 shows the labels given by the two annotators. There is a noteworthy overlap between the labels given for topic 1 and topic 4. Furthermore, annotator 1 and me labelled the topics extremely similarly. This indicates that the best performing topic model has the potential of being labelled similarly by multiple people and that for labelling in-depth knowledge of the nitrogen crisis is not required. However, it is important to note that both annotators are friends of mine that I have interacted with during both my first and second thesis. Therefore, they can have been influenced by my biases and opinions on the crisis. Although asking them to annotate leads to promising labels, it is important to test topic models with multiple independent annotators before drawing final conclusions on the interpretability of the topics in the model.

Table 4.2 | Best performing topic model with NT = 4

ANNOTATOR	TOPIC 1	TOPIC 2	TOPIC 3	TOPIC 4
AUTHOR	FARMER PROTESTS IN THE HAGUE	CONSTRUCTION CRISIS	RELATION BETWEEN EMISSIONS, AGRICULTURE AND NATURE	ROLE OF THE CABINET IN THE NITROGEN CRISIS
ANN. 1	FARMER PROTEST	PFAS AND CONSTRUCTION	CLIMATE CRISIS AND ITS CAUSES (CONSTRUCTION AND FARMING)	GOVERNMENT'S ROLE IN THE CRISIS
ANN. 2	FARMER PROTEST AT MALIEVELD	POLICY FOR REDUCING CO2 EMISSIONS	POLICY FOR REDUCTION IN AGRICULTURE EMISSIONS	PROTESTS AGAINST POLICIES/CABINET FORMATION
Topic words	protest	crisis	emission	crisis
	farmersprotest	construction	pfas	nature
	farmers	pfas	nature	policy
	boerenprotest	due	co2	minister
	tractor	climate	much	government
	people	protest	crisis	cabinet
	hague	co2	deposition	pas
	police	housing	livestock	want
	action	road	netherlands	measure
	get	year	animal	agriculture
	like	cabinet	area	pfas
	netherlands	traffic	agriculture	doe
	way	people	year	protest
	support	netherlands	climate	must
	think	emission	doe	vvd
	farmersprotests	want	people	climate
	want	speed	soil	year
	malieveld	home	air	law
	right	new	farming	debate
	time	car	want	emission

Figure 4.1 shows the ratio of tweets belonging to each topic in Table 4.2. It shows that the most frequent topic is topic 1, the topic describing the farmers' demonstrations in the Hague. The least frequent topic is topic 2. It makes sense that fewer tweets are tagged with the construction crisis topic since there were fewer construction worker demonstrations than farmers' demonstrations. Topic 3 and topic 4 are equally frequent, with 24% of the tweets each.

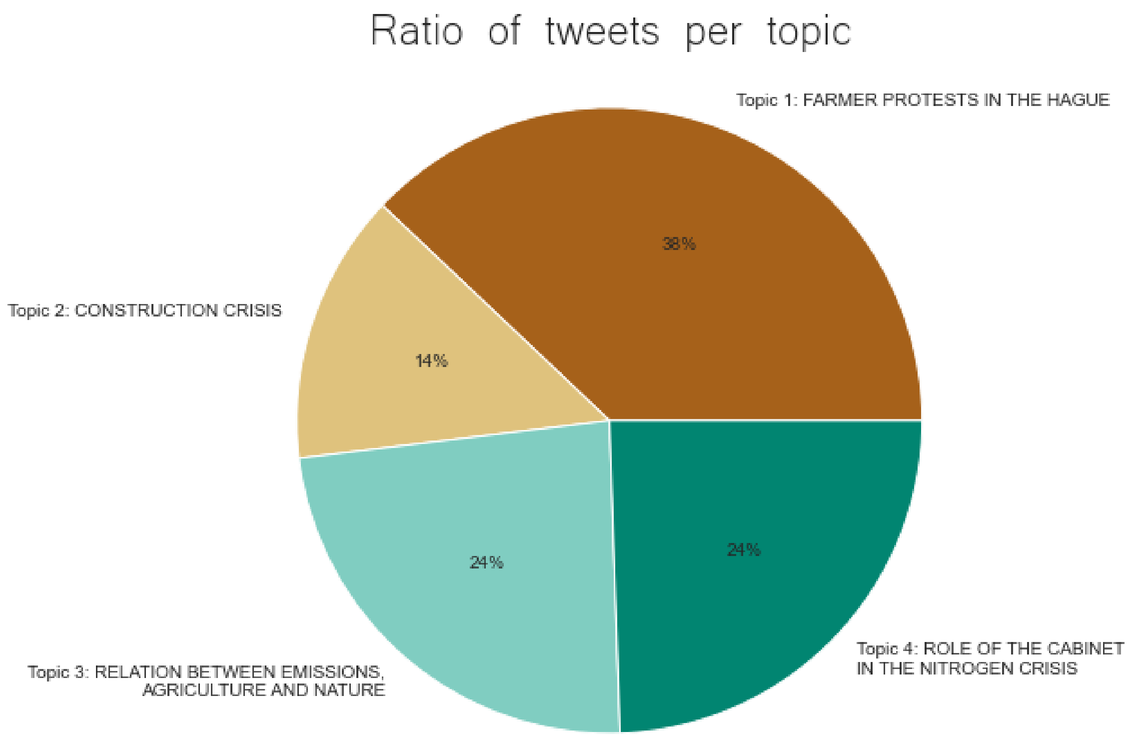

Figure 4.1 | The ratio of tweets per topic in the best performing topic model

4.2.2 Sentiment analysis and topic modelling combined

Figure 4.2 shows the mean positive, neutral and negative sentiment of each topic. It is evident that the average negativity score for tweets connected to each topic is significantly higher than the average positivity score. Topic 3 is the least negative and most positive: it has the lowest mean positive and the highest mean negative of the four subjects.

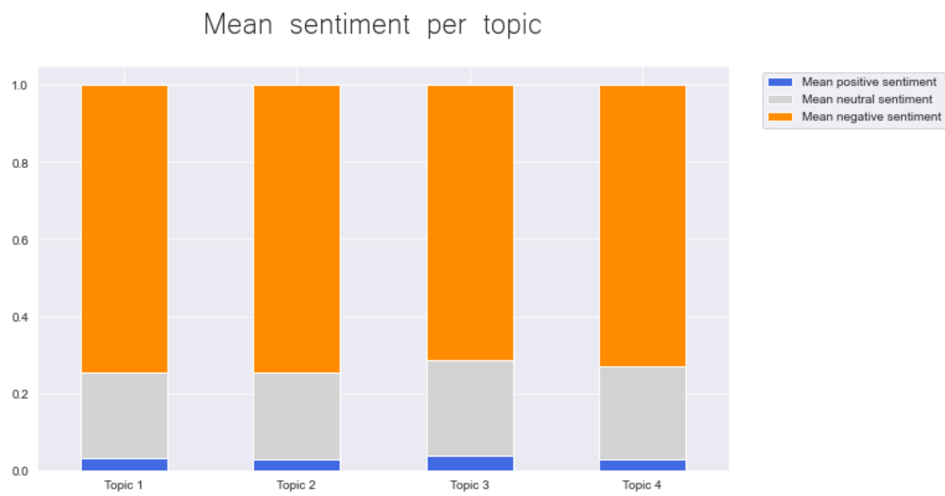

Figure 4.2 | Mean sentiment per topic in the best performing topic model

Figure 4.3 shows the combined outcomes of sentiment analysis and topic modelling. On top, in Figure 4.3.1, the number of tweets per week as a reference. The events from Table 4.1 are shown as tagged vertical lines. In the plot below that, the results of sentiment analysis are depicted. On the second graph, Figure 4.3.2, the weekly mean sentiment is plotted. The mean positive sentiment in blue, the mean negative sentiment in orange and the mean neutral sentiment in grey. The third graph, Figure 4.3.3, shows the weekly number of tweets per topic. Finally, to make the ratio of tweets per topic clearer, the graph on the bottom, Figure 4.3.4 shows the weekly ratio of tweets for each topic.

Sentiment analysis and topic modelling

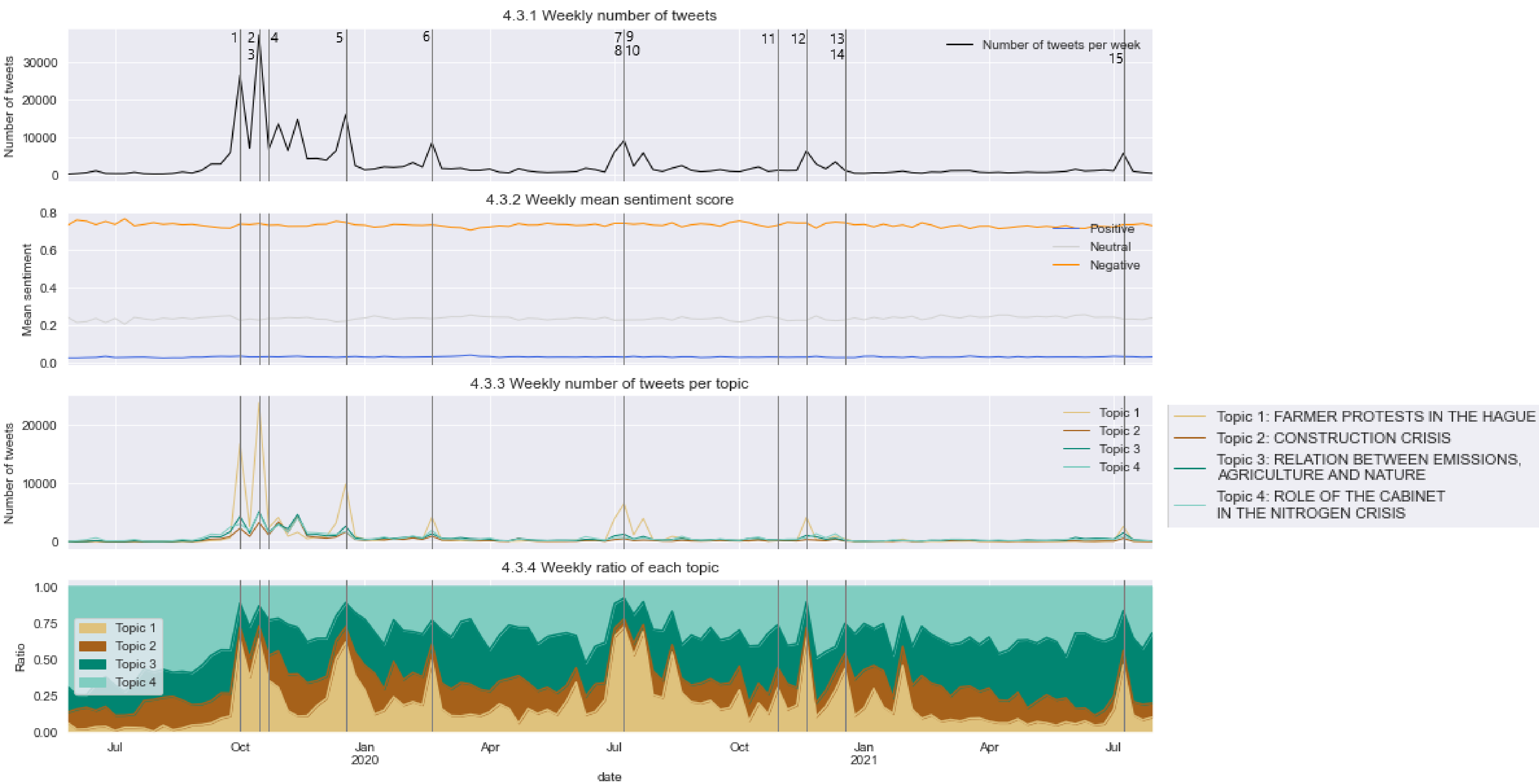

Figure 4.3 | Sentiment analysis and topic modelling results. From top to bottom: 4.3.1 Weekly number of tweets, 4.3.2 Weekly mean sentiment score, 4.3.3 Weekly number of tweets per topic and 4.3.4 Weekly ratio of each topic. Demonstrative events are plotted as numbers in 4.3.1.

Interestingly, what becomes evident when looking at the total volume of tweets in Figure 4.3.1 is that there is some relation between a peak in tweet volume and the identified events. During 12 out of the 15 identified events there is a peak in the number of tweets that week. In one week, for event 13, there is a slight peak just before the event. Also, Figure 4.3.2 and shows that tweets are immensely more negative than positive in mean sentiment, at all times. It follows that, according to the sentiment analysis results, tweets are more negative than positive throughout (nearly) the whole duration of the crisis, both in volume and in mean weekly sentiment. Furthermore, Figure 4.3.3 and Figure 4.3.4 show that topic 1, the topic that, as seen in Figure 4.1 above, has the most tweets assigned to it, is also dominant during all of the demonstrative events. Table 4.3 shows the numbers (IDs) that correspond to the events. These IDs are plotted in Figure 4.3.1 next to the vertical line that represents them. Given that subject 1 is about farmers' protests, it stands to reason that it would appear more frequently than other topics during significant farmers' protests. Topic 1 has been given a label that contains a location, The Hague, and Figure 4.3.3 shows that this topic has indeed the high peaks during the farmer demonstrations in the Hague (events 1, 6, 12 and 15). However, these peaks are not exclusive to demonstrations in the Hague, as they also appear during events 2, 5, 7-9, although not nearly as high. Only once is there a clear peak in a topic other than topic 1: after the fourth event, there is a peak in topic 3 in November 2019. This topic is about the relationship between emissions, agriculture and nature, in short: about the environmental background processes which caused the nitrogen crisis. An explanation for the emergence of this topic is that after the farmers drew nationwide attention to the nitrogen crisis with their first large-scale demonstrations, people started wondering and discussing online what caused the nitrogen crisis and whether claims made by farmers, that they are treated unfairly because other industries emit nitrogen too, were true. Apart from this exception, topics 2, 3 and 4 do not show any remarkable or dominant pattern.

Table 4.3 | IDs of each event, their name and the date they took place

ID	Event	Date
1	Demonstration on Malieveld	1th October 2019
2	Nationwide farmer protests	14th October 2019
3	Protests at RIVM in the Hague	16th October 2019
4	Farmer protests at province house and dairy firm	25th October 2019
5	Occupation of highways and supermarket distribution centers	18th December 2019
6	Farmer protests in the Hague	18th and 19th of February 2020
7	Farmers visit house of influential environmentalist to hand him a letter	6th July 2020
8	Occupation of supermarket distribution centre	7th July 2020
9	Various protests across 4 different provinces. A waste	8th July 2020

	disposal site is occupied. Minister Schouten cancels her visit to Zeeland due to security risks.	
10	Farmers blockade at Centraal Bureau Levensmiddelenhandel in Leidschendam.	11th July 2020
11	Farmer show up at home of politician	28th October 2020
12	Farmer protests in the Hague	17th November 2020
13	Highly criticised farmer protests and blockades	First half of December, 2020
14	Occupation of supermarket distribution centre and visiting of directors' house	14th December 2020
15	Farmer protests by various farmer organisations	7th July 2021

4.2.3 Topic models with a high number of topics

In section 3.4.2 I described how I selected the topic model by choosing for a coherent, interpretable topic model with stable topics. However, when looking at the topic models that had a higher number of topics but lower coherence scores, a few things stand out. While the coherent chosen topic model consists of coherent topics that I find easy to label, they describe quite general aspects of the nitrogen crisis. There is little surprising about the fact that a topic model with a relatively high coherence score contains topics that are about the farmers' demonstrations, the construction crisis, agriculture and the government. I cannot derive a political stance nor a real bias from the topic model with four topics. This changes when I look at topic models with a higher number of topics. Out of curiosity, I also analysed topic models with a higher number of topics.

As is to be expected with their lower coherence scores, topic models with a higher number of topics (NT) return fewer interpretable, though more detailed and biased topics. However, the few topics that I can interpret and label are interesting to look at: they offer a more biased perspective on elements of the nitrogen crisis. Other interpretable topics show topics that are more specific and detailed than the topics in the best-performing topic model. I selected a few of the ones I find striking and collected them in Table 4.4 below (the original complete topic models are shown in Appendix C: Topic models with higher numbers of topics).

From left to right, Table 4.4 first shows Topic 1 from the topic model with an NT of 20 which is labelled 'Demonstrations from a right-wing perspective'. It clearly describes the demonstrations, there is great overlap with the topic from the smaller topic model in Table 4.2 that received a similar label. However, the difference between the two topics is that this topic shows signs of political affiliation. The words 'fvd', 'pvv'[26], 'farmersrevolt', 'revolt',

[26] FvD and PVV are both Dutch conservative and right-wing populist Eurosceptic political parties

‘stemzeweg’[27] are affiliated with anger, specifically right-wing anger, the call for revolt and the skeptisism towards politicians who do not act the way FvD and PVV consider adequate.

Topic 11 from the same topic model shows the words ‘rivm’, ‘calculation’, ‘model’, ‘policy’, ‘report’, ‘measurement’, ‘figure’, ‘research’, ‘data’, ‘measure’, all words associated with measuring emissions, modelling and predicting what emission quantities would be harmful, and reporting on the findings as policy advice. This is one of the many tasks of the RIVM (the National Institute for Public Health and the Environment), which is why the topic is labelled ‘The role of RIVM’s models on nitrogen policy’. This subject is intriguing as it is really precise and in-depth regarding a specific organisation that is a crucial stakeholder in the nitrogen crisis.

In Topic 17, the words ‘speed’, ‘traffic’, ‘hour’, ‘maximum’, ‘car’, ‘road’, ‘jam’, ‘highway’, ‘drive’, ‘co2’ seem to refer to what the result could be of lowering the highway speed, namely less (CO_2 and nitrogen) emissions and fewer traffic jams. This policy was proposed in the Remkes report, and the word ‘remkes’ is also present in the topic (Remkes et al., 2020). I labelled the topic ‘Highway speed and traffic jams’. An alternative interpretation is that the traffic jam words refer to traffic jams caused by tractors on the highway. However, because there is no reference to protests of tractors, words that are common in other topics, I assume this is not the case.

From the bigger topic model, the topic model with 25 topics, topic 4 consists of words associated with aviation and the main Dutch airport, Schiphol: ‘aviation’, ‘traffic’, ‘Schiphol’, ‘industry’, ‘noise’, ‘co2’. During the nitrogen crisis there has been criticism on the fact that reductions in agriculture and road traffic emissions were discussed in the House of Representatives, but reductions in air traffic were not (BNNVARA, 2022). This topic seems to refer to the public discussion about this discrepancy.

Finally, topic 7 shows the signs of a typical topic about the farmers' protests in the Hague, similar to topic 1 in the same table and topic 1 in Table 4.2 above. However, this topic includes words like ‘support’, ‘good’, ‘trotsopdeboer’ (translation: proud of the farmer) and ‘proud’, which are absent in the other two topics. This indicates this topic picked up the pro-farmer perspective that was held by many farmers and spectators, especially during the first demonstrations. I am surprised not to have come across this topic earlier in smaller topic models, especially since it is often claimed farmers had high public support during the beginning of the crisis (Kalkhoven, 2021). I also find it surprising I barely came across any mentions of FDF in the topic models, since FDF is so prevalent in the news as the organiser of demonstrations. ‘fdf’ is present only twice in the topic models with higher NT (see Appendix C). In contrast ‘agractie’, another farmers’ organisation, is present twice as much in those topic models.

[27] Translation: Vote them away. Translated by the author.

Table 4.4 | Biased or detailed topics from topic models with a higher NT

Topic model with NT = 20			Topic model with NT = 25	
Topic 1	Topic 11	Topic 17	Topic 4	Topic 7
Demonstrations from a right-wing perspective	The role of RIVM's models on nitrogen policy	Highway speed and traffic jams	The role of aviation/Schiphol in nitrogen crisis	Demonstrations from a pro-farmer perspective
farmersprotest	rivm	crisis	emission	farmersprotest
protest	calculation	speed	co2	protest
farmers	deposition	traffic	aviation	boerenprotest
boerenprotest	model	due	air	tractor
support	question	hour	matter	support
citizen	policy	emission	particulate	hague
fvd	crisis	maximum	much	farmers
pvv	report	car	biomass	good
netherlands	emission	road	traffic	today
action	measurement	cabinet	gas	agractie
hague	figure	jam	schiphol	farmersprotests
agractie	doe	last	deposition	action
revolt	know	highway	power	malieveld
malieveld	government	drive	climate	people
farmersprotests	see	co2	industry	way
farmersrevolt	research	measure	calculation	like
campaign	data	remkes	doe	trotsopdeboer
climate	according	per	model	day
stemzeweg	protest	via	noise	netherlands
denhaag	measure	work	station	proud

Key Findings of Chapter 4: Results

1. News media coverage becomes less extensive after the first demonstrations.
2. From the very beginning, the farmers' demonstrations were marked by intimidation, vandalism, and threats.
3. Demonstrations often take place in clusters.
4. FDF was involved in organising nearly all of the demonstrations.
5. Negative sentiment is significantly higher than positive sentiment.
6. Although models with more topics tend to produce fewer interpretable topics, the ones that are interpretable occasionally contain greater details and indications of bias.

5 Discussion, limitations and future research

In this chapter, the results from the previous chapter are discussed in a larger context and interpreted. The meaning, importance and relevance of the results of this research are elaborated on. Then, the limitations of this research and recommendations for future research are formulated. This chapter follows the structure of the key insights listed in the Results chapter on the previous page one by one. Lastly, this chapter is closed with a section on the overarching limitations of this study and recommendations for future work.

5.1 News media coverage becomes less extensive after the first demonstrations

How social movements are depicted in news media is entirely dependent on the news media institutions (Amenta et al., 2017). Scale, participation and violence are factors that can raise the likelihood that demonstrations will receive media attention. (Amenta et al., 2017; Feinberg, Willer, & Kovacheff, 2017). As far as information is available on the number of attendees and the number of tractors present, the first four demonstrations are the only ones where these numbers reached in the thousands. It is possible that scale has been the driver of news media coverage. The largest demonstrations usually are few, but attract the majority of participants and are often well covered by news media (Amenta et al., 2017).

Information scarcity is very common in protest event analysis. Newspaper reporting is inevitably prone to the many-sided issues of selection bias (Hug, 2003; Lorenzini, Kriesi, Makarov, & Wüest, 2021; McCarthy, Titarenko, McPhail, Rafail, & Augustyn, 2008). The political stance of the newspaper and its geographic location, for instance, can affect what events are covered (Hug, 2003). Additionally, alongside political and social contexts these factors influence the way these events are reported: not only *what* is reported, but also *how* the events are framed (Amenta et al., 2017). In this research, I had no requirements for my news sources other than that they preferably came from a newspaper, no matter how big or small. I choose to work this way because, for some demonstrations, it was extremely hard to find articles and when I did, they were only reported on by small local newspapers. I even had to rely on pictures and, once, a youtube video to fill in as many categories as possible. In order to account for potential media biases without compromising information sources, I advocate evaluating potential biases of the resource in future research.

Finally, even though only a human can fill in the codebook, semi-automated approaches can be researched to speed up the process and support the coder. Selecting sources manually is an extremely time-consuming process and NLP has been demonstrated to be helpful in reducing newspaper selection bias (Hutter, 2014; Lorenzini et al., 2021). I propose to research semiautomated NLP tools to support news article selection.

5.2 From the very beginning, the farmers' demonstrations were marked by intimidation, vandalism, and threats

Although numerous sources report there was great public support for the farmers' demonstrations at the beginning of the crisis (Cornelisse, 2020; Kalkhoven, 2021), the nature of the farmers demonstrations was aggressive from the very beginning. The effectiveness of violence in social movements is a broadly studied topic (Shuman, Hasan-Aslih, van Zomeren, Saguy, & Halperin, 2022). This literature provides various answers to the question: Does the occurrence of violence during protests have a positive or negative effect on achieving the goals of the social movement? The answer largely depends on how the outcome's evaluation is defined (Shuman et al., 2022). When looking at public support, some studies show that violent protest practices can diminish public support (Feinberg et al., 2017). Considering the farmers indeed received heavy criticism for their disruptive tactics and public support was indeed reported to decrease as the crisis progressed, this may be a case where violent tactics have led to a reduction in public support (Cornelisse, 2020; Cornelisse & Klapwijk, 2020; Kalkhoven, 2021; NOS, 2020a, 2020b).

On the other hand, when looking at policy goals, there are notable examples that indicate that violence can increase support for a minority group and help to aid in the achievement of particular policy goals (Enos, Kaufman, & Sands, 2019). Such examples are observed during the nitrogen crisis as well when provinces withdraw their nitrogen policies directly after multiple demonstrations at province houses (Klumpenaar & Van Laarhoven, 2019; Van der Boon, 2019). The question arises as to whether these small successes weigh up against the significant loss of public support.

5.3 Demonstrations often take place in clusters

In literature, these clusters are referred to as 'protest cycles' (Minkoff, 1997). Actions are never isolated events, events influence one another, and so this intermittent nature of protest cycles is often observed in social movements (Minkoff, 1997; Oliver & Myers, 2003). Successful demonstrations often inspire more demonstrations, which might explain why the first four subsequent demonstrations were the largest scale events (Oliver & Myers, 2003). In this study, I only looked at demonstrations but did not formally document the external events sparking each protest cycle. For future work, it would be interesting to seek ways to document and characterise the 'sparking events' and add that information to the codebook.

This will add a new layer of information to the characterisation of demonstrations, describing not only *what* happened during demonstrative events, but also *why* they happened.

5.4 FDF was involved in organising nearly all of the demonstrations

This result indicates that FDF has been the farmers' organisation most active in organising demonstrations. FDF is known for being the most disruptive of the farmers´ organisations. Nevertheless, the selection bias of newspapers, often showing a preference for theatrical or violent demonstrations, could be the reason for this. The fact that FDF barely came up in the various topic models while Agractie has double the mentions implies that Agractie is more discussed in tweets than FDF. This is in stark contrast to how frequently FDF appears in newspaper articles. I cannot explain this result. To compare what is said on Twitter and in news articles better, for future research I propose to perform topic modelling on news articles as well. Then, for certain key actors, like farmers' organisations, political parties or politicians, it would be interesting to look at in what context they appear, and whether the topics they appear in show any bias. For this, I advise looking at topic models with higher NTs.

5.5 Negative sentiment is significantly higher than positive sentiment

The findings demonstrate that tweet sentiment is rather consistent: negativity is substantially higher at all times, while positivity is very low. As discussed in section 0, it was expected the would be more positivity during the first half-year of the crisis compared to later, as the crisis progressed and public support lowered. There are various possible explanations for this unexpected outcome. For starters, this could be because Twitter users tend to use a lot of negative language in their tweets. Various studies on positivity and negativity on Twitter conclude that negativity spreads much faster and broader than positivity (Schöne, Parkinson, & Goldenberg, 2021; Bellovary, Young, & Goldenberg, 2021; Jiménez-Zafra, Sáez-Castillo, Conde-Sánchez, & Martín-Valdivia, 2021). These studies look at the spread of sentiment of tweets from Twitter influencers or news media Twitter accounts, two of them in connection to an important political event. Although this study does not focus on the spread of sentiment, it would be interesting in future work to look at what Twitter accounts had the highest reach on Twitter, and whether the sentiment of the posts of these accounts can be connected to the higher negativity during the Nitrogen Crisis.

Secondly, the higher negativity of the dataset could be explained by the overarching topic of the tweets: the Nitrogen *Crisis*. During times of crisis, Twitter users may be more inclined to discuss the topic negatively than positively.

Since no patterns emerged in the sentiment analysis results in this study nor in my last thesis, the usefulness of applying sentiment analysis data to tweets about a policy crisis comes into question. I argue there is one more hypothesis to test before writing off sentiment analysis as a tool for this type of research: whether the sentiment analysis models I applied are not adequate for this dataset. There is very little information on how the sentiment model was established, other than what dataset was used to train it. Furthermore, it is not documented what 'positive', 'negative' and, most importantly, 'neutral' scores stand for, or how they are calculated in order to sum to one. In machine learning, context is important, and we can understand and use algorithms better if we understand how they were established.

If this information were available, the model could be taken as a base and trained further on nitrogen crisis tweets. This way, the model is altered to specifically calculate the sentiment in the nitrogen crisis dataset. However, to further train the model, annotated data is needed: tweets that are manually labelled with positive and negative scores. Establishing a training set like that is very time-consuming. Because better sentiment analysis results are not guaranteed, I propose to give research on the potential of topic modelling a higher priority than sentiment analysis.

5.6 Although models with more topics tend to produce fewer interpretable topics, the ones that are interpretable occasionally contain greater details and indications of bias

The topics shown in the best-performing topic model are quite general, they do not show surprising, niche or biased topics. As a consequence, one can wonder whether focussing on topic models with high coherence and stable topics are the most interesting. If I would want to know more about what perspectives the Twitter dataset has to offer on the various opinions and views on the nitrogen crisis, the bigger and less coherent topic models are more likely to provide information. Topic modelling traditionally focuses on creating high quality interpretable topic models, but for future research, I propose to explore the less coherent topic models with higher NTs. One of the biggest issues to overcome is how to filter out the incomprehensible topics and find the interpretable topics that show political stances or reveal specific trends during the nitrogen crisis. Also, creating bigger topic models runs the risk of generating arbitrary topics. To spare time, it would be ideal to find a common denominator for those topics and use it to make an initial selection that then can be labelled by a researcher.

Another approach for uncovering more niche topics is by slicing the dataset into smaller chunks. In this approach, more general topics, like Topic 1 about the Hague farmer protests, will not dominate lesser discussed topics, like those about smaller-scale protests. It should be noted that slicing the dataset in chunks, and thus creating more datasets and more topic models, will result in a significant increase in manual labour. In my previous thesis I propose a method for seleting the timeframe of a dataslice (Hendrikse, 2021).

Finally, since the labelling of topics is extremely prone to biases, I suggest researching how much prior knowledge and political affiliation influence how someone interprets and labels topics. For example, when filtering through the less coherent topics models, I noticed words that relate to right-wing radical politics catch my eye. Someone else could react differently to seeing the same topic model and notice other types of topics first. To produce dependable labels, I suggest having multiple annotators and study how to combine labels from different annotators to one label. Also, if for a certain topic there are too many differences between the labels from various annotators, this can serve as an indication that the topic is not coherent enough and should not be considered.

5.7 Limitations and recommendations

5.7.1 Tweet objects

The use of Twitter as a data source comes with several biases. First and foremost, only a select group of people use Twitter[28]. The demographics of Twitter users should be considered when interpreting the results of this research. Second, it is important to consider the varied motives behind tweeting. Twitter is used for expressing opinions, reporting on life events, spreading ideas and much more.

There are useful resources in tweets that this study has not used. First, hashtags emphasize words, therefore it is recommended to try giving higher weights to words that are preceded with a hashtag, and see if this improves the topic coherence of the generated topic models. Second, retweets were not used in this study, while it has been claimed that sharing tweets without adding text can be a sign of agreeing with the tweet (Goldenberg, et al., 2020). Third, many tweets contain URLs, often linking to articles on news websites. It could be interesting for decision-makers to map what news sources are spread most widely, to see where different groups on Twitter are getting news from. This could help in understanding where potential unrest originates.

5.7.2 Recommendations for characterising demonstrations

The way I created the codebook and characterised demonstrations can be further developed. First, there are instances of interdependence between some of the categories. For example, when the 'Location nr' is higher, the 'Nr of attendees' is automatically higher as well. This directly leads me to the next advancement: to avoid information loss, categories should be aggregated as little as possible. If demonstrations take place at multiple locations on the same day, collect information on each demonstration separately. This information can always be aggregated later on for providing a bigger picture.

Previously, in section 5.3, I proposed to develop a way to document what event(s) spark a protest cycle. In addition to this, I suggest researching ways to characterise and analyse the *outcome* of protest cycles as well. With this information combined, you can shape timelines and visualise what events led to demonstrations, what demonstrations led to what legislative

[28] Ironically, I myself am not on Twitter.

responses, etc. A valuable addition to the methodology of such research would be field surveys: going to demonstrations and asking participants what is going on and why they are demonstrating (Fisher et al., 2019).

Finally, I believe a productive continuation of my work is developing ways to visualise the content of the codebook. The table with demonstrations and their characteristics provides information in abundance, but taking this information in takes time. I believe a good addition is to find ways to combine some of the characteristics of demonstrations like, for example, threats, arrests and vandalism, into indicators for the scale and intensity of demonstrations.

6 Conclusion

This study aims to answer the question: *How does the intensity of emotions and discussed topics on social media relate to the characteristics of real-life demonstrations during a sustainability policy crisis?* This chapter will summarize the main findings of this study and answer the research questions.

SQ 1

How can sentiment analysis be applied for studying online social media activity?

Because in my previous research I was uncertain about the performance of the Dutch sentiment analysis tools available, in this research, I translated my dataset to English and applied a state-of-the-art sentiment analysis tool and plotted the mean positive, neutral and negative scores over time (Hendrikse, 2021). The results of the sentiment analysis show overwhelmingly negative tweets sentiment scores. This can be explained by the fact that the dataset consists of tweets about a *crisis*. However, it raises the question whether a ready-to-use sentiment analysis tool could be outperformed in picking up more subtle differences in sentiment in tweets. I recommend training a new sentiment analysis model based on annotated tweets from the dataset and see if the results are more balanced.

SQ 2

How can topic modelling be applied for studying online social media activity?

In contrast to my previous research, in this research, I looked at the content of topic models and labelled their topics. First, I defined evaluation steps for identifying the best performing topic model. I developed the topic model and labelled the topics. It turned out that this topic model only consists of general topics about the nitrogen crisis. It did not reveal any underlying trends nor did it show signs of political stance. Subsequently, the evaluation proposed guarantees coherent topics but filters details out. Additionally, there was no evidence that the sentiment of tweets per topic varied significantly.

One of the most promising findings is that topic models with higher numbers of topics but lower coherence scores prove to be more interesting sources of information on underlying topics and trends on Twitter. To build forth on this finding, future work will have to overcome the challenge of filtering out incoherent and meaningless topics in models with higher numbers of topics.

SQ 3

How can demonstrations be characterised?

A codebook was developed for characterising demonstrations. Several characteristics were defined with categories that, when filled in, give a picture of demonstrations. This codebook was developed with a focus on the scale, intensity and extremist nature of demonstrations. Before this study, studies on the Dutch nitrogen crisis were mostly empirical or based on interviews. This structural way of characterising demonstrations using news articles will prove useful in expanding our understanding of protests during the nitrogen crisis as well as for future sustainability policy crises.

MQ

How does the intensity of emotions and discussed topics on social media relate to the characteristics of real-life demonstrations during a sustainability policy crisis?

Unexpectedly, it is not sentiment analysis or topic modelling results that have the most obvious connection with the events identified: it is an increase in tweets during events. The combination of sentiment analysis and topic modelling as implemented in this research does not show interesting patterns when compared to the insights from the codebook. However, because there is an extensive amount of research applying these methods to social media data around various political events with valuable results, I argue that more experiments need to be performed with this approach before drawing final conclusions on the usefulness of the combination of sentiment analysis and topic modelling. In this report, many approaches for such improvements are suggested.

With this semi-automated approach, I aimed to look at the dynamics of online social media activity and real-life political and demonstrative events. However, with sentiment analysis results being extremely constant throughout time and topics, the results did not provide additional insights into the demonstrations and the topics discussed. I hypothesised that looking at the results of sentiment analysis and, especially, topic modelling would deepen the knowledge of someone who knows little about the nitrogen crisis. Ironically, the results suggest the opposite: extensive knowledge of the nitrogen crisis and its events seems required to train and interpret the topic modelling results and distinguish between general and niche topics.

Key Findings of Chapter 6: Conclusion

1. Sentiment analysis results are overwhelmingly negative. This can be explained by the overarching topic: the nitrogen *crisis*. This result shows applying ready-to-use sentiment analysis tools might lead to poor results. I suggest to specifically train a sentiment analysis model for the dataset at hand to pick up on the possible subtilities of sentiment on social media during a policy crisis

2. Topic models with higher number of topics but with a lower coherence show a greater potential for providing insights into the diversity of topics discussed on social media. In order to use these topic models for future research, new metrics for evaluating the usefulness of topic models need to be developed.

3. Comparing sentiment analysis and topic modelling results to the characterisations of demonstrations in the codebook shows no overlapping patterns. Extensive knowledge of the nitrogen crisis and its events seems required to train and interpret the topic modelling results and distinguish between general and niche topics

7 | Appendix

A. Plot of volume of tweets per month

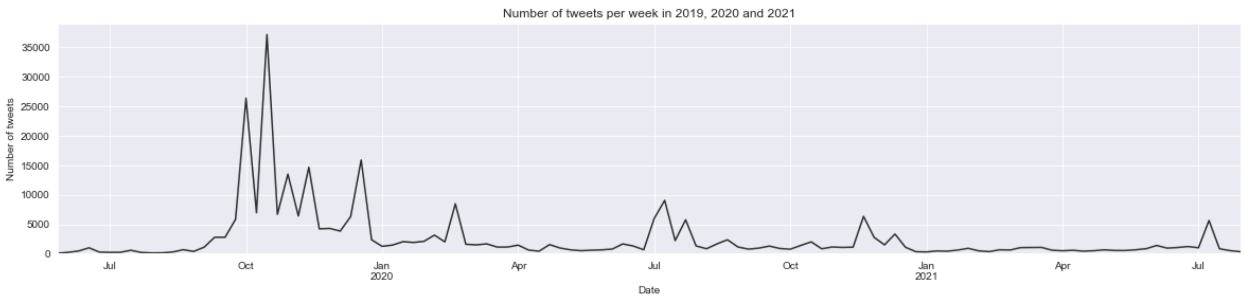

Figure 7.1 | the number of tweets about the Nitrogen Crisis per month, from July 2019 up to July 2021.

B. List of stop words

Manually added:

http
user
nitrogen
farmer
problem
and
the
that
also

Stopwords from nltk:

i
me
my
myself
we
our
ours
ourselves
you
you're
you've
you'll
you'd
your
yours
yourself
yourselves
he
him
his
himself
she
she's
her
hers
herself
it
it's
its
itself
they
them
their
theirs
themselves
what
which
who
whom
this
that
that'll
these
those
am
is
are
was
were
be
been
being
have
has
had
having
do
does
did
doing
a
an
the
and
but
if
or
because
as
until
while
of
at
by
for
with
about
against
between
into
through
during
before
after
above
below
to
from
up
down
in
out
on
off
over
under
again
further
then
once
here
there
when
where
why
how
all
any
both
each
few
more
most
other
some
such

no
nor
not
only
own
same
so
than
too
very
s
t
can
will
just
don
don't
should
should've
now
d
ll
m
o
re
ve
y
ain
aren
aren't
couldn
couldn't
didn
didn't
doesn
doesn't
hadn
hadn't
hasn
hasn't
haven
haven't
isn
isn't
ma
mightn
mightn't
mustn
mustn't
needn
needn't
shan
shan't
shouldn
shouldn't
wasn
wasn't
weren
weren't
won
won't
wouldn
wouldn't

C. Topic models with higher numbers of topics

Table 7.1 | Topic model with 16 topics

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16
protest	crisis	nature	farmersprotest	protest	construction	emission	crisis	pfas	emission	construction	crisis	policy	rivm	crisis	protest
farmersprotest	climate	deposition	animal	farmersprotest	earthmoving	crisis	farmersprotest	water	co2	crisis	housing	nature	emission	vvd	farmersprotest
farmers	netherlands	emission	protest	tractor	van	speed	good	soil	nature	pfas	climate	crisis	protest	rutte	people
boerenprotest	pfas	area	livestock	hague	road	co2	nature	standard	plant	due	care	minister	calculation	cda	police
action	co2	agriculture	food	farmers	protest	car	protest	substance	air	cabinet	etc	pas	model	party	farmers
support	people	livestock	farmers	boerenprotest	pfas	traffic	year	pfos	biomass	year	pfas	schouten	policy	d66	right
citizen	country	netherlands	want	traffic	work	measure	climate	chemical	matter	housing	tax	agriculture	deposition	protest	tractor
fvd	get	much	meat	police	new	maximum	time	via	gas	standard	shortage	state	farmersprotest	want	terror
agractie	like	crisis	agriculture	malieveld	mechanic	want	think	drinking	particulate	protest	rutte	cabinet	question	fvd	get
hague	corona	farming	farming	road	vacancy	cabinet	like	toxic	much	new	education	measure	know	farmersprotest	action
pvv	stop	doe	cow	due	service	make	see	foam	climate	sector	people	government	doe	debate	like
malieveld	policy	must	agricultural	via	company	due	people	environment	tree	project	money	law	much	cabinet	think
farmersprotests	everything	reserve	feed	today	machine	livestock	solution	fire	crisis	want	policy	approach	would	left	government
last	come	natura2000	netherlands	jam	transport	reduction	one	non	forest	government	corona	council	crisis	minister	want
netherlands	year	policy	farm	way	amp	road	would	health	power	policy	government	want	report	pvv	boerenprotest
revolt	protest	agricultural	get	morning	young	must	policy	harmful	soil	netherlands	groningen	province	see	policy	really
hour	left	measure	price	highway	malieveld	remkes	today	pan	due	house	health	new	say	netherlands	support
campaign	think	ammonia	people	groningen	field	work	pfas	chemours	year	build	youth	plan	government	baudet	way
tractor	government	sector	product	center	project	per	know	pollution	ammonia	home	year	via	measure	climate	would
farmersrevolt	etc	country	dutch	construction	engineering	drive	make	company	deposition	company	co2	must	think	say	farmersprotests

Table 7.2 | Topic model with 20 topics

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
farmersprotest	protest	crisis	co2	animal	protest	farmersprotest	construction	pfas	crisis	rivm	pfas	emission	crisis	nature	nature	crisis	netherlands	protest	vvd
protest	farmersprotest	climate	emission	farmersprotest	farmersprotest	good	pfas	standard	nature	calculation	water	livestock	housing	area	pas	speed	want	farmers	cda
farmers	people	pfas	climate	protest	tractor	protest	protest	soil	agriculture	deposition	substance	deposition	pfas	deposition	minister	traffic	get	farmersprotest	protest
boerenprotest	climate	policy	biomass	food	farmers	year	earthmoving	construction	policy	model	soil	agriculture	care	crisis	policy	due	country	boerenprotest	party
support	right	netherlands	matter	farmers	hague	people	due	fire	solution	question	chemical	nature	climate	soil	state	hour	farmersprotest	news	farmersprotest
citizen	terror	government	particulate	meat	police	like	road	new	agricultural	policy	year	much	shortage	reserve	law	emission	people	van	crisis
fvd	like	cabinet	crisis	livestock	boerenprotest	boerenprotest	project	van	livestock	crisis	much	netherlands	etc	natura2000	cabinet	maximum	crisis	hague	d66
pvv	think	year	air	feed	malieveld	see	crisis	foam	sector	report	air	farming	rutte	forest	schouten	car	come	via	want
netherlands	get	construction	gas	cow	traffic	think	new	water	want	emission	harmful	ammonia	year	plant	council	road	going	action	rutte
action	netherlands	protest	plant	price	action	time	sector	via	climate	measurement	pfos	year	people	due	measure	cabinet	think	crisis	fvd
hague	really	rutte	power	want	way	today	work	substance	government	figure	non	doe	tax	tree	government	jam	like	fdf	debate
agractie	left	people	aviation	product	road	would	malieveld	want	measure	doe	drinking	area	corona	biodiversity	permit	last	would	today	left
revolt	see	co2	energy	farm	groningen	know	company	environmental	must	know	emission	traffic	house	netherlands	crisis	highway	much	read	right
malieveld	action	country	noise	agricultural	center	day	transport	read	make	government	co2	measure	education	much	emission	drive	make	tractor	minister
farmersprotests	peasant	nature	station	farming	highway	still	housing	question	good	see	like	industry	billion	natura	approach	co2	emission	policy	say
farmersrevolt	police	want	much	subsidy	people	nice	home	municipality	emission	research	health	sector	policy	veluwe	must	measure	one	minister	pvv
campaign	government	via	tree	land	via	many	mechanic	netherlands	plan	data	pan	crisis	home	land	take	remkes	really	new	baudet
climate	time	time	wood	dutch	army	one	machine	extinguishing	approach	according	toxic	co2	cabinet	specie	province	per	time	live	boerenprotest
stemzeweg	farmerterror	due	pollution	get	government	well	builder	project	farming	protest	environment	per	money	bird	want	via	let	schouten	member
denhaag	one	make	environment	milk	today	really	vacancy	environment	pas	measure	food	reduce	construction	animal	rule	work	pfas	agricultural	think

Table 7.3 | Topic model with 25 topics

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20	Topic 21	Topic 22	Topic 23	Topic 24	Topic 25
protest	netherlands	crisis	emission	farmersprotest	protest	farmersprotest	farmersprotest	nature	crisis	crisis	pfas	nature	construction	protest	livestock	policy	crisis	deposition	pas	crisis	fvd	farmersprotest	protest	crisis
farmers	protest	climate	co2	protest	farmersprotest	protest	animal	crisis	speed	rivm	water	crisis	pfas	tractor	agriculture	crisis	housing	nature	state	year	pvv	protest	minister	netherlands
farmersprotest	farmersprotest	co2	aviation	agricultural	climate	boerenprotest	get	area	emission	protest	substance	measure	earthmoving	farmersprotest	emission	government	climate	emission	permit	hour	vvd	people	van	rutte
hague	country	pfas	air	animal	left	tractor	protest	land	car	pfas	soil	policy	standard	traffic	farming	want	pfas	area	council	last	cda	police	schouten	policy
action	people	people	matter	netherlands	right	support	people	want	due	good	standard	area	road	hague	animal	protest	care	much	ruling	protest	citizen	tractor	farmers	people
police	one	netherlands	particulate	farmers	news	hague	food	build	maximum	policy	pfos	cabinet	project	boerenprotest	nature	province	shortage	soil	nature	farmersprotest	stemzeweg	terror	action	pfas
construction	climate	year	much	subsidy	boerenprotest	farmers	cow	tree	co2	see	chemical	pas	due	road	crisis	house	etc	netherlands	project	week	outdoors	like	fdf	want
tractor	day	like	biomass	company	baudet	good	meat	house	cabinet	agriculture	toxic	deposition	sector	jam	feed	cda	rutte	ammonia	policy	time	d66	think	farmersprotest	government
malieveld	pfas	time	traffic	money	crisis	today	want	netherlands	want	think	fire	law	new	malieveld	sustainable	cabinet	people	due	new	day	party	right	boerenprotest	protest
army	time	get	gas	billion	farmers	agractie	like	construction	traffic	solution	environment	minister	soil	farmers	measure	rule	tax	agriculture	emission	trending	votefvd	get	via	cabinet
defense	every	come	schiphol	price	d66	farmersprotests	would	new	measure	farmersprotest	drinking	must	work	police	farmersprotest	brabant	corona	year	construction	question	groundinresistance	really	crisis	country
center	dutch	country	deposition	million	wing	action	boerenprotest	forest	make	make	foam	remkes	company	due	schouten	provincial	construction	doe	approach	hashtag	farmersrevolt	farmerterror	read	doe
via	get	going	power	dutch	people	malieveld	good	building	road	government	non	approach	crisis	groningen	sector	vvd	year	many	decision	morning	pvda	farm	rutte	right
people	like	everything	climate	food	fvd	people	think	reserve	drive	would	via	debate	malieveld	live	take	construction	policy	high	pfas	already	vote	would	member	farmersprotest
boerenprotest	farmers	hysteria	industry	land	like	way	come	plant	work	read	harmful	government	mechanic	morning	solution	pfas	education	effect	government	due	administrator	time	agriculture	left
netherlands	see	corona	calculation	pay	dwdd	like	know	people	highway	know	health	agriculture	protest	highway	space	party	due	per	due	pfas	trotsoponzeboeren	boerenprotest	mark	party
demonstration	crisis	think	doe	year	nieuwsuur	trotsopdeboer	eat	year	let	like	co2	want	via	way	agricultural	year	co2	livestock	law	ago	boerenprotest	action	new	climate
support	germany	etc	model	sector	nos	day	still	solution	remkes	climate	pan	doe	cabinet	via	via	measure	home	water	airport	construction	farmersprotest	way	news	think
campaign	sense	change	noise	want	amp	netherlands	day	million	must	model	pollution	say	vacancy	drive	manure	doe	cabinet	cause	environmental	one	boomer	one	agractie	get
want	think	still	station	euro	see	proud	milk	stop	reduction	thing	ban	question	builder	city	cow	new	immigration	reserve	may	per	wilders	say	agricultural	like

8 References

AD. (2019, December 16). Rutte : Holocaust- vergelijking indringend besproken tijdens boerenoverleg Catshuis, pp. 1–12. Retrieved from https://www.ad.nl/politiek/rutte-holocaust-vergelijking-indringend-besproken-tijdens-boerenoverleg-catshuis~add5d7b4/

AD. (2021). Boer gaat in hoger beroep tegen straf voor boerenprotest Groningen. Retrieved from https://www.ad.nl/drenthe/boer-gaat-in-hoger-beroep-tegen-straf-voor-boerenprotest-groningen~afe5bdab/

Allcott, H., & Gentzkow, H. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/https://doi.org/10.1257/jep.31.2.211

Amenta, E., Elliott, T. A., Shortt, N., Tierney, A. C., Türkoğlu, D., & Vann, B. (2017). From bias to coverage: What explains how news organizations treat social movements. *Sociology Compass*, *11*(3), 1–12. https://doi.org/10.1111/soc4.12460

Araujo, M., Pereira, A., Reis, J., & Benevenuto, F. (2016). An evaluation of machine translation for multilingual sentence-level sentiment analysis. *Proceedings of the ACM Symposium on Applied Computing*, *04-08-Apri*, 1140–1145. https://doi.org/10.1145/2851613.2851817

Balahur, A., & Turchi, M. (2012). Comparative experiments for multilingual sentiment analysis using machine translation. *CEUR Workshop Proceedings*, *917*(July), 75–86.

Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. (2020). TWEETEVAL: Unified benchmark and comparative evaluation for tweet classification. *Findings of the Association for Computational Linguistics Findings of ACL: EMNLP 2020*, 1644–1650. https://doi.org/10.18653/v1/2020.findings-emnlp.148

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford: Oxford University Press.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 993–1022.

BNNVARA. (2020). Extremistische boeren voeren intimidatie op. Retrieved from https://www.bnnvara.nl/joop/artikelen/extremistische-boeren-voeren-intimidatie-op

BNNVARA. (2022, July 5). Johan Vollenbroek : ' Boeren hebben een punt , Schiphol moet ook halveren .' Retrieved from https://www.bnnvara.nl/joop/artikelen/johan-vollenbroek-boeren-hebben-een-punt-schiphol-moet-ook-halveren

Bosma, A., & Peeren, E. (2021). #Proudofthefarmer: Authenticity, populism and rural masculinity in the 2019 Dutch farmers' protests. In P. Pospěch, E. M. Fuglestad, & E. Figueiredo (Eds.), *Politics and Policies of Rural Authenticity*. https://doi.org/10.4324/9781003091714-10

Bouwmeester, R. (2020, May 7). Terrorismebestrijder: "Optreden FDF draagt bij aan polarisatie." *Nieuwe Oogst*, p. 9. Retrieved from https://www.nieuweoogst.nl/nieuws/2020/05/07/terrorismebestrijder-optreden-fdf-draagt-bij-aan-polarisatie

Brantly, A. F. (2019). From Cyberspace to Independence Square: Understanding the Impact of Social Media on Physical Protest Mobilization During Ukraine's Euromaidan Revolution. *Journal of Information Technology and Politics*, *16*(4), 360–378. https://doi.org/10.1080/19331681.2019.1657047

Burns, A., & Eltham, B. C. (2009). Twitter free Iran: An evaluation of Twitter's role in public diplomacy and information operations in Iran's 2009 election crisis.

Cambria, E., Gastaldo, P., & Bisio, F. &. (2015). An ELM-based model for affective analogical reasoning. *Neurocomputing*, *149*, 443–455. https://doi.org/https://doi.org/10.1016/j.neucom.2014.01.064

Casquete, J. (2006). The Power of Demonstrations. *Social Movement Studies*, *5*(1), 45–60. https://doi.org/10.1080/14742830600621183

Cornelisse, D. (2020). Steeds minder steun voor boerenprotesten want ' het is nu wel duidelijk wat ze willen '. Retrieved February 11, 2022, from https://eenvandaag.avrotros.nl/panels/opiniepanel/alle-uitslagen/item/steeds-minder-steun-voor-boerenprotesten-want-het-is-nu-wel-duidelijk-wat-ze-willen/

Cornelisse, D., & Klapwijk, P. (2020, July 14). Mensen zijn er klaar mee : kleine meerderheid wil dat boerenprotesten nu echt stoppen. Retrieved from https://eenvandaag.avrotros.nl/item/mensen-zijn-er-klaar-mee-kleine-meerderheid-wil-dat-boerenprotesten-nu-echt-stoppen/

Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A. Y. A., Gelbukh, A., & Zhou, Q. (2016). Multilingual Sentiment Analysis: State of the Art and Independent Comparison of Techniques. *Cognitive Computation*, *8*(4), 757–771. https://doi.org/10.1007/s12559-016-9415-7

DeStentor. (2021, July 7). Boeren weer op weg naar huis: protesten trokken duizenden deelnemers. Retrieved from https://www.destentor.nl/home/boeren-weer-op-weg-naar-huis-protesten-trokken-duizenden-deelnemers~af71f498/?referrer=https%3A%2F%2Fwww.foodlog.nl%2F

Drevon, J. (2016). Embracing Salafi Jihadism in Egypt and Mobilizing in the Syrian Jihad. *Middle East Critique*, *25*(4), 321–339. https://doi.org/10.1080/19436149.2016.1206272

Driessen, P. (2019, December 13). Grootste actieboer kijkt trots terug op protest rond Eindhoven Airport : ' Generale repetitie is geslaagd .' *AD*. Retrieved from https://www.ad.nl/eindhoven/grootste-actieboer-kijkt-trots-terug-op-protest-rond-eindhoven-airport-generale-repetitie-is-geslaagd~a4583246/?utm_source=twitter&utm_medium=social&utm_campaign=socialsharing_web

Enos, R. D., Kaufman, A. R., & Sands, M. L. (2019). Can Violent Protest Change Local Policy

Support? Evidence from the Aftermath of the 1992 Los Angeles Riot. *American Political Science Review*, *113*(4), 1012–1028. https://doi.org/10.1017/S0003055419000340

Fang, J. (2016). In Defense of Hashtag Activism. *Journal of Critical Scholarship on Higher Education and Student Affairs*, *2*(1). Retrieved from https://ecommons.luc.edu/jcshesa/vol2/iss1/10

Feinberg, M., Willer, R., & Kovacheff, C. (2017). Extreme Protest Tactics Reduce Popular Support for Social Movements. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2911177

Fillieule, O. (2012). The independent psychological effects of participation in demonstrations. *Mobilization*, *17*(3), 235–248. https://doi.org/10.17813/maiq.17.3.h10631822v65820l

Fisher, D. R., Andrews, K. T., Caren, N., Chenoweth, E., Heaney, M. T., Leung, T., … Pressman, J. (2019). The science of contemporary street protest: New efforts in the United States. *Science Advances*, *5*(10), 1–16. https://doi.org/10.1126/sciadv.aaw5461

Frouws, J. (1993). *Mest en macht. Een politiek-sociologische studie naar belangenbehartiging en beleidsvorming inzake de mestproblematiek in Nederland vanaf 1970*. LUW, Wageningen.

Gelderlander. (2020a). ‘Achterhoekse FDF- voorman’ zit nog vast voor dreigen in te rijden op agent die daarna wapen trok. Retrieved from https://www.gelderlander.nl/oost-gelre/achterhoekse-fdf-voorman-zit-nog-vast-voor-dreigen-in-te-rijden-op-agent-die-daarna-wapen-trok~ac93f2bb/?referrer=https%3A%2F%2Fnl.wikipedia.org%2F

Gelderlander. (2020b). Boerenbezetting van distributiecentrum AH in Geldermalsen voorbij. Retrieved from https://www.gelderlander.nl/ede/boerenbezetting-van-distributiecentrum-ah-in-geldermalsen-voorbij~af5d82b4/

Gelderlander. (2020c). Rob Jetten spreekt schande van boerenactie tegen Johan Vollenbroek : ‘ Onacceptabele intimidatie .’ Retrieved from https://www.gelderlander.nl/nijmegen/rob-jetten-spreekt-schande-van-boerenactie-tegen-johan-vollenbroek-onacceptabele-intimidatie~a73ab7ab/

Gelderlander. (2021). Felle discussie tussen boeren en ‘ stikstofstrijder ’ Vollenbroek onder toeziend oog politie in Nijmegen-Noord. Retrieved from https://www.gelderlander.nl/oude-ijsselstreek/felle-discussie-tussen-boeren-en-stikstofstrijder-vollenbroek-onder-toeziend-oog-politie-in-nijmegen-noord~a7c32254/?referrer=https%3A%2F%2Ft.co%2F

Giachanou, A., & Crestani, F. (2016). Like It or Not: A Survey of Twitter Sentiment Analysis Methods. *ACM Computing Surveys*, *49*(2), 1–41. https://doi.org/https://doi.org/10.1145/2938640

González-Ibáñez, R., Muresan, S., & Wacholder, N. (2011). Identifying sarcasm in Twitter: A closer look. *ACL-HLT 2011 - Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, *2*(2010), 581–586.

Grasso, M. T., & Giugni, M. (2016). Protest participation and economic crisis: The conditioning role of political opportunities. *European Journal of Political Research*, *55*(4), 663–680. https://doi.org/10.1111/1475-6765.12153

Grover, P., Kar, A. K., Dwivedi, Y. K., & Janssen, M. (2019). Polarization and acculturation in US Election 2016 outcomes – Can twitter analytics predict changes in voting preferences.

Technological Forecasting and Social Change, *145*(August 2018), 438–460. https://doi.org/10.1016/j.techfore.2018.09.009

Guzman, E., Ibrahim, M., & Glinz, M. (2017). A Little Bird Told Me: Mining Tweets for Requirements and Software Evolution. *Proceedings - 2017 IEEE 25th International Requirements Engineering Conference, RE 2017*, (3), 11–20. https://doi.org/10.1109/RE.2017.88

Hakkenes, E. (2019, December 5). De boeren dreigen de voedselvoorziening rond Kerstmis lam te leggen. *Trouw*. Retrieved from https://www.trouw.nl/binnenland/de-boeren-dreigen-de-voedselvoorziening-rond-kerstmis-lam-te-leggen~b44e7ef4/?referrer=https%3A%2F%2Fwww.google.com%2F

Harlow, S., & Guo, L. (2014). Will the revolution be tweeted or facebooked? Using digital communication tools in immigrant activism. *Journal of Computer-Mediated Communication*, *19*(3), 463–478. https://doi.org/10.1111/jcc4.12062

HartvanNederland. (2020). Politie : demonstraties in Den Haag en Leidschendam rustig verlopen, (november), 1–31.

Hendrikse, M. (2021). Unravelling Twitter chaos during a policy crisis (Master's Thesis). Retrieved from http://resolver.tudelft.nl/uuid:3289b69e-efea-412e-93d7-ecf5fa1d9834

Hermenet, C. (2020, June 20). Radicale boer vindt moeizaam politiek onderdak, pp. 2–3. Retrieved from https://www.digibron.nl/viewer/collectie/Digibron/id/tag:RD.nl,20200620:newsml_pr15910819

Hong, S., & Kim, S. H. (2016). Political polarization on twitter: Implications for the use of social media in digital governments. *Government Information Quarterly*, *33*(4), 777–782. https://doi.org/https://doi.org/10.1016/j.giq.2016.04.007

Howard, P. N., Agarwal, S. D., & Hussain, M. M. (2011). When do states disconnect their digital networks? Regime responses to the political uses of social media. *Communication Review*, *14*(3), 216–232. https://doi.org/10.1080/10714421.2011.597254

Hug, S. (2003). Selection Bias in Comparative Research: The Case of Incomplete Data Sets. *Political Analysis*, *11*(3), 255–274. https://doi.org/10.1093/pan/mpg014

Hutter, S. (2014). Protest Event Analysis and Its Offspring (pp. 335–367). https://doi.org/10.1093/acprof:oso/9780198719571.003.0014

Janssen, H. (2019, December 13). "Boeren net als Joden gedeporteerd", voorman Farmers Defence Force blijft bij omstreden uitspraak. *Omroep Brabant*. Retrieved from https://www.omroepbrabant.nl/nieuws/3121852/Boeren-net-als-Joden-gedeporteerd-voorman-Farmers-Defence-Force-blijft-bij-omstreden-uitspraak

Jonsson, E., & Stolee, J. (2016). An Evaluation of Topic Modelling Techniques for Twitter. *ArXiv*, 1–11.

Jurafsky, D., & Martin, J. (2020). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. (third ed.)*. Upper Saddle River, N.J: Pearson Prentice Hall.

Jusoh, S., & Alfawareh, H. M. (2011). Automated translation machines: Challenges and a proposed solution. In *2011 2nd International Conference on Intelligent Systems, Modelling and Simulation, ISMS 2011* (pp. 77–82). IEEE. https://doi.org/10.1109/ISMS.2011.74

Kalkhoven, L. (2021). *Media-analyse Boerenprotesten*. The Hague. Retrieved from https://www.weerbaar-bestuur.nl/sites/default/files/inline-files/Rapport Media-analyse boerenprotesten.pdf

Kim, S., Weber, I., Wei, L., & Oh, A. (2014). Sociolinguistic analysis of Twitter in multilingual societies. *HT 2014 - Proceedings of the 25th ACM Conference on Hypertext and Social Media*, 243–248. https://doi.org/10.1145/2631775.2631824

Klandermans, B., Crozat, M., Hipsher, P. L., Katzenstein, M. F., Keck, M. E., Kubik, J., … Others. (1997). *The social movement society: Contentious politics for a new century*. Rowman \& Littlefield Publishers.

Klandermans, B., van Stekelenburg, J., & van der Toorn, J. (2008). Embeddedness and identity: How immigrants turn grievances into action. *American Sociological Review*, *73*(6), 992–1012. https://doi.org/10.1177/000312240807300606

Klumpenaar, S., & Van Laarhoven, K. (2019). Provincies zwichten voor boerenprotest: geen nieuwe stikstofregels voor boeren. *Retrieved October*, *1*, 2021. Retrieved from https://www.nrc.nl/nieuws/2019/10/14/ook-provincie-drenthe-schort-stikstofregels-op-na-boerenprotest-a3976663

Lau, J. H., Baldwin, T., & Newman, D. (2013). On collocations and topic models. *ACM Transactions on Speech and Language Processing*, *10*(3). https://doi.org/10.1145/2483969.2483972

Lee, C., Shin, J., & Hong, A. (2018). Does social media use really make people politically polarized? Direct and indirect effects of social media use on political polarization in South Korea. *Telematics and Informatics*, *35*(1), 245–254. https://doi.org/https://doi.org/10.1016/j.tele.2017.11.005

Lee, F. (2016). Impact of social media on opinion polarization in varying times. *Communication and the Public*, *1*(1), 56–71. https://doi.org/https://doi.org/10.1177/2057047315617763

Leidsch Dagblad. (2020). en bij CBL in Leidschendam Gebruik cookies en Heijn in Zwolle met tientallen. Retrieved from https://www.leidschdagblad.nl/cnt/dmf20200710_50940994?utm_source=google&utm_medium=organic

Looijenga, M. (2018). The Detection of Fake Messages using Machine Learning. *Retrieved September*, *1*, 2021. Retrieved from http://essay.utwente.nl/77385/

Lorenzini, J., Kriesi, H., Makarov, P., & Wüest, B. (2021). Protest Event Analysis: Developing a Semiautomated NLP Approach. *American Behavioral Scientist*. https://doi.org/10.1177/00027642211021650

Mäntylä, M. V, Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis---A review of research topics, venues, and top cited papers. *Computer Science Review*, *27*, 16–32. https://doi.org/https://doi.org/10.1016/j.cosrev.2017.10.002

McCarthy, J. D., Titarenko, L., McPhail, C., Rafail, P. S., & Augustyn, B. (2008). Assessing stability in the patterns of selection bias in newspaper coverage of protest during the transition from communism in Belarus. *Mobilization*, *13*(2), 127–146. https://doi.org/10.17813/maiq.13.2.u45461350302663v

Miao, L., Last, M., & Litvak, M. (2020). Twitter Data Augmentation for Monitoring Public Opinion on Covid-19 Intervention Measures. In *Proceedings of the 1st Workshop on NLP for Covid-19 (Part 2) at EMNLP 2020*. Association for Computational Linguistics.

Minkoff, D. C. (1997). The sequencing of social movements. *American Sociological Review*, *62*(5), 779–799. https://doi.org/10.2307/2657360

Mohammad, S. M., Salameh, M., & Kiritchenko, S. (2016). How translation alters sentiment. *Journal of Artificial Intelligence Research*, *Journal of*, 95–130. https://doi.org/10.1613/jair.4787

Mutz, D. C., & Mondak, J. J. (2006). The workplace as a context for cross-cutting political discourse. *The Journal of Politics*, *68*(1), 140–155.

Nandathilaka, M., Ahangama, S., & Weerasuriya, G. T. (2018). A Rule-based Lemmatizing Approach for Sinhala Language Maheshi. In *2018 3rd International Conference on Information Technology Research (ICITR),* (pp. 1–5). IEEE. https://doi.org/10.1109/ICITR.2018.8736134

NieuweOogst. (2021, July 7). Terugblik op boerenprotest. Retrieved from https://www.nieuweoogst.nl/nieuws/2021/07/07/terugblik-op-boerenprotest

Nieuws, R. (2019). Boerenprotest op 18 december: "Genoeg gepraat." *Retrieved September*, *10*, 2021. Retrieved from https://www.rtlnieuws.nl/nieuws/nederland/artikel/4950316/boerenprotest-boerenacties-stikstofbeleid-supermarkt

NOS. (2019a). ' Monumentale deur uit provinciehuis breken is een grens overschrijden '. Retrieved from https://nos.nl/artikel/2306147-monumentale-deur-uit-provinciehuis-breken-is-een-grens-overschrijden

NOS. (2019b). "Aantal violisten op Binnenhof is minder intimiderend dan duizend trekkers." Retrieved from https://nos.nl/nieuwsuur/artikel/2306448-aantal-violisten-op-binnenhof-is-minder-intimiderend-dan-duizend-trekkers

NOS. (2019c). Boerenprotest Den Haag voorbij, 2200 trekkers terug naar huis. Retrieved from https://nos.nl/liveblog/2304125-boerenprotest-den-haag-voorbij-2200-trekkers-terug-naar-huis

NOS. (2019d). Organisatie kijkt tevreden terug op boerenprotestdag : ' grandioos verlopen '. Retrieved from https://nos.nl/liveblog/2306317-organisatie-kijkt-tevreden-terug-op-boerenprotestdag-grandioos-verlopen

NOS. (2020a, December 13). Burgemeester Zwolle wil landelijke regels voor boerenprotesten : ' Grens overschreden '. Retrieved from https://nos.nl/artikel/2360505-burgemeester-zwolle-wil-landelijke-regels-voor-boerenprotesten-grens-overschreden

NOS. (2020b, December 14). Blokkade boeren van distributiecentrum Jumbo Breda , waren bij huis Jumbo-directeur. Retrieved from https://nos.nl/artikel/2360556-blokkade-boeren-van-distributiecentrum-jumbo-breda-waren-bij-huis-jumbo-directeur

NOS. (2021, July 7). Nieuw boerenprotest, landbouw in bepaalde gebieden verder in het nauw. Retrieved from https://nos.nl/artikel/2388278-nieuw-boerenprotest-landbouw-in-bepaalde-gebieden-verder-in-het-nauw

NRC. (2020). Leden Farmers Defence Force brengen Rob Jetten thuis voedselpakket. Retrieved from https://www.nrc.nl/nieuws/2020/10/28/vijf-leden-farmers-defence-force-brengen-rob-jetten-thuis-voedselpakket-a4017685

NU.nl. (2019). LTO Noord veroordeelt incident boerenprotest Groningen : ' zaak voor politie '. Retrieved from https://www.nu.nl/binnenland/6004098/lto-noord-veroordeelt-incident-boerenprotest-groningen-zaak-voor-politie.html?redirect=1

Ofori-Parku, S. S., & Moscato, D. (2018). Hashtag Activism as a Form of Political Action: A Qualitative Analysis of the #BringBackOurGirls Campaign in Nigerian, UK, and U.S. Press. *International Journal of Communication*, *12*(0), 23.

Oliver, P., & Myers, D. (2003). The Coevolution Of Social Movements. *Mobilization: An International Quarterly*, *8*(1), 1–24. https://doi.org/10.17813/maiq.8.1.d618751h524473u7

Omroep Brabant. (2019). 2000 boeren bij provinciehuis in Den Bosch , het protest in zeven foto ' s. Retrieved from https://www.omroepbrabant.nl/nieuws/3094865/2000-boeren-bij-provinciehuis-in-den-bosch-het-protest-in-zeven-fotos

Omroep West. (2020). ' Het is oorlog ', boeren onderweg naar Den Haag , legertrucks blokkeren weg naar Binnenhof. Retrieved from https://www.omroepwest.nl/nieuws/3998818/het-is-oorlog-boeren-onderweg-naar-den-haag-legertrucks-blokkeren-weg-naar-binnenhof

Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You*. London: Penguin. https://doi.org/10.5070/D482011835

Pluimveebedrijf. (2019, December 5). Grote voedselblokkade aangekondigd door boeren en bouwers. Retrieved from https://www.pluimveebedrijf.nl/huisvesting-pluimveehouderij/bedrijfsnieuws/grote-voedselblokkade-aangekondigd-door-boeren-en-bouwers/

Poria, S., Hazarika, D., Majumder, N., & Mihalcea, R. (2020). Beneath the Tip of the Iceberg: Current Challenges and New Directions in Sentiment Analysis Research. *IEEE Transactions on Affective Computing*, *3045*(c), 1–29. https://doi.org/10.1109/TAFFC.2020.3038167

Pörtner, H. O., Roberts, D. C., Poloczanska, E. S., Mintenbeck, K., Tignor, M., Alegría, A., … (eds.), A. O. (2022). *IPCC, 2022: Summary for Policymakers*. *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. https://doi.org/10.4324/9781315071961-11

Remkes, J. W., Dijk, J. J. van, Dijkgraaf, E., Freriks, A., Gerbrandy, G. J., Maij, W. H., … Vet, L. E. M. (2020). Niet alles kan overal.

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *WSDM 2015 - Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, 399–408. https://doi.org/10.1145/2684822.2685324

RTLNieuws. (2020). Boeren met trekkers toch op snelweg: "Never give up." *Retrieved October*. Retrieved from https://www.rtlnieuws.nl/nieuws/nederland/artikel/5027231/boerenprotest-tractor-trekker-snelweg

RTV Drenthe. (2020). Jaar van boerenprotesten: "Actievoeren zit niet in ons bloed, maar we moeten wel" - RTV Drenthe, 1–11. Retrieved from https://www.rtvdrenthe.nl/nieuws/166206/Jaar-van-boerenprotesten-Actievoeren-zit-niet-in-ons-bloed-maar-we-moeten-wel

Schelfaut, S. (2019). Stikstofplan ingetrokken na felle acties van boeren. *Retrieved October, 1*, 2021. Retrieved from https://www.ad.nl/binnenland/stikstofplan-ingetrokken-na-felle-acties-van-boeren

Schussman, A., & Soule, S. A. (2005). Process and protest: Accounting for individual protest Participation. *Social Forces*, *84*(2), 1083–1108. https://doi.org/10.1353/sof.2006.0034

Shirky, C. (2011). The Political Power of Social Media: Technology, the Public Sphere, and Political Change. *Foreign Affairs*, *90*(1), 28–41. Retrieved from http://www.jstor.org/stable/25800379

Shuman, E., Hasan-Aslih, S., van Zomeren, M., Saguy, T., & Halperin, E. (2022). Protest movements involving limited violence can sometimes be effective: Evidence from the 2020 BlackLivesMatter protests. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(14), 1–12. https://doi.org/10.1073/pnas.2118990119

Smith, L. G. E., McGarty, C., & Thomas, E. F. (2018). After Aylan Kurdi: How Tweeting About Death, Threat, and Harm Predict Increased Expressions of Solidarity With Refugees Over Time. *Psychological Science*, *29*(4), 623–634. https://doi.org/10.1177/0956797617741107

Spohr, D. (2017). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review*, *34*(3), 150–160. https://doi.org/https://doi.org/10.1177/0266382117722446

Starbird, K., & Palen, L. (2012). (How) Will the Revolution be Retweeted ? Information Diffusion and the 2011 Egyptian Uprising. In *CSCW '12: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 7–16). https://doi.org/10.1145/2145204.2145212

Steinskog, A. O., Therkelsen, J. F., & Gambäck, B. (2017). Twitter Topic Modeling by Tweet Aggregation. In *Proceedings of the 21st Nordic Conference of Computational Linguistics* (pp. 77–86).

Stokstad, E. (2019). Nitrogen crisis threatens Dutch environment—and economy. *Science*, *366*(6470), 1180–1181. https://doi.org/10.1126/science.366.6470.1180

Sunstein, C. R. (2018). *#Republic: Divided Democracy in the Age of Social Media*. . Princeton: Princeton University Press.

Tausch, N., Becker, J. C., Spears, R., Christ, O., Saab, R., Singh, P., & Siddiqui, R. N. (2011). Explaining Radical Group Behavior: Developing Emotion and Efficacy Routes to Normative and Nonnormative Collective Action. *Journal of Personality and Social Psychology*, *101*(1), 129–148. https://doi.org/10.1037/a0022728

Terpilowski, M. (2022). bitermplus. Retrieved from https://bitermplus.readthedocs.io/en/latest/index.html

The AI Community Building the Future. (n.d.). Retrieved from https://huggingface.co/

Thelwall, M., Buckley, K., Paltoglou, G. C., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, *61*(12), 2544–2558.

Trouw. (2020, December 14). Opnieuw boerenprotest : demonstranten blokkeren distributiecentrum Breda. Retrieved from https://www.trouw.nl/binnenland/opnieuw-boerenprotest-demonstranten-blokkeren-distributiecentrum-breda~be83d9ed/

Tufekci, Z., & Wilson, C. (2012). Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square. *Journal of Communication*, *62*(2), 363–379. https://doi.org/10.1111/j.1460-2466.2012.01629.x

Tumasjan, A., Sprenger, T., Sandner, P., & Welpe, I. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. *ICWSM*, *10*(1), 178–185.

Van der Boon, V. (2019, October 14). Vier provincies zwichten voor boze boeren. *Financieel Dagblad*. Retrieved from https://fd.nl/economie-politiek/1320153/boze-boeren-voor-acht-provinciehuizen

Van der Ploeg, J. D. (2020). Farmers' upheaval, climate crisis and populism. *Journal of Peasant Studies*, *47*(3), 589–605. https://doi.org/10.1080/03066150.2020.1725490

Van Rooijen, L. (2020, July 17). Van den Oever: 'Acties worden hard en FDF-style.' *Food & Agribusiness*. Retrieved from https://www.foodagribusiness.nl/van-den-oever-acties-worden-hard-en-fdf-style/

Vayansky, I., & Kumar, S. A. P. (2020). A review of topic modeling methods. *Information Systems*, *94*, 101582. https://doi.org/10.1016/J.IS.2020.101582

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.

Waikhom, L., & Goswami, R. S. (2019). Fake News Detection Using Machine Learning. *SSRN Electronic Journal*, 1076–2787. https://doi.org/https://doi.org/10.2139/ssrn.3462938

Weerd, T. De. (2021). Seeking More Appreciation by Demonstrating: Unappreciation and Farmers Protest (Bachelor's Thesis).

West, O. (2020). Twee aanhoudingen tijdens boerenprotest in Den Haag, boeren rijden toeterend langs Huis ten Bosch. *Retrieved October*, *1*, 2021. Retrieved from https://nos.nl/artikel/2360505-burgemeester-zwolle-wil-landelijke-regels-voor-boerenprotesten-grens-overschreden

Wijnants, R. (2020). Minister Schouten staakt werkbezoek Zeeland wegens dreiging boeren. *Retrieved September*, *8*, 2021. Retrieved from https://joop.bnnvara.nl/nieuws/minister-schouten-staakt-werkbezoek-zeeland-wegens-dreiging-boeren

Winterman, P. (2019). Voorman boeren moet lachen om grafkist met Jesse Klavers naam erop. *Retrieved August*, *26*, 2021. Retrieved from https://www.ad.nl/politiek/voorman-boeren-moet-lachen-om-grafkist-met-jesse-klavers-naam-erop

Xiaohui Yan, Jiafeng Guo, Yanyan Lan, X. C. (2013). A Biterm Topic Model for Short Texts. In *Proceedings of the 22nd international conference on World Wide Web* (pp. 1445–1455). https://doi.org/10.1145/2488388.2488514

Youtube. (2019). Boerenprotest 18 December 2019. Retrieved from https://www.youtube.com/watch?v=vBFwp1BQvUY&ab_channel=Gelderdrone

Zhang, W., Xu, M., & Jiang, Q. (2018). Opinion Mining and Sentiment Analysis in Social Media: Challenges and Applications. In *HCI in Business, Government, and Organizations : 5th International Conference, HCIBGO 2018 Held as Part of HCI International 2018 Las Vegas, NV, USA, July, 2018 Proceedings. 10923. Cham* (pp. 15–20). https://doi.org/https://doi.org/10.1007/978-3-319-91716-0_43

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., & Li, X. (2011). Comparing twitter and traditional media using topic models. In *ECIR'11 . , Heidelberg: Berlin Heidelberg"* (pp. 338–349). Berlin: Springer.