Public opinion on Twitter

A case study on palm oil

Sabrine van Rossum

TU Delft

Public opinion on Twitter

A case study on palm oil

Master thesis submitted to Delft University of Technology in partial fulfilment of the requirements for the degree of

Master of Science
in **Engineering and Policy Analysis**

by

Sabrine van Rossum
Student number: 4151976

To be defended in public on November 29, 2018

First Supervisor:	Dr. S. W. Cunningham	Policy Analysis
Second Supervisor:	Dr. U. Pesch	Values, Technology and Innovation

TU Delft

Preface

This report presents the results of six months' devoted research, completed at Delft University of Technology, in partial fulfillment of a masters degree in Engineering and Policy Analysis. This thesis is a combination of two complimentary fields I am primarily interested in; data science and philosophy. Whilst I have learned a lot during my extensive research, I equally realize there are endless learning opportunities in the academic world and in life in general. As I write the final sentences of my thesis, I am conceptualizing that my student days are nearing to an end. Whilst nostalgic, I am equally excited and ready to begin the next chapter of life.

I am grateful for all those who have supported me along the journey of my thesis. Firstly, I would like to thank Scott Cunningham, who has been integral in organizing code and analyzing my tweets. Further to this, his tips and comments have assisted my writing capabilities to make my report more readable. I will miss our insightful and energetic meetings, discussing topics ranging from code to the future of social media. Secondly, I would like to thank Udo Pesch who introduced me to Habermas. His help was vital when I got lost in the complexity of the Habermasian public sphere. I appreciate all the time Udo dedicated to our meetings, including the additional hours of preparation required. The conversations were fruitful and in-depth, and his feedback and critical questions kept me on track. Thirdly, I would like to thank all my friends, fellow students and especially my family. My father Theo has been invaluable, and it is thanks to our insightful conversations that my perspective on relevant topics has thrived. Additionally, my brother Mike has provided insights in Python, making my life easier by enhancing my understanding. Last but not least, my mother Linda has been always there with the greatest confidence in me.

Sabrine van Rossum
Delft, November 2018

Executive summary

The world faces many challenges today, from climate change to geopolitical tensions and from food security to artificial intelligence. These societal and environmental issues of the 21st century can be classified as wicked problems (Duckett et al., 2016) or grand challenges (Grandis and Efstathiou, 2016). Wicked problems and grand challenges can be argued to have multiple similar characteristics, including incomplete, contradictory, and changing requirements that are often difficult to recognize. Additionally, there is no indisputable public good nor single definition of equity in society, meaning there is no optimal solution to these problems possible (Rittel and Webber, 1973).

Many of these challenges are incorrigible and extremely challenging for policymakers. A better understanding of public opinion can be helpful for policymakers in dealing with challenges. Ultimately, policy is made for the public, also in the case of wicked problems and grand challenges. Some policymakers already take public opinion data into account when making policy: research shows the correlation between public opinion and policy response (Barabas, 2011; Monroe and Alan, 1998; Page and Shapiro, 1983; Pierson, 2002; Stimson et al., 1995). Public opinion formation has changed over the last centuries due to the differences in life styles and standards. Before the invention of the telephone, the telegram and the television, all interactions were face to face. The internet was invented almost thirty years ago and is argued to drastically change the way we shape and share opinions. The internet today, with ample amount of data available, can be a useful source to analyze public opinion.

A lot of information is posted on social media every second. However, policymakers hardly use social media data because they often do not know how to use it, do not see the added value or do not know what is actually being discussed. This research aims to give policymakers insight into how to better include public opinion data from social media in policy making. The main line of reasoning is based on the German philosopher Jürgen Habermas and his book *The Structural Transformation of the Public sphere* (1962). Habermas has written extensively on the public sphere, "a realm of our social life, in which something approaching public opinion can be formed" (Habermas et al., 1974, 49). This public sphere should not be seen as a market nor a coffeehouse, but as a transcending space, an abstract forum for dialogue and public opinion, for vivid debates on multiple levels in society.

The theory of the Habermasian public sphere is applied to Twitter. To do so, the traditional literature on public opinion, based on the Habermasian public sphere, is combined with more recent literature on social media. After the literature research the main knowledge gap is identified. This gap is the lacking relation between the philosophical investigation on public opinion on Twitter and an empirical investigation on actual tweets. This thesis combines the fields of philosophy and data science to contribute to the knowledge gap. Palm oil is used as a case study. The main research question is:

How can public opinion on Twitter be characterized, on the case study of palm oil?

To characterize means to describe the special, typical or telling qualities. There are four sub-questions to answer the main research question:

1. *What are possible characteristics of public opinion on Twitter?*
2. *What are relevant tweets regarding palm oil?*
3. *What characterizes public opinion on palm oil on Twitter?*
4. *What can be generalized about public opinion on Twitter from the case study of palm oil?*

The answer to the first sub-question is based on a literature review to define possible characteristics on Twitter. Possible characteristics have been proposed in three directions, all under the coat rack of the Habermasian public sphere. Firstly, the egalitarian characteristics of the public sphere; secondly, the fragmentation of the public sphere; and thirdly, dynamically emerging and dissolving publics around specific issues. The answer to the second sub-question is based on the methodology of data gathering. By using different keywords and hashtags related to palm oil, 103.500 tweets have been gathered and 65.417 unique users have been identified over a 50-day period. The answer to the third sub-question combines the possible characteristics from sub-question 1 with the tweets gathered from sub-question 2. The answer to the fourth sub-question broadens the results from sub-question 3 and generalizes the conclusion on palm oil to general characteristics. Through the four sub-questions, the main research question is answered:

The findings of this research suggest how Twitter, in its current form, does not create the optimal environment for forming something that is approaching public opinion due to five characteristics: the exclusion of more than half the world population (a), the domination by a few users (b1) (often with strategic interest (b2)) , the pollution through spam (c), Twitter's role (d) and the characteristics of tweets (e).

a. The exclusion of more than half the world population from social media is at least based on the lack of internet access, new media literacy and cultural capital. Less than 5% of the world population is on Twitter, which is a lot less than half the world population. There are many more factors why some cannot or do not want to participate. Anyone with reason and the willingness to join was argued to be welcome to join the debate in the Habermasian public sphere. The internet has for some made it more accessible to join public debate, and for some more inaccessible. To conclude, even though social media is still not for everyone, there has never been a medium before that connects so many people around the world. Twitters reach is broader and crosses more borders than any other possible incarnation of the Habermasian public sphere ever had or did before.

b.1. The dominating key players are users who get the most attention from other users, policymakers and journalists. This can either be on a specific topic, or transcending topics. The findings from this research suggest that organizations like Greenpeace, BBC Earth and the Roundtable on Sustainable Palm Oil have an above average role in guiding the thinking of the public and therewith steering public opinion. It is difficult to find a user that is saying something completely different than what the defined influencers are saying. These few key players play an above average role in defining public opinion on Twitter, and the influence of this role is not limited to Twitter. To clarify, their influence can lead to a spillover of topics discussed on Twitter to topics discussed outside Twitter. And vice versa: influential actors outside Twitter often automatically have a higher status on Twitter and therewith have more influence. Policymakers have indicated in the conducted interviews to mostly read tweets of those users influential outside Twitter, such as politicians and prominent journalists. To conclude, the role of influential players turns Twitter into a public sphere similar to the Habermasian public sphere, which was an arena of the bourgeoisie. The white, middle class male took an above average role in guiding the thinking of the public. The findings of this research suggest that on Twitter it is not the bourgeoisie but ((non-)governmental) organizations and media that have an above average role in guiding the thinking of the public.

b.2. The tweets analyzed show how the dominant key players often act with a strategic interest in mind. The public sphere should, according to Habermas, be free of power and ideology and the only common interest should be the interest in good opinion formation. It can be argued that those acting from a strategic interest do not have good opinion creation as their main interest. It is hard to distinguish private interest from strategic interest, and as was the case in the Habermasian public sphere, all citizens are private persons bringing in private arguments. But when the arguments brought in are only beneficial for a specific organization, the question is how much these opinions (should) add to public opinion. To clarify,

many actors in the debate on palm oil are multinationals like PepsiCo and Nestle. By creating a Twitter account and interacting with other users, these companies are given a face and therewith blur the lines between the private and the public sphere. Their interest is, however, still to benefit the company. When you would, hypothetically, not take any organization into account (which would drastically decrease the number of users), the problem of individuals representing organizations remains. An example is the previous CEO of PepsiCo, tweeting from her personal account strategic messages about the good intentions of PepsiCo for our planet. To conclude, times have changed since the Habermasian public sphere is developed, and in our current global capitalistic world, many companies and their spokespeople play a big role in influencing public opinion. Leaving them out will leave us with an even less representative public opinion, since they have become part of public opinion formation. We have to be aware of an increasing amount of strategic actors present on Twitter as a public sphere, who dominate public debate and therewith influence public opinion.

c. The pollution through spam is argued to play a serious role in influencing public opinion on Twitter. It is unclear how big the role is, but as can be seen in the case study of palm oil, fake accounts and hyperactive automated users are active. Fake accounts retweet specific tweets and mention certain usernames, to make these accounts seem much more popular than they actually are. An example of fake accounts in the case study of palm oil is illustrated by over 2000 retweets for an Indonesian palm oil investment fund. The accounts that retweeted the investment funds' tweets were deleted not long after they retweeted, by Twitter. The platform marked all their accounts as they were deleted. Hyperactive automated users tweet many similar messages in a very short period of time and subsequently violate the official Twitter limit. An example of a hyperactive automated user, in the case study of palm oil, is a certain user who tweeted nine times in one minute -containing nine slightly different tweets, all with the message to boycott an Indonesian palm oil company- and therewith exceeded the official Twitter limit of less than two tweets per minute (Twitter, 2018a). To conclude, the exact effects of fake accounts and hyperactive automated users is unclear but it can be argued, as can be seen from the examples, that they cause a stream of misinformation and therewith influence public opinion.

d. Twitters algorithms possibly play a role in public opinion on the platform. Technologies are never neutral and McLuhans 'the medium is the message' can be argued to be applicable. The medium influences how the message is perceived. In the case of Twitter, it does not only influence the perception but also what is communicated. Twitter, together with other internet giants like Google and Facebook, can be argued to be the gatekeeper of the internet. They control which ideas and opinions are being shared. Due to their nontransparent algorithms it is very difficult to decipher the exact influence these powerful entities have on public opinion. To conclude, it cannot be proved what Twitter is doing, and if and how they have influenced the debate on palm oil, since the algorithm is nontransparent. Examples like Morozov's accusation of censorship in Twitter's 'trending algorithm' shows the influence they can have.

e. Tweets are often retweets, contain hashtags, calls for actions, snippets of information, interpreted data and images or videos. Retweets can be argued to be echo chambers where users only retweet others with similar ideas. Hashtags can amplify any discussion and calls for action are frequently found on Twitter. A direct call for action asks people to take *direct* action. This can be the question to retweet your tweet, sign a petition or ask for a donations. All these calls of action occur frequently in the case study of palm oil. Of the most frequently retweeted tweets in this thesis' database, 57% contained a direct call of action by either signing a petition or asking for a donation. Snippets of information spoon-feed information while pretending they feed the whole story. However, it is difficult to capture a complicated story in less than 280 characters. To reinforce a message, interpreted data is added. Data is, like a technology, never neutral but it can be used to influence the public of something by using it the right way. Of the most frequently retweeted tweets, 86% uses interpreted data to get a point across. Lastly, *a picture is worth more than a thousand words* explains why

many Tweets show photos of dying orangutans, bulldozers cutting down rain forest or cute animation videos. Of the most frequently retweeted tweets, 57% contains a picture and 43% contains a video. To conclude, tweets are not just 280 characters. They are often carefully considered combinations of hashtags, calls for action, snippets of information, interpreted data and images or videos. These characteristics all help in convincing the reader of the point the user is trying to make.

The answer to the main research question shows how Twitter, in its current form, is not the optimal place for public opinion formation. Therefore, using Twitter data in policy making should be reconsidered. When all the limitations of social media data are taken into account and policymakers still want to use the data, the Python code used to gather and analyze tweets can be found on `https://github.com/sevanrossum/`. The insights from social media data can be used in different stages of the policy cycle. It can be used in the agenda setting stage to put issues on the agenda that are frequently discussed on social media. In the policy formulation stage, preferences on different policy alternatives can be analyzed using social media data. In the policy implementation stage, the level of satisfaction can be measured based on expression on social media. Lastly, in the policy evaluation stage, feedback expressed on social media on a specific policy can be used as an alternative feedback model. Not everyone likes direct feedback so analyzing feedback expressed as an opinion on social media can be a valuable addition.

This research is part of the Engineering and Policy Analysis masters program. This program deals with the grand challenges of our age and this thesis deals with the grand challenge of social media. It is a grand challenge for many reasons, including lies, propaganda, fake news and an imbalance of power. Social media is chosen as a challenge because besides a grand challenge, it can also be seen as a grand opportunity. When social media data is used correctly, it can possibly generate insight in public opinion. This, in turn, can be used by policymakers in the decision making process on other grand challenges.

This research is based on simplifications and interpretations of the extensive theory by Habermas on the public sphere. It followed a framework I developed to turn general philosophical theory into measurable constructs. Tweets are gathered and analyzed based on these constructs, to see if these tweets can be analyzed in the light of the Habermasian public sphere. The novelty of this research and therewith the main contribution to scientific literature is in the elaborate discussion and exploration, based on actual tweets, of characteristics of Twitter as an Habermasian public sphere.

Twitter and the Habermasian public sphere both do not go uncriticized, but they are chosen in this research because the Habermasian public sphere can be argued to be a good lens through which Twitter can be analyzed. Twitter itself is still of societal relevance in public opinion formation. Other choices, or other lines of reasoning, would have been possible and might have led to different outcomes. The answer to the main question, however, is more robust since most users defining the debate are not only tweeting on palm oil but on many different topics. Further research is recommended for philosophers, data scientists and mostly policy researchers.

Contents

List of Figures

List of Tables

1

Introduction

The world faces many challenges today, from climate change to geopolitical tensions and from food security to artificial intelligence. These issues of the 21st century concern a broad range of societal and environmental issues and can be classified as wicked problems (Duckett et al., 2016) or grand challenges (Grandis and Efstathiou, 2016). The term *wicked problems* was first coined by Rittel and Webber (1973). Wicked problems have different characteristics, including incomplete, contradictory, and changing requirements that are often difficult to recognize. Additionally, there is no indisputable public good nor single definition of equity in society, meaning there is no optimal solution to these problems possible (Rittel and Webber, 1973). *Grand challenges* do not have a single unambiguous definition, but they indicate the great problems of our age. Wicked problems and grand challenges indicate similar problems, large and complex.

Many of these challenges are incorrigible and extremely challenging for policymakers due to the above-mentioned characteristics. A better understanding of public opinion can be helpful for policy makers in dealing with challenges. Ultimately, policy is made for the public, also in the case of wicked problems and grand challenges. Some policy makers already take public opinion data into account when making policy: research has shown the correlation between public opinion and policy response (Barabas, 2011; Monroe and Alan, 1998; Page and Shapiro, 1983; Pierson, 2002; Stimson et al., 1995). Public opinion formation has changed over the last centuries due to the differences in life styles and standards. Before the invention of the telephone, the telegram and the television, all interactions were face to face. The internet was invented decades ago (Ballard, 2012) and has argued to drastically change the way we shape and share opinions. The internet today, with ample amount of data available, can be a useful source to analyze public opinion (Andrea and Fedra, 2016) and therewith deal with the challenges of our age.

1.1. Problem Statement

There is a lot of meaningful public opinion data available on social media, but policymakers hardly use this data. The data available covers many different subjects and subsequently many opinions on challenges. Policy makers often do not know how to use social media data, do not see the added value or do not know what is being discussed. Social media refer to applications on the internet in which users generate and share content (Kaplan and Haenlein, 2010). Among the most popular social media platforms are Facebook, Instagram and Twitter, all with their own audience and goals (Shane-Simpson et al., 2018).

Twitter has been argued to have the special status of *the* social media platform for news: “it has become a venue where newsworthy tweeters, news consumers and journalists converge to report, read, discuss and share the news” (Orellana-Rodriguez and Keane, 2018, 75). This means that the platform has a focus on the news, whereas other platform focus for example on connecting people (Facebook) or photo sharing (Instagram). Users can communicate on Twitter by sending messages of up to 280 characters called ’tweets’. Tweets are argued to have

played a major role in the 2016 presidential elections in the United States and in the Brexit referendum (Hutchinson, 2016), even though only 20% of US adults were on the platform (Statista, 2018a).

This thesis focuses on public opinion data expressed on Twitter. More than 8000 tweets are posted every second (Stats, 2018). These tweets create a lot of data from which it is difficult to separate noise from signal for all policymakers, and especially for policymakers who do not have a background in big data. Noise and signal need to be separated to use the data. While it is important to thoroughly understand public opinion and the large amount of data available on social media before it can be used in policy making.

1.2. Research Objective

This thesis aims to create insight in public opinion on Twitter. This can be used by policymakers to better include Twitter data in policy making, especially in the case of wicked problems and grand challenges, since these are often complex and involve many stakeholders. Once the policymakers have extracted public opinion data, it can be used in all phases of the policy cycle, from the formation to the implementation of public policy. It can be used as a fast and cheap stream of information, used as an addition to traditional analysis like polling.

Two sub-objectives are developed to reach the main research objective. Firstly, understanding what public opinion is and how it is expressed on Twitter is created. This is based on the line of reasoning from the German philosopher Jürgen Habermas who has written extensively on public opinion. His book, *The Structural Transformation of the Public Sphere* (1962), has been influential in the philosophy on public opinion. Numerous studies describe engagement with online spaces as a incarnation of the Habermasian public sphere (Lehti and Kallio, 2017), and therefore, Habermas' line of reasoning is continued in this research. Secondly, a script of Python code is developed to easily gain insight in public opinion expressed on social media based on chosen key words. This script can be seen as a tool and makes it possible to repeat the analyses of this thesis on another topic. The case study for this thesis is on palm oil, a wicked problem with a vivid debate online. To conclude, this thesis aims to create understanding in public opinion on Twitter and creates an easy-to-use tool for insight in public opinion on Twitter. The understanding created is based on different characteristics, which count not only for palm oil but also for other wicked problems and grand challenges.

1.3. Report Outline

This thesis consists of 9 chapters. The first chapter introduces the research problem and objective. Chapter 2 gives a literature review on public opinion; social media and Twitter; public opinion and policy making; and the subsequent knowledge gap. Additionally, a link with the Engineering and Policy Analysis master is made. This thesis is based on a case study on palm oil, so palm oil is also introduced in this chapter. Chapter 3 introduces the research question and approach per sub-question. It also mentions the scientific and societal relevance and the scope. Chapters 4 - 7 each attempt to answer a sub research question. Chapter 8 gives conclusions and chapter 9 discuses the results and reflects back on the research and the process.

2

Literature Review

A literature review is conducted from which a knowledge gap is identified. The main knowledge gap identified after analyzing literature is the relation between the conceptual Habermasian public sphere on Twitter and results based on actual tweets. Research does not often happen on the interface of philosophy and data science. Therefore, this thesis is an attempt to explore the relation on this interface: philosophy based on the Habermasian public sphere and data science on actual tweets.

Since 'public opinion' and 'social media' play a key role in this research, they are defined first. There are multiple interpretations and explanations among academic researchers as what to include in the terms 'public opinion' and 'social media'. Public opinion, in this thesis, is defined as the collective opinion of the people of a society on an issue (Speier, 1950). Social media, in this thesis, is defined as applications on the internet in which users generate and share content (Kaplan and Haenlein, 2010).

Section 2.1 elaborates on the notion of public opinion, with opinion formation, the emergence of the concept and recent literature. Section 2.2 elaborates on social media and Twitter. The choice for Twitter, the functioning of the network, criticism on Twitter and the link with literature on Habermas and Twitter are discussed. Section 2.3 elaborates on public opinion and policy making, with the policy cycle model of the policy process and ways to measure public opinion. This elaboration is based on traditional ways like surveys or focus groups, and the possibility to include new ways such as social media data. The knowledge gap, following from the previous three sections, is presented thoroughly in section 2.4. Since this thesis is part of the master Engineering and Policy Analysis (EPA), section 2.5 elaborates on the link with EPA, by explaining the grand challenge of social media and the way this thesis aims to inform and help decision makers. This thesis uses palm oil as a case study and therefore, section 2.6 introduces the wicked problem of palm oil.

2.1. Public Opinion

Researchers have written extensively on public opinion and on the correlation between public opinion and public policy (Dewey and Rogers, 2012; Monroe and Alan, 1998; Page and Shapiro, 1983; Stimson et al., 1995). This section describes opinion formation, the emergence of the concept of public opinion and more recent literature on public opinion. Firstly, opinion formation starts by every individual and a public is formed as an aggregate of individuals. Secondly, the concept of public opinion emerged for the first time in 1588 and has been researched extensively. Public opinion can be argued to emerge in so called public spheres, a term tossed by Habermas. Alternatively, it can be formed in fragments smaller than the public sphere or around certain issues that come and go. Thirdly, public opinion can lead to behavior choosing and is therefore more than only spoken word. It is important to note how individual public opinions together form public opinion, based on the idea that an aggregate of individuals forms a public (Rosenberg, 2015).

From individual opinions to public opinion

Opinion formation starts for every individual with exposure to experiences and information. The experiences and information are processed and lead via judgment to attitude (Rosenberg, 2015). To turn all these opinions into public opinion, it is important to understand the public. The public is heterogeneous and complex, and mainly based on two opposing views: one assumes that the public is an aggregate of individuals, whereas the other one takes it as a collective entity, a corpus of its own (Rosenberg, 2015). The mainstream perspective is with the public as an aggregate (Allport, 1935; Anderson, 1991) and so is this thesis. When treating the public as an aggregate, so is public opinion treated as an aggregate of individual opinions. The public as an aggregate means that there is no single public preference, since that indicates a collective entity that can express what it wants in a consistent way (Pesch, nda).

When looking at the public as the aggregate of individuals and expecting all individuals to form an opinion based on attitudes, this can influence behavior choosing of the common people. An example of this is smoking. In the 1960s it was completely accepted to smoke, whereas a shift in public opinion over the last years has made people quit. The shift in public opinion started by individuals changing their opinions and influencing others by exposing them to information. To conclude, individuals form their own attitude based on experiences and there is argued to be no such thing as a public preference. Instead, an aggregate of individual opinions forms public opinion.

The emergence of the concept of public opinion

Public opinion, first named *opinion publique* by de Motaigne in 1588 (Braatz, 2011), is argued to be based on Temple's *general opinion* and Locke's *law of opinion* (Speier, 1950). Public opinion, general opinion and law of opinion all indicate the desires, wants and thinking of the aggregate of individuals forming the public. The German philosopher Jürgen Habermas has written extensively on the emergence of public opinion in relation to the public sphere. He describes the public sphere as a "realm of our social life, in which something approaching public opinion can be formed" (Habermas et al., 1974, 49). A common conceptualization of the Habermasian public sphere suggests that private individuals are enabled to ensemble in a public body (Chen et al., 2018). As illustrated by Habermas, public opinion is formed inside coffeehouses in Europe in the 17th century (Calhoun, 1992). Coffeehouses were places where people gathered for coffee, tea and chocolate, learn about the news and discuss issues of common concern. According to Antoine François Prévost, coffeehouses were a place where one could read newspapers in favor and against governments, as seats of English liberty (Prévost, 1930). This conceptualization is in line with the above mentioned theory of the public as aggregate of individuals. The public sphere is not a market nor a coffeehouse, but a transcending space as an abstract forum for dialogue and public opinion, for vivid debates on multiple levels in society (Kuitenbrouwer, 2018).

Many criticized and extended Habermas' theory. Butsch (2016) argues how there are multiple spheres in society, and each sphere is defined as a homogeneous group. Dahlberg (2007) argues how the public is too diverse to be caught in one public sphere. As time passes and our world is becoming more connected, more spheres can emergence and possibly lead to fragmentation of the public in like-minded groups. Besides falling apart in like-minded groups, the public can also gather itself around certain issues (Dewey and Rogers, 2012). They emerge and disappear as issues come and go. Public opinion can be argued to be formed in public spheres, in fragments of public spheres or around certain issues.

Recent literature on public opinion

More recent literature does not describe public opinion as formed in coffeehouses. More recently is described how "public opinion can reflect the public attitude for some policy, and it can influence behavior choosing of the common people" (Zhu and Hu, 2018, 579). This explanation contains two parts: reflecting public attitude and influencing behavior choosing. Firstly, attitude can be seen as one's evaluation of the object with favor or disfavor (Rosenberg, 2015). As described above, attitude is formed via an individuals' processing of information and their judgment. Secondly, behavior choosing follows from one's attitude. This can be

individual or collectively, when the public organizes itself. Thus, public opinion can lead to behavior choosing and is therefore more than only spoken word.

2.2. Social media and Twitter

This section provides a literature review of social media and especially Twitter, the platform used mostly as a news outlet. The exact definition used in this thesis for social media is by Kietzmann et al. (2011): "Social media employ mobile and web-based technologies to create highly interactive platforms via which individuals and communities share, co-create, discuss, and modify user-generated content" (245). This definitions means that social media is different from traditional media in its digitalness, speed, and mostly in its user-interaction to share, create and adapt content. Participation is key and the existence of the platforms depend on its users (Kietzmann et al., 2011). In 2016 there were over 2 billion social media users (Bailey et al., 2018). Many traditional media outlets have social media editors, university degrees on social media emerge and political and cooperate leaders express their opinions on social media before any other official channel. Social media is, in other words, booming. Since social media has not been around for decades, there is a lot of uncertainty and ambiguity in the governance of social media. An interdisciplinary systems thinking approach to understand the effects of social media is a good first step towards good governance.

Twitter

Twitter launched in October 2006 and is a popular free microblogging platform (Cvetojevic and Hochmair, 2018; Java et al., 2007). Microblogging is the activity of posting frequent brief messages (Merriam-Webster, 2018f). Electricity ended the distinction between day and night, and as such, microblogging is argued to have ended the distinction between one's public and one's private life (Kittle, 2011).

The focus on Twitter in this thesis is chosen for multiple reasons. Firstly, as mentioned in the introduction, Twitter has gained the special status of *the* social media platform for news. This news status makes Twitter a suitable source to analyze public opinion because cases that make the news often generate opinions. Secondly, Twitter is an asymmetrical micro blogging platform, where one can follow someone without that person having to follow one back (Gruzd et al., 2011). This allows following people based on interest instead of personal connections, which in turn can lead to opinion formation based on interest. Thirdly, privacy regulations on Twitter do not limit open and available information for public accounts. Most Twitter users set their account on 'public' which demonstrates the awareness of the consequences of their choice (Bruns and Highfield, 2016, 114). Users want their tweets to be seen by many, not just their personal connections. If they did not want everyone to see their tweets, they would set their account on 'private'.

Messages on Twitter: Tweets

Tweets -the brief messages posted on Twitter- can contain 280 characters, a doubling since November 2017. Images, video and links to external websites can be added to the text. The tweets show up on a users time line, and anyone following this user (called a follower) can see the tweet. The standard setting on Twitter is to have a 'curated' timeline, meaning Twitter decides what you see based on their nontransparent algorithms (Romano, 2018). It is possible to change your 'curated' timeline to a chronological timeline, but this process is not straight forward.

Anyone can sign up and make an account. A few lines from their *about* section give an indication of how they describe themselves: "see what's happening in the world right now; follow your interests; hear what people are talking about; join the conversation; see every side of the story" (Twitter, 2018e). Figure 2.1 shows an example of a tweet by Greenpeace. The following numbers correspond with the numbers in the figure to explain how a tweet works.

1. Tweeted by Greenpeace, user name Greenpeace
2. The text with a maximum of 280 characters; can include URLs and hashtags
3. Hashtags, to kick start a discussion or participate in a discussion on a certain topic. Clicking on the hashtag shows more tweets with the same hashtag and can therefore be used to classify information (Chong, 2016)
4. Possibility to include an image or movie
5. Reply by Andrea Naletto @Greenpeace
6. Possibility to reply, retweet or like this tweet (all very common on Twitter), or to message Greenpeace. Retweeting allows a user to forward a tweet and post it on their own time line, and can be seen as an extension of the information beyond the original tweet's followers (Kwak et al., 2010)

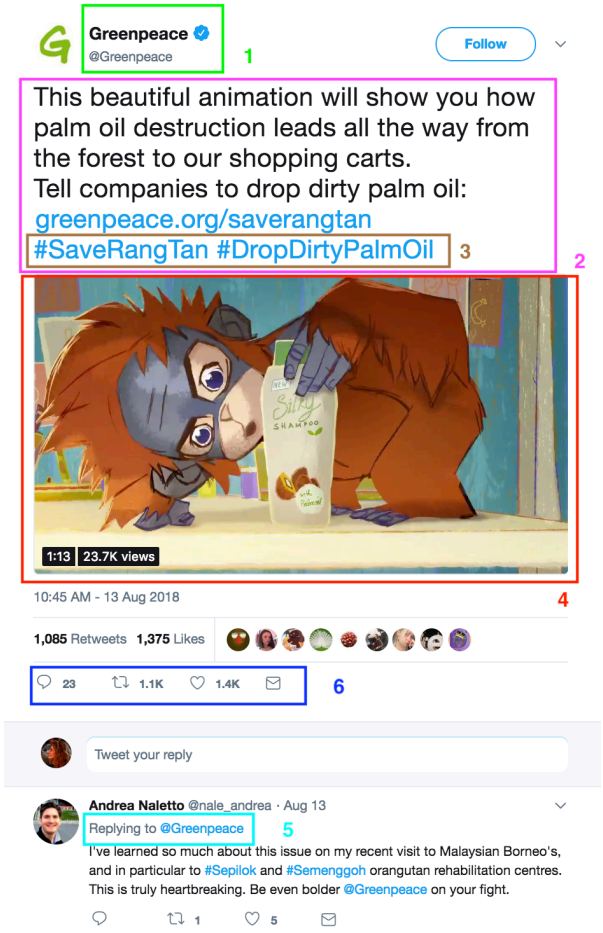


Figure 2.1: An exemplary tweet and reply (retrieved August 20, 2018)

Criticism on Twitter

Twitter does not go uncriticized. Three lines of critique are based on automated messages, alt-right groups and the large volume of tweets. Firstly, automated messages and spam are a problem on Twitter. A report by Nexgate mentions how 1 in every 200 social media posts is spam (Inuwa-Dutse et al., 2018). A more recent study mentions how on Twitter 15% of the active users are bots (Varol et al., 2017). These automated messages and spam can be a problem for measuring public opinion, since these might bring too much noise into the

debate. These messages can influence public opinion strongly, as was the case with the role of Cambridge Analytica in the Trump election (Granville, 2018). Secondly, argued is how Twitter is increasingly used as an alt-right hate-speech hub (Stokel-Walker, 2018). Therefore, argued is how users on Twitter are not a representative public. Thirdly, the large volume and real-time characteristics of tweets can results in noisy postings and information overloads (Lo et al., 2016). The critiques based on automated messages, alt-right groups and the large volume of tweets are only a handful of many more critiques. To conclude, the main take away of this section is that Twitter does not go uncriticized.

Habermas on Twitter

Different researchers wrote on Twitter as an Habermasian public sphere on numerous aspects. Conceptual models are developed for public-sphere like activities on Twitter under certain hashtags (O'Hallarn et al., 2018); investigated is conceptually if online spaces can function as a (re)incarnation of the Habermasian public sphere (Lehti and Kallio, 2017); analyzed is if and how online spaces play a role in facilitating emotions in political discussions (Jackson and Valentine, 2014). These three articles all link social media or Twitter to the Habermasian public sphere but stay on the conceptual level. No data analysis of tweets is conducted in these studies.

Other researchers used data from Twitter to say something about the potential of a public sphere. The effectiveness of Twitter as a platform for political discourse is analyzed (Kasmani et al., 2014); machine learning techniques are used to measure political homophily on Twitter and see if Twitter functions as an echo chamber or a public sphere (Colleoni et al., 2014). These articles do not stay on the conceptual level but move far way from the concept of the Habermasian public sphere and mostly focus on data. No conclusions are made in these studies if Twitter functions as a public sphere, based on the tweets gathered.

In the book *The routledge companion to social media and politics* (2016), Bruns and Highfield devote a chapter to the question *if Habermas is on Twitter.* They show, based on tweets, different publics in the Australian Twittersphere and publics forming around hashtags. Additionally, they show insight in the layered nature of public spheres (Bruns and Highfield, 2016). What is not there yet, however, is a more general view on the characteristics of the Habermasian public sphere. Important characteristics of the public sphere on Twitter, such as the inclusiveness of Twitter as a public sphere; the equality of users; the role and effect of certain users with more influence than others; and the effect of rhetorical devices used in tweets (such as images and selectively interpreted data) on public opinion, are not discussed.

These explorations presented show how different researchers combined Twitter and Habermas. Some have stayed on the conceptual level (Jackson and Valentine, 2014; Lehti and Kallio, 2017; O'Hallarn et al., 2018) and others have moved away from the Habermasian concept to deep-dive in the data (Colleoni et al., 2014; Kasmani et al., 2014). Bruns and Highfield (2016) attempted to combine the conceptual Habermasian public sphere with a quantitative analysis of tweets. Based on these and other articles and books, the knowledge gap is identified. A thorough description of this knowledge gap can be found in section 2.4. In brief, the conceptual and academic characteristics of the Habermasian public sphere are not yet in depth analyzed and combined with the empirical data of tweets.

2.3. Public opinion and policy making

Researchers wrote extensively on public opinion and on the correlation between public opinion and public policy (Dewey and Rogers, 2012; Monroe and Alan, 1998; Page and Shapiro, 1983; Stimson et al., 1995). This section looks at the way public opinion is taken into account in public policy making. It firstly describes the policy cycle with its different stages and secondly describes different traditional and new ways to measure public opinion.

The policy cycle model of the policy process

Lasswell was one of the first researchers to write on the policy process as a series of steps in a cyclical model of decision making in the 1950s (Howlett and Giest, 2015). The policy cycle model has been adapted many times, but is mostly seen in five stages as can be seen in figure 2.2. There are many different ways and models to analyze the policy making process, but I do not think there is the need to elaborate on other models than the policy cycle explained above and visualized in figure 2.2, since the point I want to make concerns the different stages in the policy cycle in which public opinion data from social media can be taken into account.

The first stage, agenda setting, refers to the stage when a problem is sensed and multiple solutions are put forward by policymakers. The second stage, policy formulation, refers to the development and ranking of the different specific policy options. The third stage, decision making, refers to the stage where governmental actors decide upon a specific course of action. The fourth stage, policy implementation, refers to the stage where the decisions are put into effect. The fifth stage, policy evaluation, refers to the evaluation by governmental and non-governmental actors, and can lead to a new cycle (Howlett and Giest, 2015).

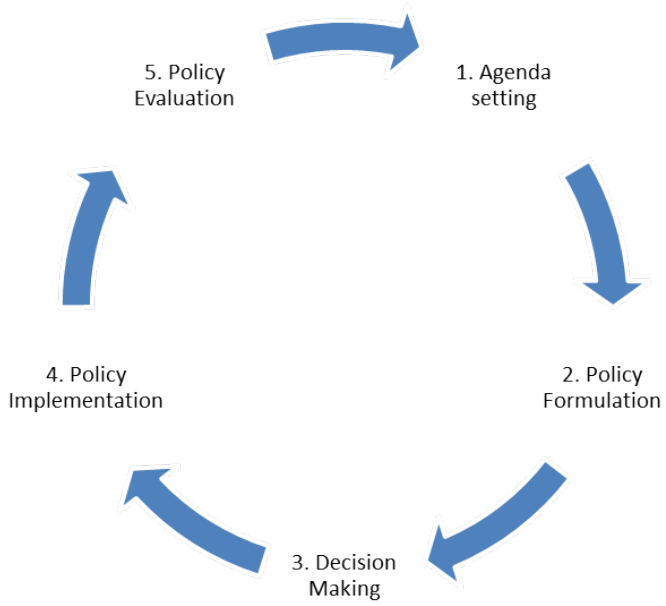


Figure 2.2: Five stages of the policy cycle as described by Howlett and Giest (2015)

It is useful for policymakers to understand citizens' preferences, measure their levels of satisfaction with policies and get feedback on their actions (Andrea and Fedra, 2016). The citizens' preferences, satisfaction and feedback are applicable in different stages of the policy cycle in figure 2.2. Firstly, policymakers can use the preferences on different policy alternatives in the policy formulation stage. Secondly, policymakers can use the levels of satisfaction to monitor opinions during the policy implementation stage. Thirdly, policymakers can use citizens' feedback in the policy evaluation stage (Andrea and Fedra, 2016). Additionally, what topics are important to citizens' can be used in agenda setting. This sections showed how citizens' preferences and opinions can be taken into account in the policy cycle. But before the preferences and opinions can be taken into account, they have to be measured.

Ways to measure public opinion

The preferences and opinion of citizens are often measured in surveys (Andrea and Fedra, 2016) and in focus groups. Different types of surveys exist. These surveys can be online, over the telephone or face-to-face and are given to a small and representative group in society. A focus group is a small group of diverse people in conversation about a certain topic. There are limitations to both. Regarding surveys, a survey is based on a representative sample of the

group of people being studied. But it is difficult to find a representative sample of the public, since the public is an aggregate of individuals instead of a corpus of its own (as explained in section 2.1). Additionally, due to the way questions are framed (often closed questions) people give answers they think are socially desirable. Regarding focus groups, a public is selected either by a dedicated or open invitation (Pesch, nda). This way of selecting might attract only the vocal members of a community of the ones with a strong opinion on a certain topic. Besides, internal group dynamics can have an effect of the opinions formed (Pesch, nda). Due to these limitations, other tools could be used additionally to surveys and focus groups.

Social media data can be used as an additional tool in understanding citizens' preferences and opinions in policy making, but often policy makers do not use social media data. Social media data should not be seen as a panacea to replace surveys and focus groups, but as an addition. Multiple researchers have used Twitter data in policy making: Signorini et al. (2011) tried to build a measure of public concern related to the H1N1 virus based on Twitter data. Ceron (2015) created an indicator to monitor political sentiment in Italy. Social media data is used in these cases because it can monitor real-time information and can be seen as a fast and cheap stream of information (Ceron and Negri, 2015). This data can be used in different stages of the policy cycle in figure 2.2. O'Brien et al. (2017) step up the game of using public opinion data to actively engage participants and encourage contributions to the future food system in Birmingham. Their use of data can fit in the stages of agenda setting as well as policy formulation. These examples show how social media data is used in policy making. These are, however, the exception rather than the rule and many policymakers do not use social media data. This thesis informs policymakers on how to use social media and what public opinion on social media entails. Policymakers can use the insights from social media in multiple stages of the policy cycle, from the stage of agenda setting to the stage of policy formulation and from the stage of policy implementation to the stage of policy evaluation (see figure 2.2).

2.4. Knowledge Gap

The literature review in the previous sections has led to a knowledge gap presented here. The conceptual character of the Habermasian public sphere is not yet combined in-depth with the empirical data of actual tweets. Research does not often happen on the interface of philosophy and data science. This thesis attempts to explore a relationship between the philosophical investigation of the Habermasian public sphere and the empirical investigation of tweets, on the case study of palm oil. Herewith, it combines philosophy with data science. Different researchers combined Twitter and Habermas. Some stayed on the conceptual level (Jackson and Valentine, 2014; Lehti and Kallio, 2017; O'Hallarn et al., 2018) and others moved away from the Habermasian concept to deep-dive in the data (Colleoni et al., 2014; Kasmani et al., 2014). Bruns and Highfield (2016) attempted to combine the conceptual Habermasian public sphere with a quantitative analysis of tweets. Based on these and other articles and books, the knowledge gap is identified: a combination of the conceptual Habermasian public sphere with empirical proof of actual tweets.

The novelty of this thesis is in the elaborate discussion and exploration, based on actual tweets, of characteristics of the Habermasian public sphere. Examples of these characteristics are the inclusiveness of Twitter as a public sphere; the equality of users on Twitter; the role and effect of certain users with more influence than others; and the effect of rhetorical devices used in tweets (such as images and selectively interpreted data) on public opinion. Over 100.000 tweets are gathered to explore if Twitter, based on data, can function as a public sphere. To conclude, this thesis attempts explore an interface between philosophy (via the Habermasian public sphere) and data science (via tweets).

2.5. Link with Engineering and Policy Analysis masters degree

The Engineering and Policy Analysis masters degree (EPA) deals with international grand challenges. As defined by Broekhans (2018), an international grand challenge has the following five characteristics: an urgent decision is needed; there is a lot of data, information

and expertise involved; there are many possible problem definitions and solutions; there are many parties involved; and there is no single problem owner or responsible authority. There are many similarities between wicked problems and grand challenges, based on the complexity and large amount of stakeholders involved. This thesis deals with the grand challenge of social media. As described by professor Lewandowsky from the University of Bristol, "Having a large number of people in a society who are misinformed is absolutely devastating and extremely difficult to cope with" (Gray, 2017). Lies, propaganda and fake news all create a world in which there is no set of facts that people with different viewpoints agree upon. And when the prerequisite of the same starting point (based on a set of facts that people agree upon), is not met, it is hard to address any of the other grand challenge and wicked problems the world faces today. Misinformation is only one of the problems involved in the social media challenge: an imbalance in power is another one. Google, Facebook and Twitter are argued to be the gatekeepers of the internet (Verhagen, 2018). Besides the grand challenge of social media, this thesis uses the case study of palm oil. Palm oil is defined as a wicked problem in section 2.6, and there are many similarities between wicked problems and grand challenges. Both are large, complex and are best approached from an interdisciplinary systems thinking perspective.

Social media as a grand challenge

The five characteristics of a grand challenge, based on Broekhans (2018) introduced in the previous section, are elaborated upon here. Firstly, an urgent decision is needed. Certain companies function as the gatekeepers of our online world (Verhagen, 2018), and since our online and offline worlds are becoming more intertwined, the power of the gatekeepers is increasing. Additionally, more people are becoming connected every day (Berners-Lee, 2018), making the need for a decision more urgent. Since many people base their opinions on what they hear and see, it is important that the information accessible is aiming to inform citizens. Secondly, there is a lot of data, information and expertise involved. More data is generated every day and everything we do online is tracked by Google and Facebook. This data is not owned by you and me, but by Google and Facebook. The expertise and knowledge on this data is therefore also not with the individual user but with the big tech companies. The data they have leads to more expertise and knowledge, due to the advancement of their algorithms based on more data. Thirdly, many possible problem definitions and solutions are proposed. It can be argued how it is great that social media connects the world and opportunities are created that did not exist before. On the other side, it can be argued how these opportunities are only for the ones already benefiting. Many solutions are proposed, from governments stepping in or taking over to boycotting Facebook and Google, but with not much effect yet. Fourthly, many parties are involved: almost all companies using social media for advertising, politicians use social media for campaigns and mostly individuals use social media every day, even though the consequences are up until this day unknown. Fifthly, there is no single problem owner or responsible authority. The internet does not stop at geographical borders and we do not have a world government yet. These characteristics show the urgency and relevance of the grand challenge of social media.

Informing decision makers

Besides a grand challenge, social media can be seen as a grand opportunity when the social media data generates insight for policy makers on other wicked problems and grand challenges. This thesis aims to inform decision makers of public policy on the consequences of opinion formation on social media, and how to use social media in decision making. Understanding social media is a first step in its good governance. Additionally to creating understanding, this thesis aims to create a tool to help policymakers extract Twitter data on any topic they are interested in. The tool can be found on `https://github.com/sevanrossum`. The data extracted can be used in different stages of the policy cycle in figure 2.2. During the EPA program, we learned many different decision making models. I do not think there is the need to elaborate on other models than the policy cycle explained above and visualized in figure 2.2, since the point I want to make concerns the different stages in the policy cycle in which public opinion data from social media can be taken into account.

It is useful for policymakers to understand citizens' preferences, measure their levels of satisfaction with policies and get feedback on their actions (Andrea and Fedra, 2016). The citizens' preferences, satisfaction and feedback are applicable in different stages of the policy cycle in figure 2.2. The next section describes where social media data fits in the policy cycle. It is important to note that the use of social media data should only be seen as an *addition* to traditional tools such as surveys and focus groups. In the agenda setting stage topics that are frequently discussed on social media can possibly be put on the agenda. But firstly, the topics discussed have to be investigated to find out why they are being discussed. An example is the #MeToo discussion on social media. By gaining a lot of attention on social media, policymakers can set the issue of sexual violence on the agenda. The policy formulation stage can use social media data to analyze different preferences on different policy alternatives. The sentiment can be analyzed on different alternatives to get an indication which alternatives rate positively and which negatively. An example of this is the development of a new area. What do citizen express on social media about a shopping area, an entertainment area or a park? The policy implementation stage can be strengthened by measuring the level of satisfaction during implementation. An example of this is the test stage for a prohibition of taxi's around Leidseplein in Amsterdam, to prevent nuisance. Social media data can be used to see if people are expressing their opinion on their level of satisfaction on the prohibition. The evaluation stage can use feedback. Not all citizens are willing to give direct feedback, but when expressing an opinion on social media, this can be seen as providing feedback in a more accessible way. An example of this is the expression of feedback on new guidelines on the use of renewable energy in Europe. Social media data can be used to find unsolicited feedback.

This research is part of the Engineering and Policy Analysis masters degree. This program deals with the grand challenges of our age and this thesis deals with the grand challenge of social media. It is a grand challenge for many reasons, including lies, propaganda, fake news and an imbalance of power. Social media is chosen as a challenge because besides a grand challenge, it can also be seen as a grand opportunity. When social media data is used correctly, it can possibly generate insight in public opinion. This, in turn, can be used by policymakers in the decision making process on other grand challenges, as is ex amplified here on the case study of the grand challenge of palm oil.

2.6. Palm oil: a wicked problem case study

The case study for the empirical investigation is based on the debate on palm oil. This case study is chosen because palm oil can be described as a wicked problem and the online and offline debate is vivid. Palm oil is described as a wicked problem for two reasons. Firstly, because of its characteristics, including incomplete, contradictory, and changing requirements that are often difficult to recognize (Rittel and Webber, 1973). Palm oil is in 50% of the goods we buy in the supermarket (WWF, 2018) and is the leading source of vegetable oil world wide (Hansen et al., 2015). The world does not know how to deal with the growing demand. Secondly, there is no indisputable public good nor single definition of equity in society, meaning there is no optimal solution to these problems possible (Rittel and Webber, 1973). Palm oil is argued to be good for economic development band not so good for the environment due to forest degradation and deforestation (Purnomo et al., 2018), negatively affecting the natural habitat of people and animals living in the forest. Both characteristics are explained in-depth.

Incomplete, contradictory and changing requirements

The challenge of palm oil has many faces. Palm oil is in 50 percent of the products we buy in the supermarket, and increasingly used as a bio fuel (WWF, 2018). The high usage percentage makes us highly dependent on the oil grown mostly in Indonesia and Malaysia. Palm oil has been used by mankind for thousands of years. An immense expansion in its cultivation has happened over the last fifty years, to the extent that palm oil is now the leading source of vegetable oil worldwide (Henson, 2012). Palm oil makes up 29% of the global fats and oils, and the demand is expected to exceed 250 million tonnes in 2050 (Masani et al., 2018). This estimate, however, is difficult, since a large share of the production is illegal and therefore

difficult to measure (Purnomo et al., 2018).

The world does not know how to deal with the increasing demand of palm oil and therefore frequently changes the requirements. The World Bank has funded the palm oil industry for 45 years until they suspended their funding due to the conflicts the palm oil plantations caused, the rain forest destruction and more abuses (Rescue, 2018). The Round table for Sustainable Palmoil (RSPO) emerged as a multi-stakeholder initiative to deal with the problem of unsustainable palm oil (Dentoni and Bitzer, 2015). The RSPO unites stakeholders from the seven sectors of the palm oil industry to develop and implement global standards for sustainable palm oil (RSPO, 2018), but they are argued to be ineffective in achieving conservation objectives (Ruysschaert and Salles, 2014) and therewith a sustainable palm oil sector. There is a lot of ambiguity on what is sustainability within the palm oil sector and what the effects are of palm oil plantations, due to the large illegal business surrounding it. To conclude, the world is highly depend on palm oil and only becoming more dependent. The RSPO emerged as a multi-stakeholder initiative to deal with the unsustainable palm oil cultivation, but they are argued be ineffective in achieving a sustainable palm oil sector.

No indisputable public good nor single definition of equity

The production and consumption of palm oil have many different effects on society. Palm oil is argued to be good for economic development and not so good for the environment due to forest degradation and deforestation (Purnomo et al., 2018). Forest degradation and deforestation affect the living habitat of people and animals living in the forest, such as indigenous tribes, orangutans, elephants and tigers. The United Nations developed the Sustainable Development Goals (SDGs) in September 2015 to end poverty, protect the planet and ensure prosperity for all (Priyadarshini and Abhilash, 2018; UN, 2018). These goals can be seen as wicked problems or grand challenges. The current production and consumption of palm oil touch on many of the SDGs, where it represents a solution for some goals like economic prosperity (goal 8) but a problem for environment-related goals (goal 15). To conclude, the contradiction in the SDGs shows how there is no indisputable public good: the cultivation affects some positively (e.g. multinationals, land owners in Indonesia) and some negatively (e.g. indigenous tribes and orangutans in the forest).

Debate on palm oil

The debate on palm oil has multiple sides, at least a scientific and popular side. The scientific side has grown: there has been an exponential growth in research on palm oil and sustainability, from 11 ISI article publications (ISI stands for International Scientific Indexing) in 2004 to 713 in 2013. The growth is mainly due to increased research on the technical topics associated with palm oil residues, and very few related to biodiversity, economic, social and environmental impacts (Hansen et al., 2015). The scientific side is in stark contrast with the popular side, driven mainly by NGOs and with increased media coverage on the topic of deforestation (Padfield, 2010). Part of the popular debate happens on social media and many tweets are posted daily regarding palm oil. To conclude, the debate of palm oil has many faces. This thesis aims to show insight in part of this debate, namely the debate on Twitter.

The difficulty of policy making in the case of palm oil

Due to the characteristics of palm oil as a wicked problem, it is difficult for policy makers to make good policy. Difficulties are the increasing demand for palm oil; ineffectiveness of the RSPO; the conflicting interests between the producing and consuming countries; and between the NGOs and multinationals. The debate has multiple sides that do not communicate properly with each other, leading to a compartmentalization of the debate. Policy makers on different levels could use public opinion data better to include this in their policy making.

3

Research Question and Approach

3.1. Research Questions

The main research question is:

How can public opinion on Twitter be characterized, on the case study of palm oil?

In this question, characterizing means to describe the special, typical or telling qualities of public opinion on Twitter (Merriam-Webster, 2018b). Palm oil is chosen as a case study to gather and analyze empirical data. The answer to this question aims to give policymakers insight in how public opinion can be characterized and used in policy making. To answer the main research question, there are four sub-questions. The sub-questions are mentioned in the list below and elaborated on in section 3.2. Section 3.2 also introduces a framework by which the rest of this thesis is structured.

1. *What are possible characteristics of public opinion on Twitter?*

   Sketching possible characteristics by an elaboration on the information from the literature review and therewith bringing together different lines of reasoning. This chapter defines theoretical concepts and ope rationalizes them into constructs, based on the theory of the public sphere by the German philosopher Jürgen Habermas.

2. *What are relevant tweets regarding palm oil?*

   Gathering tweets on the topic of palm oil for the case study.

3. *What characterizes public opinion on palm oil on Twitter?*

   Combining the first two sub-questions and therewith applying the theory to the tweets, on the case study of palm oil.

4. *What can be generalized about public opinion on Twitter from the case study of palm oil?*

   Generalizing the results from the previous sub-question, away from palm oil. These generalized results can be used to inform policy makers on the nature of social media, before they can use this data in decision making for wicked problems.

3.2. Research Methodology

The research methodology in this thesis is based on a framework developed for a combined research on philosophical theory and empirical data. The framework is shown in figure 3.1 and shows the different steps and chapters this thesis follows. The framework describes seven steps, as can be seen on the left (from A-G). On the right are two columns indicating the sub-question and the chapter dealing with the specific part of the framework. The far right

shows an hourglass as the shape of this thesis. It starts out with general theory on public opinion on Twitter (sub-question 1), gathers and analyzes specific data on the case study of palm oil (sub-question 2 and 3) and goes back to general conclusion on public opinion on Twitter (sub-question 4).

Before explaining the methodology per sub-question, the role of the Habermasian public sphere has to be clarified. The theory by Habermas on the *Structural Transformation of the Public Sphere* is used in two ways in this thesis. Firstly, it is used as a coat rack for analyses to come up with general characteristics. Secondly, the public sphere *in itself* is more than a coat rack. It is a lens through which Twitter can be analyzed and new insights and characteristics are created. Platforms on the internet have the potential to function as modern public spheres. They can be open, democratic and egalitarian. Analyzing Twitter via the Habermasian public sphere thus functions as a coat rack and as a lens, generation additional insights.

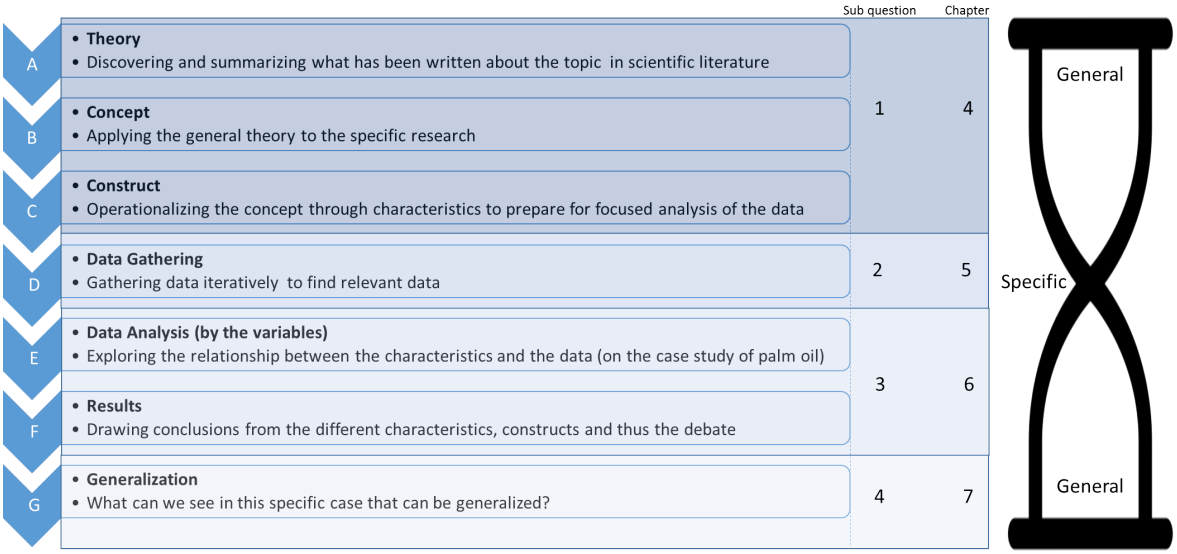


Figure 3.1: Guiding framework in this thesis

Sub-question 1: What are possible characteristics of public opinion on Twitter?

The first sub-question distills broad theory into three specific and applied concepts of public opinion on Twitter. These three are the Habermasian public sphere, fragmentation and issue publics. Firstly, the Habermasian public sphere is a place where something approaching public opinion is formed. An example could be the coffeehouse, as explained in section 2.1. Secondly, fragmentation of the public happens, as argued by opponents of Habermas, when the public is too diverse to be caught in one public sphere. According to them, the public sphere fragmentizes in smaller spheres called sphericules. Thirdly, issue publics are publics that gather around certain issues. The publics emerge and disappear as issues come and go. An example is the issue of the Olympic Games. Around the Olympic Games, many people are talking about the games but when they are over, so are most of the conversations around them. Public opinion can be argued to be formed in public spheres, in fragments of public spheres or around certain issues.

Section 4.1 describes the three above mentioned concepts in-depth, where general theory is conceptualized. Step C, constructs, are more specific than step B, concepts. Constructs operationalize the concepts by translating the concepts into more concrete (and almost measurable) characteristics. These constructs function like a bridge between the raw and broad theory and the specific data (section 4.2). Cooke and Campbell describe constructs as the operationalized variables used for observation (Calder et al., 1982).

Some concepts, like height, are easily turned into a construct (centimeters) and measured. A more complicated example is knowledge. The theory of knowledge contains the concept of

intelligence, which is about the forming of ideas on showing more or less intelligent behaviour. Intelligence as a construct contains characteristics to define and measure intelligence more specifically, like the ability to comprehend situations, to respond to words and numbers and to react to social situations (Parankimalil, 2014). The methodology for this sub-question is a combination of literature reviews and interviews with policymakers who make use of social media data to measure public opinion to define possible characteristics of public opinion on Twitter. The aim of this sub-question is to clarify the concept-construct structure and use these as the bottom building blocks for the rest of this thesis.

Sub-question 2: What are relevant tweets regarding palm oil?

The second sub-questions gathers tweets relevant in the debate on palm oil. The process of tweet gathering is iterative and a visualization can be seen in figure 3.2. There are five steps: installing the software, defining hashtags and keywords, spot-check preliminary tweets, install a server and getting a first impression of the tweets. Chapter 5 describes the process of data gathering extensively.

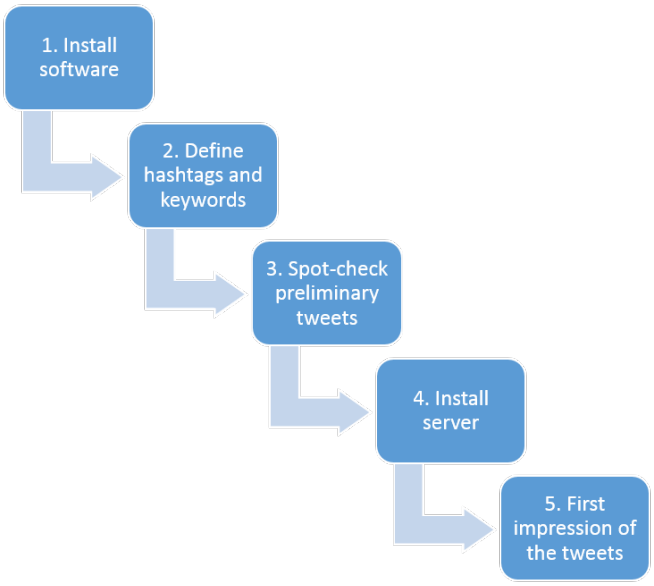


Figure 3.2: Iterative process of collecting tweets

The first step, installing software, contains multiple steps. Python, an open source, easy to use and widely accepted programming language (McClarren, 2018) is run in the Jupyter Notebook environment. Notebooks are documents produced in the Jupyter Notebook environment and contain computer code and text (Ingargiola, 2015), making them easier readable for humans. To analyze tweets, Tweepy is installed. Tweepy is a Python library for accessing the Twitter API. API stands for Application Programming Interface, and the Twitter API used in this thesis searches against a sampling of recent Tweets published in the past 7 days (Twitter, 2018c). When the first step is done, all the necessary preparations are taken to start searching for tweets.

The second step, defining hashtags and keywords, is based firstly on literature on palm oil and secondly on interviews with experts. Section 2.6 gave an introduction to the literature and section 5.2 gives an elaboration on who is being interviewed to define the hashtags and keywords.

The third step, to spot-check preliminary results, is based on looking at the tweets and checking for notabilities. An example of a notability is when a tweet has been retweeted a lot, but without a direct link to palm oil. This can happen when it contains the words palm and oil separately but not together and is not concerning the topic to be addressed in this thesis. If it turns out there is structural problem with one of the defined hashtags and keywords, these will be changed in the second step. See the upward arrow in section 3.2. Additionally, this

step is checking if there are enough tweets collected to conduct this analysis. The minimum is set at 50 tweets an hour.

The fourth step, installing a server, aims to collect tweets 24 hours a day. This will be a virtual server to collect tweets for 50 days in total, from July 23 to September 11. Tweets collected are saved in so-called JSON files (JavaScript Object Notation) per 500 tweets so they are ready for analysis.

The fifth step, first impression of the tweets, is the last step before the in-depth analysis and is checking the hashtag frequencies, word frequencies, tweet frequencies and users. The most frequent words and hashtags should make sense in the palm oil debate and the tweet frequencies should be over 50 tweets an hour. The tweet frequencies are plotted in time series and the hashtag- and word frequencies in word clouds. The aim of this sub-question is, through these five steps, to gather and prepare tweets for more in-depth analysis in the next sub-questions.

Sub-question 3: What characterizes public opinion on palm oil on Twitter?
The third sub-questions analyzes the data to explore the relationship between the constructs defined in sub-question 1 and the data gathered in sub-question 2. See step E and F in the framework in figure 3.1. Chapter 6 is structured so analysis are conducted per construct. This sub-question analyzes the data to draw conclusions on the different constructs. This is achieved by analyzing the data using Python in Jupyter Notebook and additional libraries for data analysis. See appendix A for an overview of all the libraries and packages used with their version number. The most frequently used libraries are Pandas for data structuring and analysis, Matplotlib for making graphs, WordCloud for making word clouds and NLTK (Natural Language Toolkit) for processing language. The aim of these analyses is to show results and draw conclusions on the different constructs to define characteristics on the specific case of palm oil.

Sub-question 4: What can be generalized about public opinion on Twitter from the case study of palm oil?
The fourth sub-question zooms out from the specific results on palm oil and aims to conclude what these results mean for general public opinion on Twitter. See step G in the framework in figure 3.1. The generalization is achieved in two ways in chapter 7. Firstly, the relation between the three concepts is investigated and broadened to general debate instead of only palm oil. These three concepts are the Habermasian public sphere, fragmentation and issue publics. Secondly, the structure of concepts is released and overarching themes are investigated. These can be based on specific characteristics from within the concepts but can also be overarching or based on a combination of multiple characteristics inside different concepts. The aim of this sub-question is to come up with general characteristics and interpretation of these characteristics, to define what this means for public opinion on Twitter.

3.3. Scientific and Societal Relevance

This thesis has two objectives: to add scientific value and to add societal value. Regarding scientific value, the conceptual character of the Habermasian public sphere is not yet combined with the empirical data of actual tweets in existing studies. This thesis attempts to explore a relationship between the philosophical investigation of the Habermasian public sphere and the empirical investigation of tweets, on the case study of palm oil. Different researchers combined Twitter and Habermas. Some stayed on the conceptual level (Jackson and Valentine, 2014; Lehti and Kallio, 2017; O'Hallarn et al., 2018) and others moved away from the Habermasian concept to deep-dive in the data (Colleoni et al., 2014; Kasmani et al., 2014). Bruns and Highfield (2016) attempted to combine the conceptual Habermasian public sphere with a quantitative analysis of tweets. The novelty of this thesis is in the elaborate discussion and exploration, based on actual tweets, of characteristics of the Habermasian public sphere. Examples of these characteristics are the inclusiveness of Twitter as a public sphere, the social equality of users on Twitter, the role and effect of certain users with more influence than others and the rhetorical devices used in tweets. Over 100.000 tweets are gathered to explore, based on data, if Twitter can function as a public sphere.

Regarding societal value, by showing insight in public opinion on Twitter on the case study of palm oil, the aim of this thesis is to inform policymakers on how to use public opinion extracted from Twitter in decision making. The public opinion data on social media can be applied to any wicked problem or grand challenge and in any part of the policy cycle, from rating citizen's preferences in the formulation phase to help monitor opinions during the implementation phase (Andrea and Fedra, 2016). The programming language used to analyze the tweets is Python. Python is open, easy to use and widely accepted (McClarren, 2018). The Python code used to analyze the tweets has to be clear, well-documented and easy-to-use, so it can be easily applied to any wicked problem or grand challenge. Moreover, as a side objective, this thesis will generate insight in the debate on palm oil. The characteristics of the Twitter debate on palm oil will become clear, such as the main actors, the main topics and the general opinion.

3.4. Scope

The duration of this thesis is six months so choices are made. This section describes firstly the demarcation on public opinion, secondly the choice of Twitter and thirdly the case study of palm oil. Theoretically defining public opinion can in itself be a thesis and has been researched widely. The main line of reasoning is based on Habermas, a very influential thinker on the public sphere. Others, like John Dewey and John Locke, have also been influential in the debate and are therefore briefly mentioned. Habermas is playing a central role because his theory is very suitable for analysis in the 21st century: multiple research has written on the (re)incarnation of the Habermasian Public Sphere (Bruns and Highfield, 2016; Lehti and Kallio, 2017). The public sphere as a place where something approaching public opinion is formed, can be very suitable to Twitter.

Demarcation of Twitter

Twitter is chosen for multiple reasons. The three reasons mentioned in section 2.2 are based on Twitter as a news platform, as an asymmetrical microblogging platform and the high amount of public instead of private accounts. Additionally, the Twitter Application Programming Interface (API) is used to gather tweets. The Twitter API is easy to use and accessible. By the free Twitter API used, it is not possible to use data from more than 7 days before the starting moment of gathering. This naturally limits the dates of data gathering, which will be 50 days: from July 23 - September 11. 50 days are chosen to gather at least 100.000 tweets. There are no geographic boundaries: instead, all tweets gathered are based on specific key words and hashtags. See chapter 5 for the process of data gathering. The Twitter debate does not cover the whole debate, since real life, traditional media and other social media cannot be excluded. These all fall outside the scope of this thesis. However, public opinion expressed on Twitter is argued to represent an 'activated public opinion'. This activated public opinion, in turn, can anticipate trends in wider public opinion (Andrea and Fedra, 2016). Therefore, the data gathered on Twitter over 50 days regarding palm oil can give insight in public opinion on palm oil and general characteristics of public opinion on Twitter.

A case study on palm oil

The case study for the empirical investigation is based on the debate on palm oil. This case study is chosen because palm oil can be described as a wicked problem (as explained in section 2.6) or a grand challenge, and the online and offline debate is vivid. It transcends national borders and involves many stakeholders (like multinationals, citizens, farmers and (non)-governmental organizations). Other case studies could have been chosen but the focus of this thesis is on the system of social media and not on the case study of palm oil. Therefore, even if a different case study was chosen, the characteristics of Twitter should still show.

4

Concepts and Constructs

To answer the first sub-question, *What are possible characteristics of public opinion on Twitter?*, literature is reviewed and concepts are developed. Subsequently, these concepts are operationalized into constructs. Figure 4.1 shows the selection of the framework from chapter 3 that accompanies this chapter highlighted in green. The consecutive phases A, B and C show the order and logic of this chapter. Based on literature, three concepts are developed to apply general theory (from section 2.1) to the specific research of public opinion on Twitter. These three concepts will be elaborated on in-depth and are the Habermasian public sphere, fragmentation and issue publics. They are operationalized in constructs.

It is important to emphasize how these descriptions and concepts are abstractions of much more encompassing and complicated theories. The description have an ideal-typical character, which is necessary in order to operationalize the empirical research. These concepts are all related to the public sphere. They can be argued to be based on criticism on the public sphere or on additions and interpretations, as introduced in section 2.1.

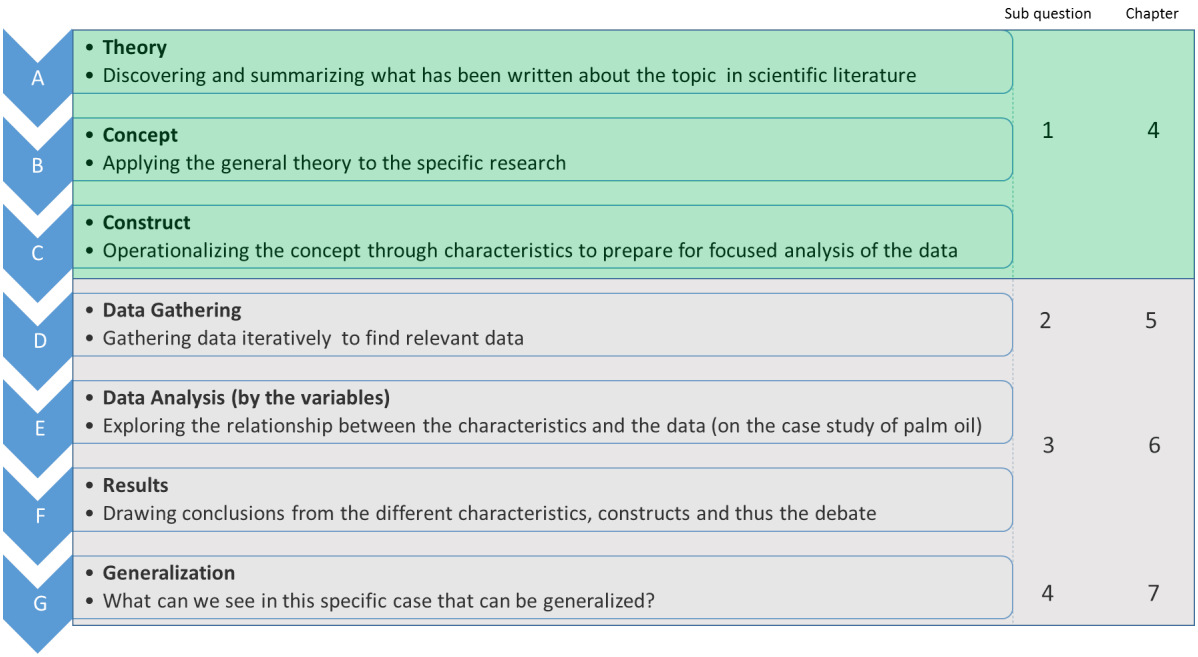
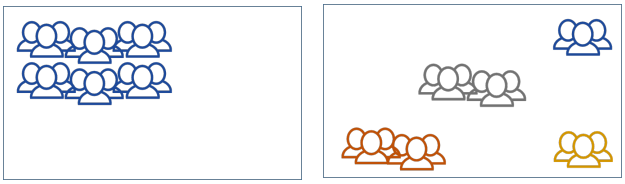


Figure 4.1: Framework introduced in chapter 3 with the phases A, B and C highlighted

The first concept, Twitter in an incarnation of the Habermasian public sphere, is explained in section 4.1.1. In short, the Habermasian public sphere is described as the place where something approaching public opinion is formed (Calhoun, 1992). This sphere is applied to Twitter, to see if Twitter is an incarnation of the public sphere. Figure 4.2a is a simple visualization of a public sphere, as those forming public opinion. It is not necessarily true that everyone has the exact same opinion, but by aggregating all individual opinions, something approaching public opinion is formed and therefore, the persons in the figure are in the same color and in the same area.



(a) The Habermasian Public sphere

(b) Public sphericules (fragmentation)

Figure 4.2: The public sphere and public sphericules

The second concept, Twitter causes a fragmentation of the public, is explained in section 4.1.2. This concept is based on the addition by Nancy Fraser to Habermas' theory on the existence of multiple spheres in society. A recent case study on social media debates on Traditional Chinese Medicine shows that in the online public sphere, "debates and conversations do not facilitate viewpoint exchange, but they reinforce antagonism between groups. [...] In short, different networks of expertise are getting farther and farther away from each other in terms of opinion exchange, resulting in the fragmentation of public sphere at large" (Chen et al., 2018, 8,27)". This means that people only talk to those who reinforce what they already think, leading to a fragmentation of the public in like-minded groups. The term used in literature for fragmentation of the public sphere are public sphericules: "social fragments that do not have critical mass [but] share many of the characteristics of the classically conceived public sphere" (Cunningham, 2001, 135). The classically conceived public sphere is split up in different social fragments, which keep the same characteristics (which will be explained later). A simple visualization can be seen in figure 4.2b, with different groups of people (shown in different colors) in different directions (different places in the rectangle).

The third concept, Twitter creates space for issue publics, is explained in section 4.1.3. This zooms further in on the public sphericules as shown in 4.2b. Issue publics are publics that form around certain issues. The publics emerge and disappear as issues come and go. An example is the issue of the Olympic Games. Around the Olympic Games, many people are talking about the games but when they are over, so are most of the conversations around them. Public opinion can be argued to be formed in public spheres, in fragments of public spheres or around certain issues. The public is smaller than the public in a public sphericule, which are both smaller than the public in the public sphere (since that is the whole public). Another example to clarify the three concepts is climate change as a public sphere, global warming as a public sphericule and the 2018 heat wave in the Netherlands an issue around which a public formed. This issue comes and goes as time passes by and the hot summer is over, which is the case in many issues currently at stake. Issue publics are shorter-lived and more dynamic than the notions of public sphere and public sphericule (Dahlgren, 2009, 74).

(a) Issue Publics, t=0 (b) Issue Publics, t=1

Figure 4.3: Issue publics at t=0 and t=1

Additional to the third concept is the notion of the attention economy. We are living in a the so called 'attention economy' where we have access to more information than ever and more distractions all fighting for our attention (Atchley and Lane, 2014). For a fruitful public debate, however, there are only a limited number of issues we can discuss. Coordinated action on issues ask for aligning the issues we are discussing. This is an interpretation of the attention economy in social organizational terms. When looking at the psychology of the limitations in our brains, we have limited attention spans. According to a (non-peer reviewed) research by Microsoft Canada, the average human attention span has decreased from 12 seconds in 2000 to 8 seconds in 2013. This is argued to be less than the average attention span of a goldfish (Galloway, 2017). Figure 4.3 shows a simple visualization of issue publics. Figure 4.3a shows issue publics at t=0. The public is split up in smaller groups and in more different directions compared to the public sphericules in figure 4.2b. Figure 4.3b shows the public at t=1, where they are shuffled around towards new issues since these just come and go.

To summarize, the different concepts are elaborated on in section 4.1. Thereafter, they are translated into constructs in section 4.2 using different characteristics. Table 4.1 is developed to keep an overview of the different concepts and constructs. The table shows the definition, characteristics used in operationalizing (which will be explained in section 4.2) and the visualization (based on figures 4.2 and 4.3).

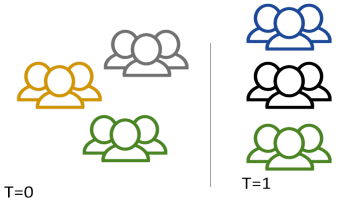
	Concept & construct 1	**Concept & construct 2**	**Concept & construct 3**
Definition	Twitter is an incarnation of the Habermasian public sphere	Twitter causes a fragmentation of the public	Twitter creates space for issue publics
Explanation	A public sphere is a transcending space, an abstract forum for dialogue where something approaching public opinion is formed	When the public is too diverse to be caught in one public sphere, the public sphere fragmentizes into smaller particles called sphericules. Sphericules have the same characteristics as the public sphere.	Issue publics are publics that gather around certain issues. The publics emerge and disappear as issues come and go.
Visualization			
Characteristics	- Social inequalities are bracketed - Public sphere is a new center of institutional authority - Access is granted to all citizens - Those forming public opinion are presupposed reasonable	- The public on Twitter is fragmented - The fragmented public on Twitter is caused by Twitter	- Twitter functionalities (retweets and hashtags) create space for issue publics - Certain issues push others away in a time where our attention is scarce

Table 4.1: Overview of the different concepts with their characteristics and a visualization

4.1. Concepts

The concepts used in this thesis function as the bottom building blocks on which the data analysis can be build. The figures 4.2 and 4.3 show the relation between these concepts. It is important to emphasize again how the concepts and images are all abstractions and simplifications of reality, to operatonalize my empirical research. Habermas is a red line throughout this thesis. The Habermasian public sphere shows how the public is an aggregate of individuals forming a public opinion. The first concept is *Twitter is an incarnation of the Habermasian public sphere* and can be found in section 4.1.1. Public sphericules are based on Fraser's addition to Habermas' theory and show the possible fragmentation of the public. The concept is *Twitter causes a fragmentation of the public* and can be found in section 4.1.2. Issue publics are, in the time where our attention is scarce, dynamic and volatile. The concept is *Twitter creates space for issue publics* and can be found in section 4.1.3. These three concepts are not an opaque coverage of public opinion but are used as a lens to analyze public opinion.

4.1.1. Concept 1: Twitter is an incarnation of the Habermasian public sphere

This section firstly explains what is meant by the Habermasian public sphere, secondly discusses some of the major critiques to the public sphere and thirdly applies the public sphere to Twitter. Numerous studies describe engagement with online spaces as a incarnation of the Habermasian public sphere (Lehti and Kallio, 2017). As defined by Habermas himself, the public sphere is a "realm of our social life, in which something approaching public opinion can be formed" (Habermas et al., 1974, 49). It is a sphere of people, who from their private interest and space, enter and therewith shape a public (Habermas et al., 1974). The public sphere is not a market nor a coffeehouse, but a transcending space as an abstract forum for dialogue and public opinion, for vivid debates on multiple levels in society (Kuitenbrouwer, 2018).

Those forming public opinion are presupposed reasonable, making the rational-critical debate a key feature. Here, may the best argument win and not the one made by the man of highest status (Calhoun, 1992). This argument should be free of ideology and power, with the common interest of good opinion formation as the only interest (Kuitenbrouwer, 2018). The 'rational' part of the rational-critical debate "was supposed to transform voluntas into a ratio that in the public competition of private arguments came into being as the consensus about what was practically necessary in the interest of all" (Habermas et al., 1991, 83). This means that voluntas (loosely translated as desires) are changed into ratio (loosely translated as reason). Turning desires into reason creates reasonable arguments. A simplification to explain this is based on wanting a new school. You want a new school in your neighbourhood (a desire), and therefore you come up with an explanation and justification on why you want this school (reason). This rational part is key to the process of public opinion formation. The 'critical' part of the rational-critical debate is based on the judgment by private individuals on political issues.

To conclude, in the rational-critical debate actors are presupposed reasonable. They are expected to bring in private arguments, free of ideology and power, with the common interest of good opinion formation as the only interest (Kuitenbrouwer, 2018). They should not fear critical judgment on political issues. The emergence of public opinions is argued to have happened inside coffeehouses in Europe in the 17th century (Calhoun, 1992). Coffeehouses were places where people gathered for coffee, tea and chocolate, learn about the news and discuss issues of common concern. The possibility to have these open discussions made these coffeehouses a center for (critical) political and intellectual discussion.

Characteristics of the public sphere

Public spheres can be argued to have three clear characteristics: social inequalities are bracketed; public spheres are more than a new place to talk; and access is granted to all citizen. This is based on an ideal-typical simplification of the theory on the Habermasian public sphere, which is a necessary step in the preparation of the literature for linking it later with the actual data. Firstly, social inequalities are bracketed. As Habermas describes, "a kind of social intercourse that, far from presupposing the equality of status, disregarded status

altogether" (Calhoun, 1992, 12). To explain what is meant with this, firstly status is defined and secondly the disregarding of status is interpreted. Status is defined as *a position or rank in relation to others* (Merriam-Webster, 2018h). When this definition of status is taken into account and social inequalities are indeed bracketed, everyone is not only treated but actually seen as equals. This thought is in line with egalitarianism, preaching that all humans are equal in fundamental worth (and thus social status). Bracketing social inequalities means that the public sphere is of and for everyone, not just of and for the privileged class in society. This leads to everyone being able to speak to each other, regardless their social status.

Secondly, public spheres are more than a new place to talk. They are new centres of institutional authority. An institutional authority is an established organization with the power to influence. Institutional, as based on the word institution, is defined as "an established organization or corporation especially of a public character" (Merriam-Webster, 2018e). The authority is the "power to influence or command thought, opinion, or behavior" (Merriam-Webster, 2018a). Public spheres can be seen as sovereign, autonomous institutional authority. This means that where normally the state authority has the monopoly of interpreting information in a way favorable to them, now these public spheres can go against the authority of the state by interpreting information in a different way. The power of going against authority can either be expressed in a counterweight to traditional authority, by speaking out against certain structures and calling for structural change as has been the case in many revolutions in history. Alternatively, it can be expressed as a direct call for action towards others, for example by calling to not use airplanes anymore to fight climate change. These calls for action can influence behavior choosing directly (Zhu and Hu, 2018), as introduced in section 2.1. A new centre of institutional authority is more than a new place to talk: it is a place where people either speak out against certain structures or where people make direct calls for action to influence behavior choosing.

Thirdly, "access is granted to all citizens" (Habermas et al., 1974, 49). This builds further upon the first characteristic and works towards an egalitarian public sphere. Anyone with reason and the willingness to join was welcome to join public debate. Reason is a statement offered in explanation or justification (Merriam-Webster, 2018g). These statements should, according to Habermas, not be based on an ideology or power (Kuitenbrouwer, 2018). All citizens are seen as private persons who can bring in arguments that will eventually define public opinion. The arguments brought into the public sphere should be reasonable, including a statement in explanation or justification without being based on an ideology or power. These arguments together consequently form something approaching public opinion.

These three characteristics -bracketing social inequalities, new centres of institutional authority and access is granted to all- are an ideal-typical simplification of the theory on the Habermasian public sphere. Habermas envisioned the public sphere as an open, democratic and egalitarian space. These three characteristics are an attempt to approach what Habermas meant and to operationalize the theory enough to explore a relation with the tweets gathered.

Criticism on the Habermasian public sphere

The Habermasian public sphere does not go uncriticized. Some critics argue how there are multiple spheres in society, and each sphere is defined as a homogeneous group (Butsch, 2016). The public is too diverse to be caught in one public sphere. As time passes and our world is becoming more connected, more spheres can emergence and possibly lead to fragmentation of the public in like-minded groups (Dahlberg, 2007). Besides falling apart in like-minded groups, the public can also gather itself around certain issues. They emerge and disappear as issues come and go. Public opinion can be argued to be formed in public spheres, in fragments of public spheres or around certain issues.

On of the outspoken critics of *The Structural Transformation of the Public Sphere* is Nancy Fraser (Calhoun, 1992). She explicitly criticizes all three characteristics (Fraser, 1992) introduced above: social inequalities are bracketed; public spheres are more than a new place to talk; and access is granted to all citizen. Firstly, on bracketing social inequalities, Fraser argues how "such bracketing usually works to the advantage of dominant groups in society and to the disadvantage of subordinates" (Calhoun, 1992, 120). This means that most people

do not have any advantage bracketing social inequalities because this labelling only works in the advantage of the bourgeois: the white, middle class, male. Others have added how communication is carried on by an elite, and by contrast, “ordinary people-the public- are cast in the role of audience member who for the most part are merely able to watch the events unfolding on this ”virtual stage of mediated communication” (Bruns and Highfield, 2016, 56). This means that the ordinary people do not have a say and therefore, the public sphere is not a very egalitarian place.

Secondly, the line between the public and the private issue is not so clear. Some issues are not new centres of institutional authority because they will only concern (private) issues. Private arguments, based on ideology of power, are not in the interest of all nor in the interest of good opinion formation. The arguments were just in the interest of a few, not making it a public issue but keeping it public.

Thirdly, not everyone is welcome in the coffeehouse. Once in, anyone can open their mouth, but not anyone can come in. Women, for example, were excluded. The issue of gender has been criticized by many others, including Benhabib, Eley and Ryan (Negt et al., 1988). Habermas has noted the exclusion of women but has not in depth dealt with the criticism (Calhoun, 1992). The public sphere, therefore, became an arena of the bourgeois -white, middle class, male- who decided for all. The bourgeois became the traditional elite who defined the debate. They take a bigger role in directing the thinking of the public by acting as a spokesperson.

Nancy Fraser is only one of the strong critics of Habermas. She criticized all three characteristics on the Habermasian public sphere as introduced in the previous section. Her main points of critique regard the disadvantages of bracketing social inequalities, the thin line between public and private issues and the exclusion of women. My personal opinion regarding the public sphere, taken into account the criticism, is that the public sphere embodied in a coffeehouse is not the optimal place for forming public opinion. I also realize, however, how new the possibility was for open critical discussion in the coffeehouses in early Europe. So even though the public opinion formed in coffeehouses is not concerning the whole public and not everyone had an equal say, it is better than its predecessors.

Twitter as a public sphere

The internet has the potential to be a public sphere: open, democratic and egalitarian. Twitter is part of the internet so before looking at Twitter as a public sphere, the internet as a public sphere is briefly looked into. The internet is by some argued to have created room for discussion (Reese and Shoemaker, 2016) as a public sphere. Even further, it is suggested to be a virtual (re)incarnation of the public sphere “with an infrastructure that promises unlimited and unregulated discourse” (Papacharissi, 2008, 3). The authors of *The Permanent Campaign: New Media, New Politics* (Elmer et al., 2012) agree and argue how the internet has fewer central nodes, fewer gatekeepers and fewer agenda setters than traditional media, leading to a more horizontal and accessible playing field. The potential is there, for the internet, to be a new open, democratic and egalitarian public sphere.

On social media and especially Twitter, where users have an open and free space for discussion, the formation of a public sphere is even more facilitated than on the internet in general. Users can have conversations in the online space about issues that concern them all (Hoskins, 2013). Opponents of social media as a new public sphere argue how cultural capital (Docherty, 2015) and new media literacy (Papacharissi, 2008) are needed to participate in the online debate, besides of course internet access.

Cultural capital is mostly gained through an exclusive education system, excluding certain groups like in the eighteenth century (Docherty, 2015). New media literacy involves different skills that are crucial for living in the participatory (digital) society (Koc and Barut, 2016). These skills include reading and writing media and critical thinking (Koc and Barut, 2016). The need for cultural capital, new media literacy and internet access do not make Twitter automatically a public sphere in which all can participate.

There has been elaborate research on the Habermasian public sphere with its characteristics and critiques. Less consensus us on Twitter as a pubic sphere. The same goes for the elite: “Traditional leaders and featured actors withing public debate and the bourgeois

public sphere (journalists, the traditional media, and politicians) are often slower to officially adopt newer channels for discussion from the internet in general to specific platforms such as blogs, YouTube, or Twitter" (Bruns and Highfield, 2016, 126). This means that if the traditional elite is too slow to have a direct influence in the age of Twitter, others can fill that gap. For Twitter, these can be the so called influencers. They "can create a chain-reaction of influence that is based on word-of-mouth approach and reach to a very large scale of users" (Alp and Şule Gündüz Öğüdücü, 2018, 211). The large chain-reaction of influence can be seen as the glue of Twitter networks (Patel, 2014).

The internet and Twitter have, with their open, democratic and egalitarian possibilities the opportunity to be a modern public sphere. However, all those with a lack of internet access, cultural capital and new media literacy are excluded. Additionally, there are certain users on Twitter that take the role of the traditional elite. These so called influencers can be seen as the temporarily bourgeoisie with the possibility to steer public opinion.

4.1.2. Concept 2: Twitter causes a fragmentation of the public

This concept firstly explains fragmentation of the public and secondly applies this to Twitter. Fragmentation is the process of breaking parts off or detaching parts (Merriam-Webster, 2018c). The term used for fragmentation of the public sphere is public sphericule. Sphericules are "social fragments that do not have critical mass [but] share many of the characteristics of the classically conceived public sphere" (Cunningham, 2001, 135). The characteristics are, as described in section 4.1.1: bracketing social inequalities; new centres of institutional authority; and grating access to all. The public sphericules do not address the public as a whole but move around thematic debates (Bruns et al., 2015).

Habermas himself suggests, however, how different debates not necessarily lead to fragmentation and can even counter it: "a larger number of people tend to take an interest in a larger number of issues, the overlap of issue publics may even serve to counter trends of fragmentation" (Habermas, 2006, 244). This means that the overlap of issues counters fragmentation. An important distinction here has to be made. Habermas suggests that a large number of people tend to take interest in a large number of issues. These large number of issues do not necessarily exclude other issues, but when reading upon the attention economy in section 4.1.3, it turns out we have limited attention and no space for all the issues around. Additionally, there is only a limited amount of issues we discuss together and which we can align our actions upon. When only a limited number of issues is discussed in the public space, the quality of public debate can increase since issues are discussed more in-depth with more elaborate arguments. So in line with Habermas, different issues don't have to lead to fragmentation. But when attention is scarce, (digital) distractions are all around, we can only participate in certain debates and not in all.

Fragmentation on Twitter

This section looks at different sphericules on Twitter through a fragmented public sphere. Public sphericules are fragments of the public sphere with the same characteristics but a smaller public. Social media, which has grown exponentially during the course of the last decade (Warner-Søderholm et al., 2018), is not only used for social purposes as the name indicates but also to a large extent used to follow the news by instant updates. The instant updates cause small thematic debates. When thematic debates inside a public sphere become smaller, so do the different sphericules. Argued is how social media can, and often do, fragmentize political discourse (Papacharissi, 2002, 9). It is even argued how different sphericules consists of like-minded people and because the expertise stays within the sphericule, there is no exchange of information (Chen et al., 2018). Benjamin Barber argued already in 1998 about the consequences of new media:

> "I want to remark on the consequences of some key attributes of new media, including their speed, their reductive simplicity and tendency to (digital) polarization, [...] the solitariness of their user-interface, their bias towards images over text, their point-to-point, lateral immediacy and consequent resistance to hierarchical mediation, their partiality to raw data rather than informed knowledge, and their inclination to audience-segmentation rather than to a single, integrated community of users/viewers" (Barber, 1998, 1)

Even though this quote is not specifically about Twitter, is is very applicable. A scroll through my own timeline shows how many tweets contain images and not all tweets can be based on informed knowledge. It is even argued how some tweets with URL are retweeted before even clicking the URL (Orellana-Rodriguez and Keane, 2018).

Another applicable addition is made by McLuhans differentiation between hot and cool media (Chandler and Munday, 2011). Hot media are high in sensory data and ask for less sensory involvement. They are 'spoon-feeding' the content. Cool media need more sensory involvement and participation is key. Twitter can be argued to have characteristics of both. In a cold medium, participation is indeed key and the existence of the platform depends on its users. Argued is how active engagement makes social media more influential than traditional media (Unger et al., 2018). On the contrary, Twitter can be argued to be hot medium. Only a small community of users posts the most tweets, where all the others are passively taking in information (Kittle, 2011). One does not need to actively go around to fill in any blanks and one does not have to have the feeling they are missing out. This is due to the feeling that the information that reaches one is informed knowledge and thus one is informed.

This concept, *Twitter causes a fragmentation of the public*, looks at if fragmentation of the public sphere is happening. With that, the public becomes fragementized as well. Public sphericules are fragments of the public sphere with the same characteristics but a smaller public. Certain characteristics of new media- such as their speed, their reductive simplicity and their partiality to raw data rather than informed knowledge, can be argued to add towards fragmentation of Twitter. Additionally, Twitter can be argued to be a cold medium because the platform exists due to its participation and a hot medium because it spoon-feeds content.

4.1.3. Concept 3: Twitter creates space for issue publics

Issue publics are publics that emerge and disappear around topics as the topics come and go and Twitter can be argued, with its diversity in short-lived topics discussed, to create space for issue publics. John Dewey has written extensively on publics emerging around issues. The public is, according to Dewey, that which forms itself around an issue and consists of anyone who is effected by or concerned with the issue (Dewey and Rogers, 2012). It is not a pre-existing group that is capable of democratic decision making, nor is it something formed by the various political institutions such as parliament, the state, law or national election. It is an emerging public. When we base issue publics on the notion of publics forming around issues by Dewey, it can be argued that publics can only form around a limited number of issues. Especially when we want to achieve coordinated action on these issues. We have to align and coordinate the issues between individuals, and to align, we have to decide upon issues to discuss. When only a limited number of issues is discussed in the public space, the quality of public debate can increase since issues are discussed more in-depth with more elaborate arguments.

Besides the sociological claim based on the process of public debate, there is also the natural limitation of our brain. Microsoft Canada's claim introduced before (how our current average attention span is smaller than the average attention span of a goldfish) is hard to prove and might be misleading. But Webster points to the scarcity of attention: "the supply of human attention is limited. As more media compete for attention, the audiences they seek become relatively scarce" (Webster, 2014, 1). The limited supply of human attention means that we can only engage with a limited number of issues, based on the limitations of our brain. We have, due to the internet, more and cheaper (or even free) access to information, but there is only so much we can process even when we are fully focused. And being in a state of full focus is scarce, since on social media, many different distractions are vying for our attention (Atchley and Lane, 2014).

Even though there are limitations to the amount of issues in public debate, based on the process of public debate and our limited brains, there still are different topics entering our mind and public debate. These issues can be argued to create space for issue publics. "Issue publics form especially around shorter-lived topics and events and are thus considerably more temporary and dynamic than some of the other formations we have already encountered - they "emerge, exist for varying durations, and then eventually dissolve" (Dahlgren, 2009, 74) as public debate moves on" (Bruns and Highfield, 2016, 108). This means that issue

publics are shorter lived and more dynamic than the previously mentioned public spheres (section 4.1.1) and public sphericules (section 4.1.2), having a clear emergence, existence and disappearance as issues come and go. Issue publics can be seen as subsets of the wider public spheres and public sphericules.

Twitter creates space for issue publics

This section looks at issue publics on Twitter. Before looking into these, it is import to define attention on Twitter. The definition used here is based on the definition by Orellana-Rodriguez and Keane (2018): "We use the term attention to cover all forms of engagement with a news-tweet whether that is simply looking at/reading the tweet (i.e., impression) or actively engaging with it (i.e., by liking, quoting, or retweeting it)" (79). This definition shows how different forms of engagement can happen on Twitter, from which some are easier to measure than others (e.g., retweets are counted below a tweet but it is not easy to see how many users have read your tweet). Two functions on Twitter can be argued to create space for issue publics: retweeting and hashtags. Before explaining these two concepts, it is important to note how the traditional public spheres theory is based on a domination by the mass media. The theory of the public sphere is based on a time when the internet did not exist yet. Today, more specific publics can form around more narrowly defined interests with new forms of media such as social media. Twitter can create a space for specific debates between actors with similar interest, whereas in the time of the mass media, it might not have been as easy to find different people interested in all your different interests. An example is someone interested in Trump, climate change, racism and the future of food. On Twitter, this person can join different debates on the different topics. In an offline world, it is more difficult to find people around you interested in all these different topics.

When looking at the two functions that can possibly create space for issue publics, the first function is retweeting. Retweeting can leads towards Taylor's imaginaries (Taylor et al., 2004). Imaginaries are describing, among other things, the way in which people imagine their social existence and how they fit together with others (Pesch, ndb). This imagination, or perception, can generate a common interest (to be) and therewith create a public. Individuals can see themselves as a greater whole, a greater public. It is likely that retweets often happen based on the user's picture of the imagined audience (Bruns and Highfield, 2016). By retweeting ones imaginaries, one creates a debate around an issue. Retweeting can be argued to function like echo chambers where users only retweet others with similar ideas. The second function, the hash symbol # is called a hashtag and is used before any keyword as an attempt to kick-start or join a discussion on a theme of interest. A hashtag itself can be described as a "linguistic segment as well as a clickable hyperlink" (Giaxoglou, 2018, 13), making it possible to form a group around a certain hashtag. The group formed by a hashtag can be described as an ad hoc public (Bruns and Burgess, 2011).

An internet skeptic called Morozov has described the relation between issues and publics even more outspoken in his book *To Save Everything, Click Here* (2013). He describes Twitter's trend feature as censorship. The feature, as described by Twitter, tailors trends by an algorithm just for you, based on your interest (Mauskopf, 2012). This algorithm creates publics around issues. Morozov says the following:

> "Consider Twitter's Trends feature, a filter that relies on several signals and algorithms to determine which topics are "trending" across the platform. Whether something becomes a trend on Twitter is important for one simple reason: once the story achieves this much coveted status, it attracts even more attention, spilling into national and global conversations far beyond Twitter. In this sense, Twitter too is an engine, not a camera; it does not just reflect realities—it actively creates them. Consequently, when some important discussion generates a lot of buzz but never rises to the status of a trend, it's very common to see Twitter accused of censorship. This was the case with the Occupy Wall Street discussion—much of it happening under the #OccupyWallstreet tag—which continuously failed to register on Twitter's radars" (Morozov, 2013)

This quote shows how the determination of 'trending' topics is not transparent. Twitter creates instead of reflects realities, as an engine instead of a simple camera. On Twitter,

the popular issues create a public around them. These publics form either through Twitter's trending feature; because they actually are discussed frequently; or through a combination of the previous two. Twitter's trending feature shows how Twitter theoretically is in the position to create space for issue publics, and therein also create issue publics that suit the platform.

4.2. Constructs

The constructs used in this thesis are operationalizations of the concepts from section 4.1 and can be seen as the building blocks on top of the concepts. See figure 4.1. Operationalizing the concepts happens with the help of characteristics specific to the construct and as preparation to link the data with the theory. The following sections all describe a construct based on a specific concept. Section 4.2.1 translates concept 1: *Twitter is an incarnation of the Habermasian public sphere* into a construct, section 4.2.2 translates concept 2: *Twitter causes a fragmentation of the public* into a construct and section 4.2.3 translates concept 3: *Twitter creates space for issue publics* into a construct. These constructs all contain certain characteristics to operationalize the theory. Again, it is important to emphasize the ideal-typical character of these characteristics. The constructs are going to be applied to the case study of palm oil. Table 4.1 is a summary of the different concepts/constructs with their characteristics and visualization.

4.2.1. Construct 1: Twitter is an incarnation of the Habermasian public sphere

The first construct is based on the Habermasian public sphere. In brief, the Habermasian public sphere is an open, democratic and egalitarian place where something approaching public opinion is formed. To turn the first concept into a construct, three characteristics and one additional prerequisite (4) are distilled from the concept of the Habermasian public sphere in section 4.1.1. After mentioning these four characteristic, the sections below will describe a more specific construct per characteristic.

1. Social inequalities are bracketed
2. Public sphere is a new centre of institutional authority
3. Access is granted to all citizens
4. Those forming public opinion are presupposed reasonable

1. Social inequalities are bracketed

When social inequalities are bracketed, one can speak to one other regardless one's social status. This is, in the most basic sense, the case at Twitter. Anyone can send a message to president Trump, retweet his tweet or reply to what he says. It implies everyone is equal while at the same time, the fact that Trump would most probably not follow one back leads can be interpreted as a non-egalitarian connection. To clarify, Trump probably does not know of your existence whereas you actively follow him. Two checks are conducted to see if social inequalities are bracketed.

Firstly, to map the public sphere of Twitter to see what actually happens with regards to status, the network of those debating in the palm oil debate is mapped. Secondly, in line with Fraser's critique on how bracketing social inequalities only works to the advantages of already privileged groups in society and how the normal people cannot participate but can only look from the sideline. To translate this division to the age of Twitter, influencers in the debate are mapped. They might not be the traditional elite- the bourgeoisie: white, middle class men- but this section is going to map who they are. The network and the influencers cannot be seen separate, since influencers are part of the network, though in a more central role than some other users.

An increasing part of the Twitter network is currently based on users posting spam and autonomous posts. Users can be divided into humans and social bots, which can both be divided further into legitimate and spam (Inuwa-Dutse et al., 2018). See figure 4.4 for a visualization of the different users. Legitimate users operate within the Twitter usage policy and interact with moderate frequency. Spam users act hyperactive and often produce irrelevant content. Additionally, they potentially violate Twitter's term of use (Twitter, 2018d).

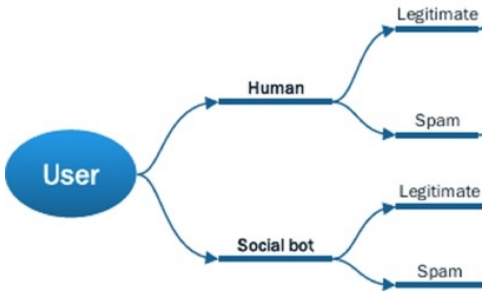


Figure 4.4: Different types of Twitter users (Inuwa-Dutse et al., 2018, 507)

The network

Mapping the Twitter network can show insight in all those participating in the debate on palm oil. There are multiple ways of plotting the network. One of these ways is plotting who is all replying to who, to create insight in key players in the network. It can also show what gave users their key role, either because they have replied to frequently or are being replied to differently.

Finding influencers

Influencers have to be defined before they can be found. Influencers are those who contribute most to the diffusion of information (Alp and Şule Gündüz Öğüdücü, 2018, 212) and can create a chain-reaction of influence, to reach a lot of users (Carter, 1956). Argued is how 1% of the twitters users, serving as influencers, controls a fourth of the total information diffusion (Alp and Şule Gündüz Öğüdücü, 2018). An attempt to translate influencers into a construct is based on user-level resources and tweet-level resources (Francalanci and Hussain, 2017), see table 4.2. User-level resources say something about the user like the number of followers, whereas the tweet-level resources say something about their tweets like the number of retweets. A combination of user-level-resources and tweet-level-resources can be interpreted as the level of influence a user has.

User-level Resources	Tweet-level Resources
Following	URLSs
Followers	Hashtags
Favorites	Retweets
Tweets	Favorited
	Mentions

Table 4.2: User- and tweet-level resources (Francalanci and Hussain, 2017)

To find the influencers in the palm oil debate, these resources will be mapped for different users. Not all resources are equally important. A user with a high in-degree (high number of followers) is generally more seen as an influencer, since that user can reach more people. It is, however, not necessarily true that the most socially influencing person is the person with the most followers (Alp and Şule Gündüz Öğüdücü, 2018). In conclusion, a combination of key users showing up in the network analysis and the influencers based on the different resources can be argued to be influencers. These influencers can be seen as the contemporary bourgeoisie, in the position to steer public opinion.

2. Public Sphere is a new centre of institutional authority

This section is aiming to explain how to find out if Twitter is a new centre of institutional authority, like the Habermasian public sphere. This is based on interviews and on analyzing the data. Firstly, policymakers are interviewed. The goal of these interviews is to find out if the government is using social media data (especially Twitter) to measure public opinion, and if so, how and on what topics. Additionally, the questions aim to find out if policy makers have seen movements as counterweights to the traditional institutional authority (them). Counterweights can show by speaking out against certain structures and calling for structural change as has been the case in many revolutions in history.

Secondly, it is interesting to see if social media in itself can be the means of authority, so instead of using the data, forming a new centre. Hayes shows with her case how the public has taken over a social media campaign from the New York Police Department (NYPD) (Hayes, 2017). In this case, the NYPD has started a campaign on social media to rise awareness throughout the city about the hard and important work the police was doing. This backfired and turned into a protest on social media against police brutality by the public. Depending on how much power a public is getting, it *could* become a new centre of institutional authority. Another way to actually become the centre of authority can be seen from massive behavioral changes like a boycott on the use of palm oil. People can make multiple calls for action on Twitter: they can address multinationals to stop making products with palm oil, they can address governments to regulate better or they can call on their fellow citizens to stop using palm oil as consumers.

To conclude, a combination of these to analyses aims to create insight in Twitter as a new centre of institutional authority. Interviews with policymakers show if they use social media data and if they see the organization of a counterweight. Analysis of the tweets shows if a public is organizing like in the case of the NYPD or if there are frequent calls for action.

3. Access is granted to all citizens

It is important to note that the public debate on Twitter does not include the whole public on the issue of palm oil. The offline debate, and the online debate on other media, cannot be forgotten, but are excluded in this thesis. This exclusion makes it very difficult to grant access to all citizen, since not have access to Twitter. As argued in section 4.1.1, cultural capital (Docherty, 2015), new media literacy (Papacharissi, 2008) and internet access are necessary to participate in the debate on Twitter.

The prerequisites of internet access, cultural capital and new media literacy are limitations of Twitter as a representative of a public sphere and therefore of the public debate on Twitter. The opinions expressed on social media, however, can represent an activated public opinion which in turn can anticipate trends in public opinion (Andrea and Fedra, 2016). To see who all has a say in the debate the users participating in the debate on palm oil are analyzed on a deeper level. The different actors will be grouped and mapped, to see what kind of actors are participating and if groups or individuals are excluded. To see who is participating, a distinction is made between parties participating in public or in private interest. An example of participating in private interest is narrating a strategic argument to benefit your company instead of an argument to benefit society.

4. Those forming public opinion are presupposed reasonable

Based on the rational-critical debate, where voluntas are translated into ratios, people are presupposed reasonable. Reasonable means having reason, and reason is a statement offered in explanation or justification (Merriam-Webster, 2018g). These statements should, according to Habermas, not be based on an ideology or power (Kuitenbrouwer, 2018). To find out how reasonable those forming public opinion are, tweets will be analyzed manually. The top ten retweeted tweets will be taken as case studies, to see if the authors are acting reasonable. This will be based on the arguments in a tweet, if an image or video is used, and how positive or negative a tweet is (the polarity).

The analysis of tweets conducted in this characteristic aims to create insight in the explanation and justification within a tweet. The arguments, use of media an polarity all add up to the possible explanation in a tweet. Additionally, the tweets are checked for ideological arguments or arguments made out of power.

4.2.2. Construct 2: Twitter causes a fragmentation of the public

The second construct is based on a fragmentation of the public. In brief, Twitter can cause a fragmentation of the public due to the many thematic debates that are happening. The public sphere breaks apart in separate fragments called public sphericules, which have the same characteristics as public spheres. These are the bracketing of social inequalities and new centre of institutional authority. There are two characteristics in the concept of fragmenting the public. Only if the first characteristic is true, the second one can be analyzed:

1. The public on Twitter is fragmented
2. The fragmented public on Twitter is caused by Twitter

1. The public on Twitter is fragmented

The public on Twitter can be seen as fragmented because they are part of different sphericules. These are social fragments with like-minded people and not much information exchange to other sphericules. To check if the public sphere is divided into public sphericules, the first check is a count of the hashtags. Hashtags give an impression of the topics discussed frequently and can be seen as the main topics defining public sphericules. Generally, only 20% of the tweets contain hashtags (Orellana-Rodriguez and Keane, 2018). Therefore, the second check is a word frequency check. The word frequency also gives an impression of the topics discussed and therewith the public sphericules, especially when analyzed in combination with the hashtag frequency.

After these two basic checks to get an idea of the data, a more advanced technique of topic modelling is used, mainly to see if no main topics have been missed. When there are multiple ways of discussing a certain topic, they will not show up at the hashtag count but they will at topic modelling. An example is the word computer. There are many words that are closely related or synonymous for computer, such as pc and laptop. The main assumption in all topic models is the existence of a topic level beyond the observable word level (Li et al., 2018b). The topics shown can be interpreted as the different public sphericules. The topic models used are based on unsupervised machine learning, meaning there are no set topics beforehand. This technique can make the topics generated difficult to explain. On short texts, however, topic modelling is difficult due to the noise problem (Li et al., 2018a,b). The noise problem consists of missing, misspelled or incorrect terms. On Twitter, there are a lot of abbreviations used which add to the noise problem. This can be partly overcome by excluding Twitter-specific words like 'RT' (retweet), 'U' (you) and 'clk' (click).

To see if the public on Twitter is fragmented -a prerequisite for the next characteristic, where this fragmentation is caused by Twitter- three analyses are done. Firstly, hashtags are counted and can be seen as different public sphericules. Secondly, the same is done for certain words. Thirdly, the more advanced technique of topic modelling, in which the topics can be seen as sphericules, is conducted to see if no topics have been missed in the first two analyses.

2. The fragmented public on Twitter is caused by Twitter

After there are signs of fragmentation of the public sphere on Twitter, this characteristic checks if the fragmentation is caused by Twitter. This characteristic is based on the argumentation of Barber (1998) on the consequences of new media. From these, three applicable features are taken. The top ten retweets will be tested by these three features (image use, raw data and segmentation).

1. Their bias towards images over text

   Digital media are argued to be carriers of image and sound instead of word and thought. The famous idiom "a picture is worth more than a thousand words" says it all. Quick images, however, might not be conducive to your own thinking. These images can be argued to spoon-feed the message. No own interpretation or filling in blanks is necessary.

2. Their partiality to raw data rather than informed knowledge

   Raw means before any preprocessing. Raw data is generally not usable and difficult to interpret. When it is, it is not raw anymore but filtered, selected, edited and imbued

with meaning. The idea that the raw data does not need to be filtered or interpreted will lead to a data-overload rather than more knowledge. New media are often based on raw data rather than interpreted data.

3. Their inclination to audience-segmentation rather than to a single, integrated community of users/viewers

   Digitization prefers parts of knowledge over integrated knowledge or so called information snippets. Data can be interpreted in the way most convenient for the individual but without creating much common ground.

Different tweets are analyzed based on these different features. These features can be combined with the theory of McLuhan on hot and cool media (see section 4.1.2). Hot media are high in sensory data and ask for less sensory involvement. They are 'spoon-feeding' the content. Cool media need more sensory involvement and participation is key. Tweets which contain multiple features are pointing towards hot media, since only a small community of users posts most of the tweets, whereas other are taking information in like warm cookies. A combination of three features as described by Barber (1998) -the use of images, raw data and snippets of information- attempts to answer the question if the fragmentation visible on Twitter is caused by Twitter and the nature of tweets. These features can be argued to spoon-feed content, making it a hot media, according to Chandler and Munday (2011).

4.2.3. Construct 3: Twitter creates space for issue publics

The third construct is based on issue publics. In brief, issue publics are publics that emerge and disappear around issues as these issues come and go (Dahlgren, 2009). To see if Twitter creates space for issue publics, this construct is divided into two characteristics. The first characteristics is based on the possible functionalities of tweets, whereas the second characteristic builds upon the theory of the attention economy. The attention economy consists of two parts. The first part considers the limited human attention due to our brains, whereas the second part considers the limitations due to the process of public debate. The result of these two parts is that our attention is scarce and different parties are calling for it. With scarce attention comes our limited capability to talk about a selection of issues and not about all issues. This construct checks if, because one issue is emergent, other issues get snowed under.

1. Twitter functionalities (retweets and hashtags) create space for issue publics
2. Certain issues push others away in a time where attention is scarce

1. Twitter functionalities (retweets and hashtags) create space for issue publics

Before finding out if Twitter functionalities create space for issue publics, the existence of issue publics on Twitter is investigated. To do so, tweets are mapped against the time to see if they are peaks in the tweets. If peaks exists, the three characteristics of issue publics (they emerge, exist for varying duration, and then eventually dissolve) are applied to different peaks. Subsequently, the cause of these peaks is investigated and checked if it is caused by certain retweets and hashtags. If the peaks are caused by retweets and hashtags, Twitter functionalities can be argued to create space for issue publics.

The line between issue publics and sphericules can be thin. Certain topics might exist longer and pop up in different forms over the whole period of tweets gathered, whereas others might only exist briefly. It is also possible that inside public sphericules exist issue publics. Hashtags -with the aim to create a debate around an issue- can be used for both, depending on the situation and the scope of the hashtag. An example of a hashtag for an issue public is #NetherlandsFirst, which was a frequent hashtag during the Winter Olympics to take a shot at president Trump. This hashtag was only tweeted during the Olympics and more specifically, when the Dutch were winning or had won. Therefore, it depends on the situation if a hashtag is causing an issue public or a sphericule.

This characteristic aims to find out if retweets and hashtags create space for issue publics. To do so, firstly peaks in the amount of tweets are found and secondly, these tweets are analyzed to see if they are issue publics. If they have a clear emergence, existence and

disappearance, they can be argued to be issue publics. The public disappears as the issue disappears.

2. Certain issues push others away in a time where attention is scarce

This characteristic build upon the previous one by analyzing the tweets in peaks to see what happens to tweets on other topics. They can either push each other away (1) or boost the discussion as a whole (2). Pushing away can happen when there are no other issues discussed than the one that is peaking. The whole discussion can get a boost when a very popular hashtag or retweet leads to an increased amount of discussion on the whole debate. This can lead to a larger total share of tweets on palm oil on Twitter.

It can be checked if issues pushing others away or boosting the whole debate can happen within the sphere of palm oil, but not with subjects out of this sphere since there are no tweets gathered on other topics. This can be seen as a limitation of this thesis, but checking what happens within the sphere of palm oil is a first step. Retweets and hashtags are only one way to check the effect of tweets vying for our attention. Not all forms of attention can be measured. Simply looking at or reading a tweet cannot be measured in this thesis.

Additionally to the natural development of certain issues and the use of hashtags and retweets, Twitters trend feature can amplify the debate on certain topics. Twitter can use that feature to shine light on certain topics and not on others. Since Twitter's Trend feature is an nontransparent algorithm (Morozov, 2013), it is unclear what the effect of this feature has had on the debate and therewith public opinion formation on Twitter. It is not possible to measure what the effect is of making a topic trending even though that could be insightful due to the nontransparency of the feature.

In conclusion, this characteristic aims to show the effects on topics when taking the attention economy into account. This thesis can only see what the effect is within the palm oil debate. Three options are possible: issues push each other away, issues boost the bigger theme they are part of (such as #NetherlandsFirst boosting #Olympics) or there is no influence based on issues based on their fluctuating popularity.

5

Finding relevant Tweets

To answer the second sub-question, *What are relevant tweets regarding palm oil?*, tweets are gathered and basic analyses are conducted. Figure 5.1 shows the selection of the framework from chapter 3 that accompanies this chapter highlighted in green. Phase D, data gathering, is highlighted. Data gathering means tweet gathering in this thesis and happens by an iterative process.

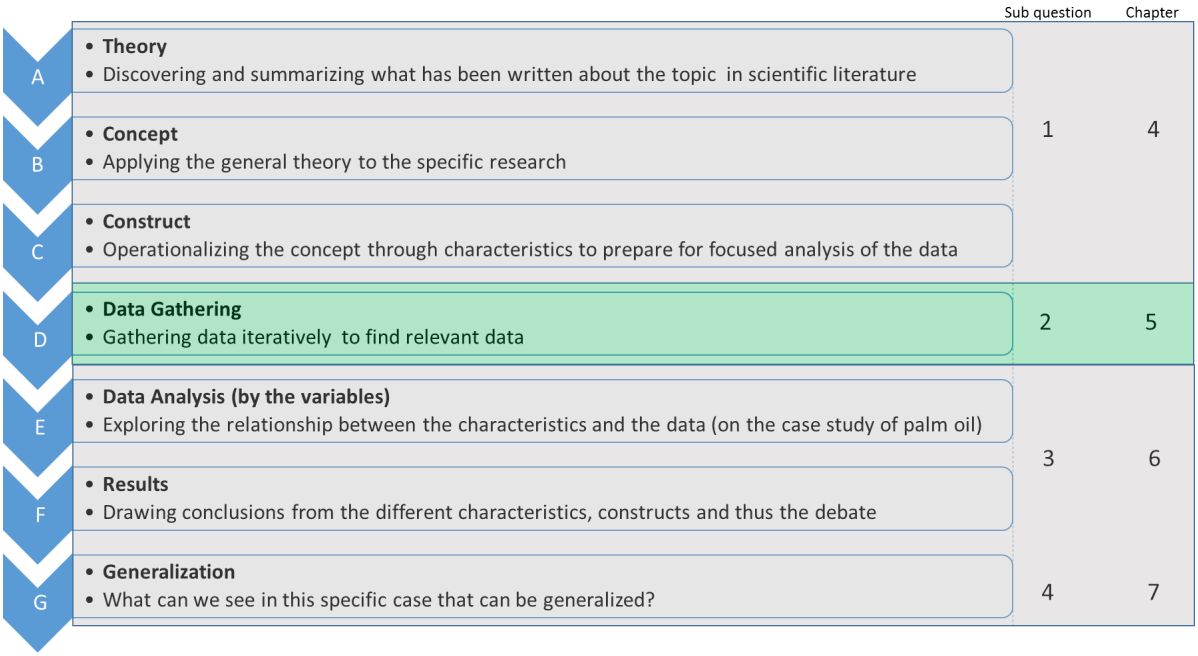


Figure 5.1: Framework introduced in chapter 3 with the phases D highlighted

A visualization of the iterative progress can be seen in figure 5.2. There are five steps involved: installing the software, defining hashtags and keywords, spot-check preliminary tweets, install a server and get a first impression of the tweets. All these steps are elaborated upon in the next subsections. See `https://github.com/sevanrossum` for the corresponding code.

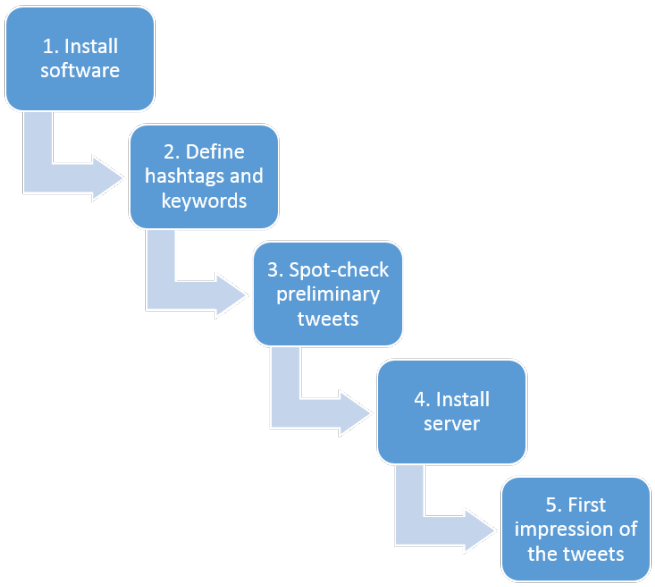


Figure 5.2: Iterative process of collecting tweets

5.1. Step 1: Installing the software

Python is an open source, easy to use and widely accepted programming language (McClarren, 2018). Multiple libraries are installed to analyze and visualize the data, which can be found in appendix A. The Tweepy library is installed for accessing Twitter data. Tweepy uses the Twitter Application Programming Interface. The Twitter API used searches against a sampling of recent Tweets published in the past 7 days (Twitter, 2018c). It returns a collection of relevant Tweets matching the query of hashtags and keywords as described in section 5.2. All the collected tweets contain lots of information, like the screen name, user ID, hashtags, text, followers count, following count, location, description, and more. See for the exact information gathered per tweet appendix B. Important to note is how the so-called standard search API is based on relevance and not completeness, meaning that not all tweets might show up in the search results. This is not expected to be a problem in my case since the data is meant to link the theory to Twitter in an explorative way. If the case study was playing the most important role in this thesis instead of understanding the debate, I would have used a paid version of the Twitter API that returns a complete set of tweets instead of only the ones classified as relevant by Twitter. After the first step is finished, all the preparations are taken to gather, analyze and visualize the data.

5.2. Step 2: Defining hashtags and keywords

Hashtags and keywords are identified based firstly on literature on palm oil and secondly on interviews with experts. Section 2.6 gives an introduction to the literature. The interviewees are coming from the government, science and business, to cover as much as possible on the palm oil debate. From government, mister van der Elst, a policymaker from the Ministry of Agriculture, Nature and Food Quality is interviewed. His focus on on greening agricultural supply chains. From science, mister Lindenboom is interviewed in an informal way to direct this thesis. He aims to help make the industry more sustainable for environment and people, based on technological solutions. From business, mister de Legé from Palmares Partnership

aims to increase value from waste in the palm oil chain in Malaysia. A summery of these interviews can be found in E.Based on their input, the hashtags and keywords in table 5.1 are defined.

Keywords	Hashtags
palm oil	#palmoil
oil palm	#oilpalm
palmoil	#rspo
palm oil plantations	#RefinedPalmOil
rspo	#palmolie
mpob	-
palmolie	-
minyak kelapa sawit	-
aceite de palma	-

Table 5.1: Keywords ans hashtags used to collect tweets

Abbreviations used in the keywords and hashtags are Roundtable on Sustainable Palm Oil (RSPO) and Malaysian Palm Oil Board (MPOB). Palm oil in different languages is palmolie (Dutch), minyak kelapa sawit (Malay) and aceite de palma (Spanish). Tweets with these hashtags are collected manually from July 23 - 30. After the second step is finished, a list of keywords and hashtags is defined to gather tweets regarding palm oil during the first week.

5.3. Step 3: Spot-check preliminary tweets

Spot-checking preliminary results happens by looking at the tweets and see if there are notabilities. Almost 10% of the first 16.000 tweets gathered over the first week is a retweet of the tweet shown in figure 5.3. This tweet does not have to do with palm oil, but with palm sugar and cold pressed oil. Since the search query also searches for the separate words, this tweet (with its retweets) has to be filtered out. There are no other notabilities so the other hashtags and keywords can be used for the next two months. These first 16.000 tweets contain enough tweets with hashtag palm oil. 52% of the tweet do not have hashtags, emphasizing the importance of also searching for keywords. 95 tweets per hour are collected on average. The exact information gathered per tweet can be found in appendix B. After the third step is finished, the tweets gathered over the first week are analyzed.

Figure 5.3: Tweet by Actor Karthi (retrieved September 11, 2018)

5.4. Step 4: Installing a server

A server is installed to collect tweets 24 hours a day. A virtual server via Digital Ocean is used. This already has Python 3 installed, so only Tweepy needs to be installed and the scripts with the hashtags and keywords need to be uploaded. The program 'Screen' is used to let the project run for 50 days. The technical details of the server can be found in appendix C. Once the Server is up and running, tweets collected will be saved in a so-called JSON file (JavaScript Object Notation) per 500 tweets so they are ready for analysis. The server is running from July 30 to September 11. After the fourth step is finished, the server is set up to gather tweets non-stop.

5.5. Step 5: Getting a first impression of the tweets

Tweets are gathered for 50 days. This section plots the tweets over time, analyzes the accounts and languages, and visualizes word- and hashtag frequencies in wordclouds. 86 tweets are collected on average per hour, leading to a total of 103.500 tweets. Figure 5.4 shows the amount of tweets plotted over time. There are multiple peaks visible, which will be analyzed further in chapter 6. There are 65.417 unique accounts in these tweets. Of these accounts, 75% is English, 12% in Spanish and 3% in Indonesian. The rest are in many different languages, 45 in total, and the next three languages are Dutch, French and German.

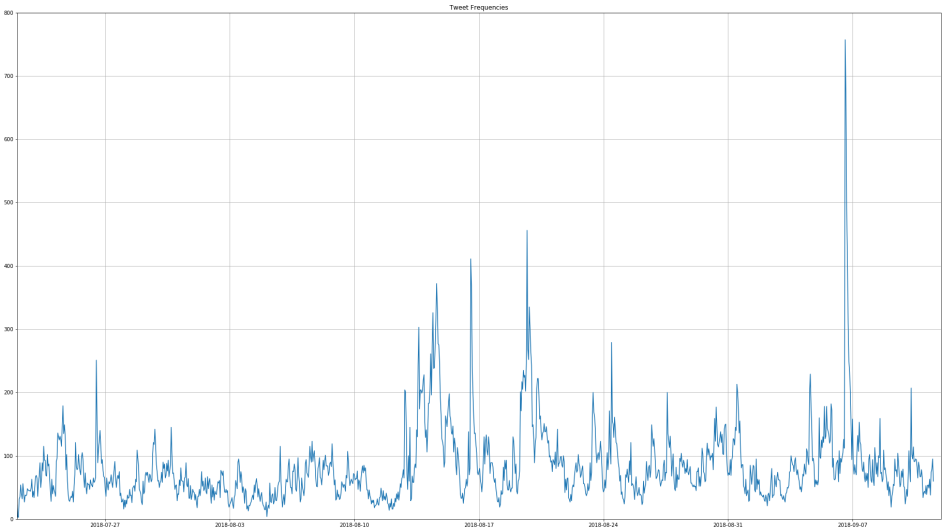


Figure 5.4: Tweets per 12 hours, from 22/07 - 11/09

The word clouds in figure 5.5 and 5.6 give a first impression of the most frequent words and hashtags in the tweets gathered. The words in the word cloud shown most prominently have been mentioned the most. As can be seen in figure 5.5, besides the obvious palm, oil, #palmoil, palma and aceite (Spanish for palm and oil), destroyed, orangutans and @bpdp_sawit show up prominently. The @ in @bpdp_sawit means that people have mentioned bpdp_sawit, either to send them a public message or to refer to them. bpdp_sawit stands for Badan Pengelola Dana Perkebunan Kelapa Sawit and is an Indonesian palm oil plantation fund. As can be seen in figure 5.6, the most frequent hashtags are palm oil, plantation, orangutans, sawitbaik and biodieselhematdevisa. Orangutans are often mentioned to be endangered due to deforestation and therewith loss of their living environment. Sawitbaik means in Indonesian 'palm is good', and is often tweeted in combination with the Indonesian palm oil plantation fund bpdp_sawit. Biodieselhematdevisa means 'biofuel to save foreign exchange' and is

also used often in tweets and retweets by bpdp_sawit on the topic of the new Presidential Regulation on biofuel, to decrease fuel imports (BPDP, 2018). After the fifth step is finished, 103.500 tweets are gathered, a first impression is defined and the tweets are ready for further analyses.

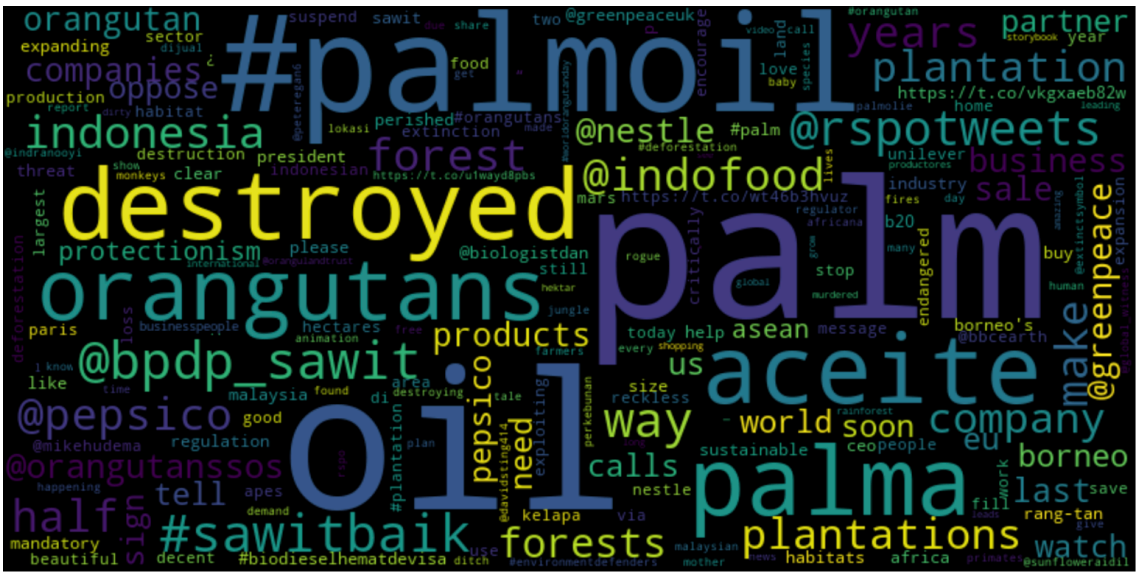


Figure 5.5: Wordclouds of most frequent words, from 22/07 - 11/09. The bigger the word appears in the image, the more frequent the word has been used

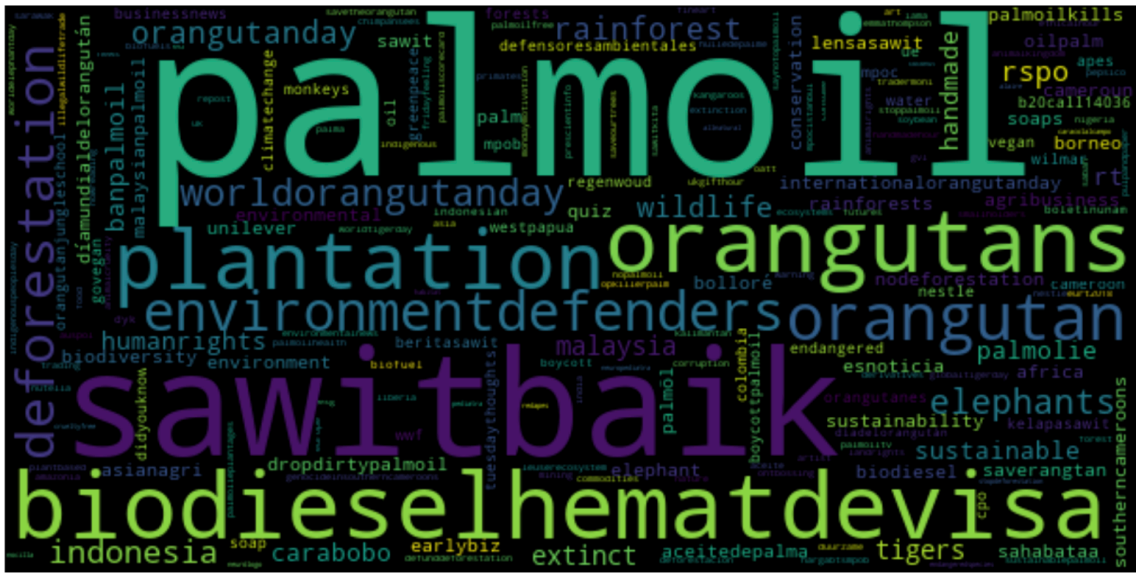


Figure 5.6: Wordclouds of most frequent hashtags, from 22/07 - 11/09. The bigger the word appears in the image, the more frequent the hashtag has been used

To conclude, the process of data gathering happens in an iterative way and is visualized in figure 5.2. After the first step is finished, all the preparations are taken to gather, analyze and visualize the data. After the second step is finished, a list of keywords and hashtags is defined to gather tweets regarding palm oil during the first week. After the third step is finished, the tweets gathered over the first week are analyzed. After the fourth step is finished, the server is set up to gather tweets non-stop. After the fifth step is finished, 103.500 tweets are gathered, a first impression is defined and the tweets are ready for further analyses.

6

Driving the debate

To answer the third sub-question, *What characterizes public opinion on palm oil on Twitter?*, tweets are analyzed based on the constructs. Figure 6.1 shows the selection of the framework from chapter 3 that accompanies this chapter highlighted in green. The phases E and F are highlighted. The previous phased included theory, concepts, constructs and the data gathering. This step combines the steps A-D to come up with results based on the palm oil debate.

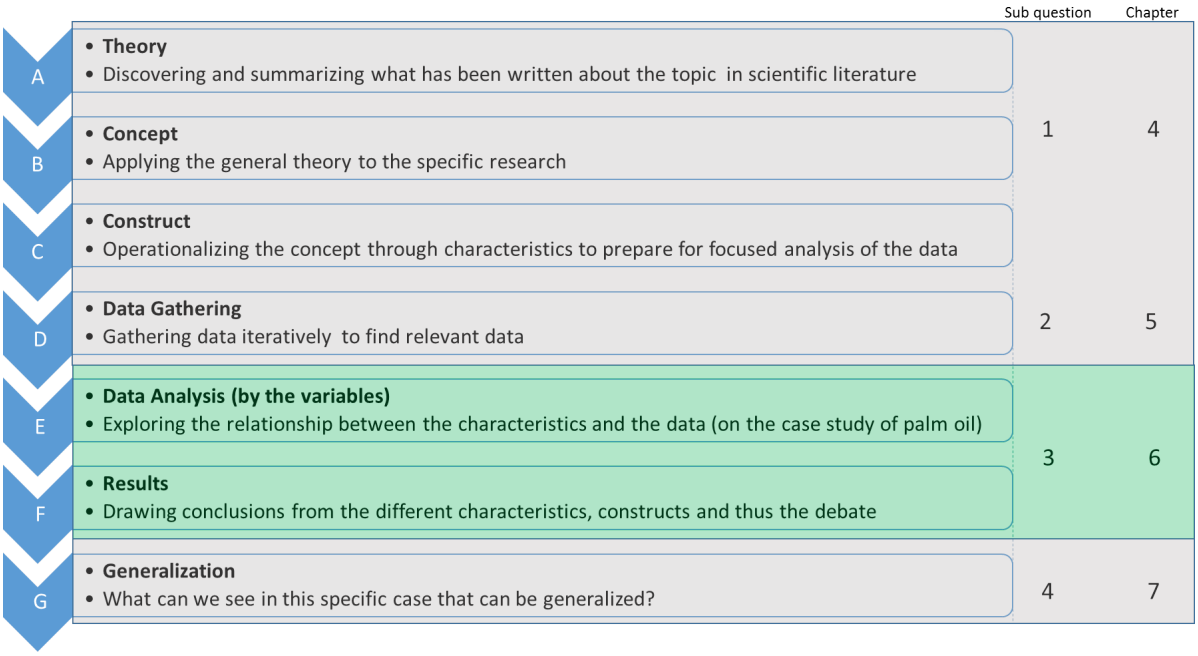


Figure 6.1: Framework introduced in chapter 3 with the phases E and F highlighted

The different sections in this chapter build further on the different concepts and constructs. Section 6.1 builds further on section concept and construct 1: Twitter is an incarnation of the Habermasian public sphere. Section 6.2 builds further on concept and construct 2: Twitter causes a fragmentation of the public. Section 6.3 builds further on concept and construct 3: Twitter creates space for issue publics. All these sections make use of the data gathered as described in chapter 5. Before explaining the results per construct in depth, the next section sketches a rough overview of characteristics found based on the constructs. Additionally, the most frequently retweeted tweets are introduced.

Overview

The constructs can be seen as the lens through which Twitter is analyzed. See table 4.1 for the overview of the concepts-constructs with their characteristics. The first construct is: *Twitter is an incarnation of the Habermasian public sphere.* The public sphere is where something approaching public opinion is formed. The egalitarian base of the public sphere preaches that all humans are equal in fundamental worth and thus in social status. Where the bourgeoisie in the Habermasian public sphere were the most influential actors to form and influence public opinion, multinationals and NGOs take that role on the Twitter debate on palm oil. Additionally, fake accounts try to pollute the debate by giving certain actors more influence than others in a certain direction. An example from the palm oil debate is the large amount of fake accounts that have tried to promote the ideas by the Indonesian investment fund Badan Pengelola Dana Perkebunan Kelapa Sawit. Besides the fake accounts and other types of spam, many tweets use other measures to grasp the attention of Twitter users. Many tweets contain a video or image to gain attention. To conclude, Twitter has the potential to be a Habermasian public sphere when taken the limitations of the traditional public sphere into account (such as the exclusion of some and the role of the elite). Spam, however, has a polluting effect and nothing like it existed in the traditional public sphere.

The second construct is: *Twitter causes a fragmentation of the public.* There are certain features that people use in Tweets to win the war on our limited attention and therewith fragment the public, such as retweets and hashtags. It can be seen, however, that the public on Twitter in the case of palm oil on the highest level is not very fragmented but on the lower level it is. Higher means more high over like a helicopter view, whereas lower means on a specific topic. Not much fragmentation is present on the higher level (most tweets are about deforestation) but there is on the lower level (such as the orangutan, Indonesia and RSPO). These different tweets are not covering the whole debate on palm oil but only a fraction. The features of tweets that can cause fragmentation are the use of snippets of information, the interpretation of data and the use of images and videos. All of these features are frequently used to influence public opinion. To conclude, the public on Twitter can be argued to be fragmented on very specific topics but more on one line on the highest level of deforestation. Features of tweets cause the fragmentation on lower levels.

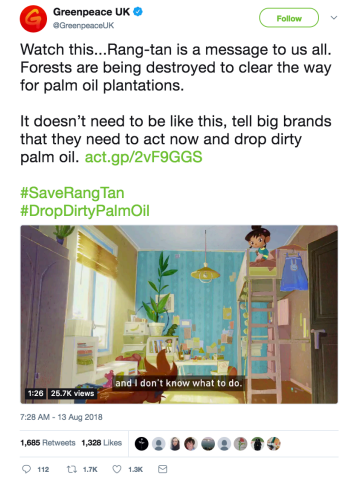
The last construct is: *Twitter creates space for issue publics.* Issue publics are publics that emerge around issues, and issues are short-lived and dynamic topics. Certain peaks are visible (when many tweets were posted in a small time slot) and some of these are caused by specific Twitter functionalities such as retweets and hashtags. Other peaks do not have a clear cause and can have many different smaller causes. Twitter definitely creates space for issue publics.

Retweets

Multiple analyses in this chapter are based on the top ten most frequently retweeted tweets. To give an idea of what these tweets are, see the figures 6.2 and 6.3. These will be referred to throughout the whole chapter. Three tweets are by the same organization, the Indonesian Investment Fund Badan Pengelola Dana Perkebunan Kelapa Sawit (tweets in figures 6.2a, 6.2b and 6.2d). The other tweets are by Greenpeace UK (figure 6.2c), BBC Earth (figure 6.2d), Mike Hudema (a Greenpeace activist, figure 6.2f), Stephen Fry (a writer and activist, figure 6.2g), Peter Egan (an actor, figure 6.2h), Daniel Schneider (a biologist and activist, figure 6.3a) and Greenpeace (figure 6.3b).

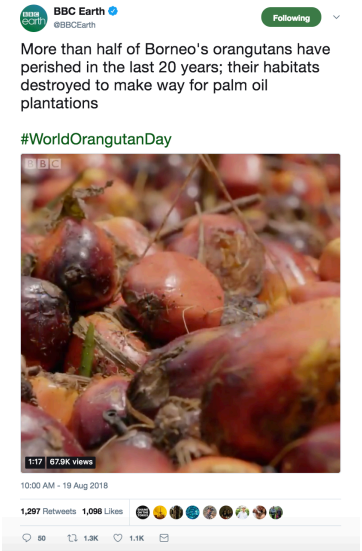
(a) Number 1 most frequently retweeted tweet (retrieved September 25, 2018)

(b) Number 2 most frequently retweeted tweet (retrieved September 25, 2018)



(c) Number 3 most frequently retweeted tweet (retrieved September 25, 2018)

(d) Number 4 most frequently retweeted tweet (retrieved September 25, 2018)



(e) Number 5 most frequently retweeted tweet (retrieved September 25, 2018)

(f) Number 6 most frequently retweeted tweet (retrieved September 25, 2018)

(g) Number 7 most frequently retweeted tweet (retrieved September 25, 2018)

(h) Number 8 most frequently retweeted tweet (retrieved September 25, 2018)

Figure 6.2: Eight out of ten most frequently retweeted tweets

(a) Number 9 most frequently retweeted tweet (retrieved September 25, 2018)

(b) Number 10 most frequently retweeted tweet (retrieved September 25, 2018)

Figure 6.3: Two out of ten most frequently retweeted tweets

6.1. Construct 1: Twitter is an incarnation of the Habermasian public sphere

The public sphere, as defined by Habermas, is “a realm of our social life, in which something approaching public opinion can be formed” (Habermas et al., 1974, 49). This section looks into Twitter as a public sphere, where something approaching public opinion can be formed. Four characteristics are defined for the Habermasian public sphere. These will be elaborated on in the next sections.

1. Social inequalities are bracketed (section 6.1.1)

   Showing the different users participating and especially looking at those with an influential role to see if everyone is seen as equal and no differences are made between different users.

2. Public sphere is a new centre of institutional authority (section 6.1.2)

   Interviewing policymakers to see what they use the data for and if they see the public organizing into a counterweight to the traditional institutional authority. Additionally, tweets are analyzed to see if they contain a call for action.

3. Access is granted to all citizens (section 6.1.3)

   Mapping different groups of users to find out what categories of users are active. Additionally checking if these users participate out of public or strategic interest.

4. Those forming public opinion are presupposed reasonable(section 6.1.4)

   Analyzing tweets to see what argumentation is used, if additional material (such as a video or image is used) and if they are classified as positive or negative.

6.1.1. Social inequalities are bracketed

Bracketing social inequalities means that the public sphere is from and for everyone, not just of and for the privileged class in society. To do so, the network and the influencers are looked into. Showing the network aims to give insight in who is talking to who and who all play a central role in the debate. By showing influencers, the aim is to translate the elite in the traditional public sphere (the spokesmen, the people with the most power) into influencers in the Twitter public sphere. There are many different ways to map the network and to find influencers, but due to limited time and resources, a selection has been made. Because of the activity by automated messages and the so called bots, a separation is made between messages that gain influence through these bots and through actual human users.

The network of actors involved

The first method to find out if social inequalities are bracketed is based on mapping the different actors involved. This method creates insight in what users are active and possibly what status these have. There are 65.417 unique users active in the 103.500 tweets gathered. A deeper analysis of users is based on the so called reply graph. A reply graph shows the replies between the different users. Figure 6.4 is a reply graph and shows how many tweets have been replies between the nodes, with the nodes as users and the lines (edges) the actual replies. The network in figure 6.4 is only part of the total reply network, namely the giant reply network. In network theory, the giant component is "a connected graph that includes a large fraction of all documents in the data set that can be connected (Bollobás and Riordan, 2004)" (Small, 2009, 332). The giant reply network is chosen in this case because of the large, but not too large, fraction connected. The whole reply graph would be too big to draw conclusions from with the naked eye and this giant component included a large fraction of data points.

The giant reply network in figure 6.4 is a network with many replies between users. This graph does not show who is replying to who. The most central nodes are surrounded by a green square. These correspond with the users in table 6.1. This table also splits out if the user is replied to or has replied. Nestle and Kitkat play a very central role but hardly reply to other users (6 and 0 times). RSPOtweets, however, is active in replying and has replied to more users than users have replied to them (117 versus 87). PepsiCo is replied to more than it has replied to (151 versus 69), but can be seen as active. By replying on Twitter, they give their organization a face on Twitter and give other users the idea they are willing to participate in the debate. It is therefore possible to play an important role in the reply network with and without actively participating. To conclude, the network consists of 65.417 unique users from which some have a more influential position than other. This influential position can be gathered without tweeting yourself, like Kitkat en Nestle.

Name	Is replied to	Has replied to
1. PepsiCo	151	69
2. Nestle	143	6
3. RSPOtweets	87	117
4. KitKat	29	0

Table 6.1: The names of the users and the amount the user has replied to, and is replied to, for the users with a green square in figure 6.4

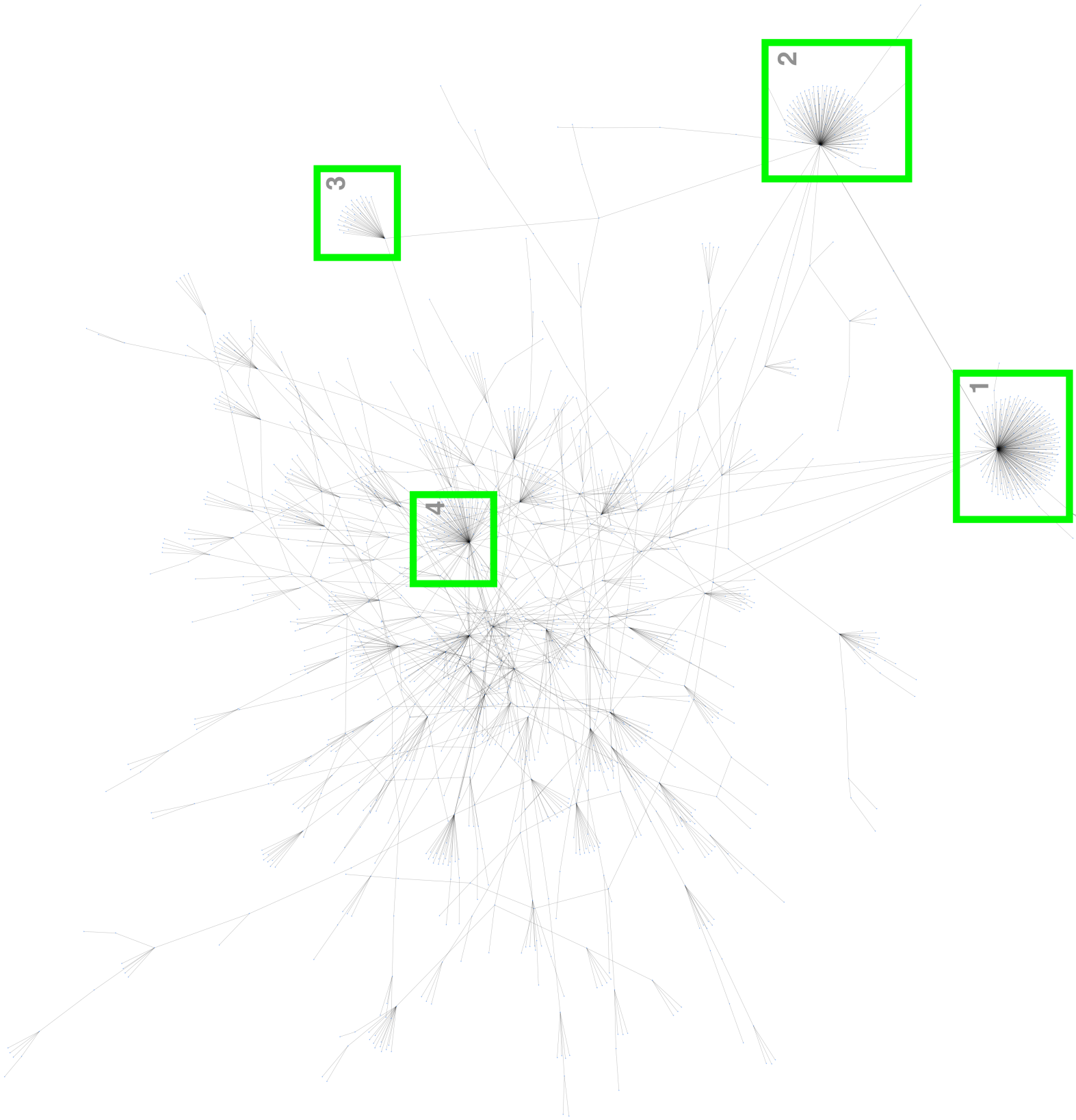


Figure 6.4: Giant reply graph with key actors shown with a green box

The possible influencers on Twitter

The second method to find out if social inequalities are bracketed is based on finding possible influencers. Argued is how "ordinary people -the public- are cast in the role of audience member who for the most part are merely able to watch the events unfolding on this virtual stage of mediated communication" (Bruns and Highfield, 2016, 56). This means that the ordinary people do not have a say and there, it is not a very egalitarian place. Others, in the traditional public sphere the elite, on Twitter influencers, define what happens. It is important to note that in this thesis there is no such thing as 'the' public, but in this case, it means all the ordinary people. To see if these ordinary people are cast in the role of audience member, as described by Bruns and Highfield (2016), the influencers are investigated. The influencers are firstly defined based on three measures, and secondly analyzed based on the table 4.2. The three measures to find influencers are listed below, followed by the top four per category in table 6.2:

1. Central role in reply network (see figure 6.4)
2. Mentioned the most by other users in this database
3. Retweeted the most in this database (see figures 6.2 and 6.3)

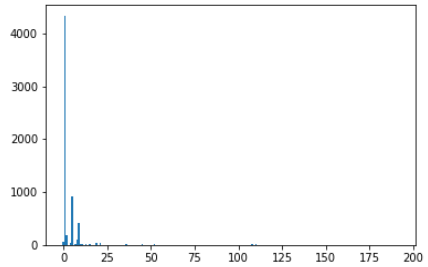
This analysis leads to a top nine users when the doubles are taken out. These are called influencers. See appendix C.1 for a detailed description of the influencers, their tweet- and user level resources are given (based on table 4.2). Not all resources are equally important. Generally, a high in-degree (followers, users who follow you) is seen as more influential than a high out-degree (friends, users who you follow). All users in the table have a higher in-degree than out-degree. When looking at Indofood and KitKat, both with 0 tweets in this database (meaning they have not tweeted about palm oil), it shows that users can even play an influential role in opinion formation *without* participating. The user bdpd_sawit is a bit different but will be elaborated on in the next section. To conclude, no ranking is made within these 9 influencers. They all contribute lots to the diffusion of information and can be described as the glue of Twitter networks.

Central role in reply network	**Mentioned the most**	**Retweeted the most**
PepsiCo	bpdp_sawit	bpdp_sawit
Nestle	RSPOtweets	Greenpeakce UK
RSPOtweets	indofood	Mike Hudema
KitKat	PepsiCo	BBC Earth

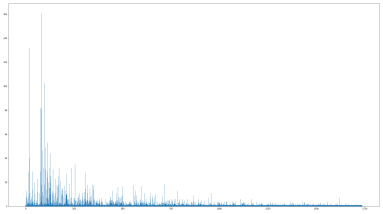
Table 6.2: The users who can be classified an influencers, based on three different categories

Spam and bots

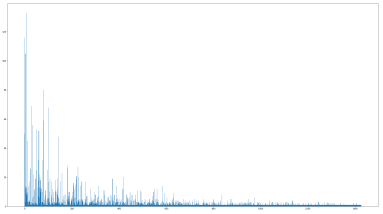
Spams, bots and automated messages are present on Twitter and thus also in the debate on palm oil. A report by Nexgate mentions how 1 in every 200 social media posts is spam (Inuwa-Dutse et al., 2018). Additionally, a more recent study mentions how on Twitter 15% of the active users are bots (Varol et al., 2017). Bots are described in the literature as "automated social media accounts governed by software but disguising as human users" (Kudugunta and Ferrara, 2018, 312). The user Badan Pengelola Dana Perkebunan Kelapa Sawit (bpdp_sawit) appeared out of nowhere based on user mentions and retweets on July 18 and conquered a prominent position in the debate. But, as can be seen in figure 6.5a, only by followers with no other followers. The other three figures in figure 6.5 show other accounts of frequently retweeted users to show what other patterns look like. The emphasis is not so much on 0 but much more spread out.



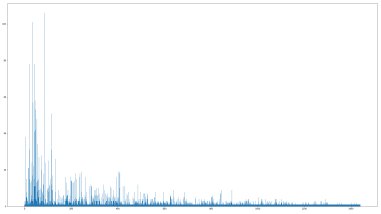
(a) The amount of followers per users that have mentioned bdpd_sawit



(b) The amount of followers per users that have mentioned RSPOtweets



(c) The amount of followers per users that have mentioned indofood



(d) The amount of followers per users that have mentioned PepsiCo

Figure 6.5: The amount of followers per users for the four most frequently mentioned users

In my data, the tweet in figure 6.6 shows up 2409 times. In the official Twitter retweet number (under the tweet) are only 249 retweets. The only possible explanation is that Twitter deleted many of the accounts, including their likes and retweets. A selection of users that have retweeted the tweet in figure 6.6 shows indeed that these accounts do not exist anymore. It is highly unlikely these all deleted themselves. It is more likely that, if Twitter recognized the accounts as spam, they deleted them.

Figure 6.6: Tweet by Badan Pengelola Dana Perkebunan Kelapa Sawit that has been retweeted many times by bots, until Twitter deleted the bots and retweet count went down from 2409 to 249 (retrieved at September 14, 2018)

Additionally, a conspicuousness of spam-posting accounts is their screenname. This name contains many digits and the usernames an corresponding screen names are inconsistent (Inuwa-Dutse et al., 2018). See figure 6.7 for a selection of screennames and usernames for retweeted accounts (that do not exist anymore). The screennames contain letters and digits randomly, and the usernames are inconsistent and all Arabic.

Figure 6.7: Screennames followed by usernames of the users that have retweeted Badan Pengelola Dana Perkebunan Kelapa Sawit. There is a dissimilarity between username and screenname and the screennames contain many digits

Two other tweets that show up much more than the official retweet numbers are:

- President to Sign Regulation on Expanding Mandatory B20 Soon https://t.co/Wt46b3HvuZ #SawitBaik #BiodieselHematDevisa #PalmOil #OilPalm (retweeted 1974 times in my data versus 179 in official data)
- Businesspeople Encourage Decent Work in Palm Oil Sector https://t.co/U1wayD8PBS #SawitBaik #PalmOil (retweeted 1276 times in my data versus 765 in official data. Many of these accounts are temporarily restricted)

These three tweets point to specific articles on the website of Badan Pengelola Dana Perkebunan Kelapa Sawit which has a counter how often articles have been read. The three frequently retweeted messages have been clicked on around 8.000 times, whereas other articles have been clicked on around 500 times. Plotting the total amount of tweets over time against the tweets over time generated by potential bots (mentioning Badan Pengelola Dana Perkebunan Kelapa Sawit) can be seen in the graph in figure 6.8. This graph shows how these retweets are published very gradually and following the flow of all other tweets, instead of being published all at the same time.

Badan Pengelola Dana Perkebunan Kelapa Sawit is the Indonesian palm oil fund (also called the Commodity Pool Operator (CPO) Fund). All taxes on palm oil products and export are gathered by this organization and reinvested in the sector. For 90% this happens through a subsidy for palm oil bio fuel producers. The organization is part of the Indonesian ministry of finance but also representatives of other ministries are present. To conclude, it can be seen that the user Badan Pengelola Dana Perkebunan Kelapa Sawit has gained its influence through fake accounts and therefore, their status of influencer can be argued to be unfairly obtained. There are possibly more spam users and fake accounts in this dataset and Badan Pengelola Dana Perkebunan Kelapa Sawit only shows that spam is happening. It can be concluded that fake accounts are active in the debate on palm oil and try to influence public opinion.

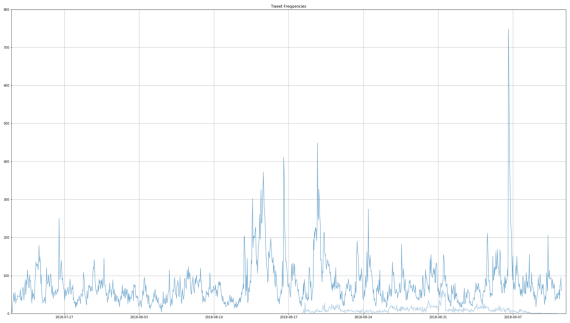


Figure 6.8: The total amount tweets minus the retweets of bdpd_sawit plotted against the time (top line) versus the amount of tweets by retweets of bdpd_sawit (bottom line)

Conclusion characteristic 1: Social inequalities are bracketed

Bracketing social inequalities means that the public sphere is of and for everyone, not just for the privileged ones in society. There are 65.417 users participating in the debate on palm oil. This characteristic has analyzed public status based on the network and the influencers.

Regarding the network of users participating, not all of these users have been investigated in-depth so it is difficult to make definite conclusions if social status is indeed bracketed. What can be concluded is that it seems that the amount of followers does not matter to enter the public sphere of palm oil. Even people with zero followers can enter, but this does not automatically mean they are also heard. All the users with zero followers, mentioning Badan Pengelola Dana Perkebunan Kelapa Sawit, are most probably not heard because their tweets do not show up on any timeline. The network in figure 6.4 shows how there are four key actors. Three out of four are multinationals, namely PepsiCo, Nestle, and Kitkat. The

other actor is RSPOtweets, which has has a very active role and gained its status by replying more than it has been replied to. This is different for Kitkat and Nestle, which have hardly participated in the debate on palm oil.

Influencers are based on their central role in the reply network, mentioned the most by other users in this database and are retweeted the most in this database. Here, fake accounts have had influence in the debate. Twitter has deleted these accounts but since the data collection is real time, the tweets are captured before they were deleted. These have tried to make the Indonesian investment fund Badan Pengelola Dana Perkebunan Kelapa Sawit show up prominently in the debate but Twitter did not let them. Besides the above mentioned multinationals and RSPOtweets, the palm oil company Indofood, the NGO Greenpeace UK, the Greenpeace activist Mike Hudema and the media channel BBC earth have influential roles in the debate of palm oil and therewith influence public opinion.

To conclude, these findings suggest that social inequalities are not completely bracketed. Anyone with internet access, new media literacy and enough cultural capital can *in theory* participate. In practice, however, not much has changed since the 18th century. Argued is how the ordinary public in the coffeehouses of the 18th century did not have a say but how communication was carried out by an elite, the bourgeoisie. This made it, even though Habermas wanted it to be, not a very egalitarian place. The same goes for Twitter. Many influential users are multinationals and some have gained their influential role through spam. Their influential role forms public opinion.

6.1.2. Public sphere is a new centre of institutional authority

A new centre of institutional authority is more than a new place to talk. These places are sovereign in relation to the state. This sovereignty means that where normally the state authority has the monopoly of interpreting information in a way favorable to them, now these public spheres can go against the authority of the state by interpreting information in a different way. To see if public spheres go against the authority of the state, this section analyzes possible new centres of institutional authority in two different ways. Firstly, in a direct way of analyzing a new centre of institutional authority, interviews with policymakers will give insight in if and how social media data is used. The interviews create insight in if they see social media being used to organize an alternative public, which could in turn transform into a new centre of institutional authority. This is called a counterweight to traditional authority and speaks out against certain structures. It is important to realize that there is not 'one policymaker'. Different ministries and municipalities have different policies with regards to social media data, but these interviews show how some look at using this data. Two interviews are conducted: one with a policymaker at the Dutch ministry of social affairs and employment (SZW) and one at the Amsterdam municipality with a social media advisor. These interviews do not link directly to the case study of palm oil because I was not able to find policymakers that make use of public opinion in policy making on palm oil. This is probably due to the European decision making around Europe instead of only Dutch. Both interviews are expanded upon in the next sections and both bodies have very different strategies. SZW does not use social media to measure public opinion but does look at what offline influencers tweet. The Amsterdam municipality uses it, among traditional techniques such as surveys and focus groups, to get insight in what the public is talking about. They both don't see groups online framing a clear counterweight to the institutional authority (them), but that does not mean it does not happen. It might just slide off the radar because they are not analyzing tweets at all or not analyzing the right tweets.

Secondly, in a more indirect way, Twitter can become the centre of institutional authority by calling for action. Many different calls for action are possible, for example users calling upon each other to stop consuming palm oil or on companies to stop producing (products with) palm oil.

Interviews with policymakers on the use of social media data - ministry of social affairs and employment

According to the interview at the Ministry of Social Affairs and Employment, social media data is not used to measure public opinion for multiple reasons. See appendix D for the transcript of the interview. Firstly, they don't believe it gives an accurate view of the average citizen. Besides, the amount of active users in the Netherlands is not enough to give an overall average view. Secondly, if the ministry wants to know something about a certain topic, they organize focus groups for in-depth conversations.

Twitter is used, however, by policymakers to find out what a select group around the Hague shares online. Politicians, journalists and unions are very active and policymakers follow their tweets to see what topics are mostly discussed. They define who to follow based on who is influential in the world outside Twitter, and follow them on Twitter. The relation between the journalists and politicians on Twitter is described as the traditional chicken egg dilemma, a causality that works two ways. Journalist write about topics discussed by politicians and politicians discuss what journalists write about. Based on this interview, Twitter cannot be seen as a new centre of institutional authority since the public is not taken into account by policymakers. This is a prerequisite for realizing them as a counterweight. Instead, they are merely seen as audience members. The different policies for different ministries have to be emphasized though, so maybe it is different for others.

Interviews with policymakers on the use of social media data - Amsterdam municipality

According to the interview at the Amsterdam municipality, social media data is used for multiple reasons. See appendix D for the transcript of the interview. The municipality uses the tool Coosto to monitor everything that happens on social media. Coosto gives real time updates based on certain keywords from Twitter, Facebook and Instagram in an easy-to-use interface. Every week they evaluate the data to see if their communication strategy is fulfilling

their key performance indicators. There is no direct effect on policy making since that takes a long time and is often already determined, but there are certainly possibilities to use the data from social media to get instant feedback on certain policy decisions. The municipality is working towards a more responsive system where they actively act based on the input from social media and other data sources.

Besides the advantages, they also see disadvantages of Twitter by the increasing voice of extreme right groups and censorship. Firstly, they describe the increase of alt-right activists. Research indicates how an army of Russian trolls increase the publicity around these right activists (van der Noordaa and van de Ven, 2018). Secondly, they recognize censorship on politically sensitive discussions such as 'zwarte piet' (black Pete), where the image is hidden first and you get a warning that the image may contain sensitive content. Twitter, in this case, decides with an (nontransparent) algorithm what is sensitive. Based on the previous two sections can be concluded that policymakers have different strategies in using social media data. There is still a lot to gain in the way they conducted analyses but in their current state, it cannot be argued they see a new centre of institutional authority emerging in a direct way.

Data to find calls for action

This section analyzes an indirect possibility to form a new centre of institutional authority through analyzing tweets. Calls for action are used as a measure, since a call for action can organize a public.

New centers of institutional authority can be possible on Twitter. Tweets are analyzed to find calls for actions as an indicator for forming a new centre. Calls for action can happen in multiple categories. A call for action can be very direct, semi direct or indirect. Examples of direct calls are the ability to buy a piece of land to replant forest or to sign a petition. An example of a semi direct call is 'to tell big brands to drop dirty palm oil' (as made by Greenpeace, see figure 6.2) since not many people will go out and actually tell the big brands. Lastly, a shout out for awareness can be seen as an indirect call for action by. A well-known example is the #MeToo. No direct actions are attached, but it raises awareness and can indirectly cause action.

The top ten retweets together make up 12% of the total tweets. They are analyzed to see if they contain a call for action and if so, in what category. See table 6.3 for the user, tweet, number of retweets ans possible call for action. As can be seen in the table, four tweets contain a direct call for actions. Others are less direct, for example the tweet by Mike Hudema: "One palm oil company has destroyed an area of forest half the size of Paris. PepsiCo, Mars, Nestle, Unilever - companies that have promised not to buy palm oil from forest destroyers - are all supporting the destruction. That needs to stop. #deforestation #PalmOil". This tweet does not contain a direct call for action but does express a worry and an expression of how deforestation needs to stop.

The tweets that do call for action do it in two ways: by signing a petition to tell multinationals to stop 'dirty' palm oil and by making you buy part of the land to replant forest. It falls outside the scope of this thesis to check what the effects are since Habermas' theory is based on communication and not direct action. Section 4.2.3 looks further into the retweet function and therewith the creation of an ad hoc public. If the ad hoc public created is large enough, it can eventually become a new centre of institutional authority.

The words boycott, stop and petition can also be argued to be direct calls for action (besides the above mentioned calls to sign a petition and to donate money). Boycott appears almost 1000 times, stop 2123 times and petition 800 times, which is respectively 1, 2 and 1% of the times. Due to limited time and resources, not all tweets are checked for more calls for action. To conclude, calls for action are made in the analyzed tweets in different categories. These urge people to do something in a direct, semi-direct on indirect way. These calls for action can influence public opinion which in turn can lead to behavior choosing to make it more than only spoken word, but actions are not measured in this thesis.

User	Tweet	Number of retweets	Call for action?
Badan Pengelola Dana Perkebunan Kelapa Sawit	Tweet 1 (figure 6.2a)	2409	-
Badan Pengelola Dana Perkebunan Kelapa Sawit	Tweet 2 (figure 6.2b)	1947	-
Greenpeace UK	Tweet 3 (figure 6.2c)	1589	Petition for big brands to drop dirty palm oil
Badan Pengelola Dana Perkebunan Kelapa Sawit	Tweet 4 (figure 6.2d)	1276	-
BBC Earth	Tweet 5 (figure 6.2e)	1176	-
Mike Hudema	Tweet 6 (figure 6.2f)	808	-
Stephen Fry	Tweet 7 (figure 6.2g)	764	Buy the land to replant forest
Peter Egan	Tweet 8 (figure 6.2h)	745	Buy the land to replant forest
Daniel Schneider	Tweet 9 (figure 6.3a)	703	-
Greenpeace	Tweet 10 (figure 6.3b)	674	Petition for big brands to drop dirty palm oil

Table 6.3: Top ten most frequently retweeted tweets and possible call to action (generated September 12, 2018)

Conclusion characteristic 2: Public sphere is a new centre of institutional authority

Public sphere is a new centre of institutional authority means it is more than a new place to talk. It is a place sovereign in relation to the state. This means that where normally the state authority has the monopoly of interpreting information in a way favorable to them, here these public spheres can go against the authority of the state by interpreting information in a different way. Social media can be used to organize a completely different public, like the well known #MeToo-discussion. The #MeToo-discussion discussion can be argued to have influenced public opinion on sexual harassment in the work place. In only becomes a new centre of institutional authority when consequences are attached to the organized public but influencing public opinion is certainly a first step.

Interviews are conducted at the ministry of social affairs and employment (SZW) and at the Amsterdam municipality. Both have very different polices: SZW uses social media mostly to see what politicians, journalists and unions post. They clearly do not represent the whole public but can represent the activated public opinion, who can anticipate trends in wider public opinion (Andrea and Fedra, 2016). So even though policymakers do not take public opinion data directly from the whole public, they are aware of influencers and the activated public opinion. The Amsterdam municipality uses social media to monitor what is going on in the city. The monitoring does not directly influence policy making but indirectly and in the long term, it might. The data shows different calls for action such as calls to sign a petition and to donate money to buy a piece of the rain forest. It is not investigated what the actual effect of these calls is. The debate on palm oil has in the monitored time (from July 23 - September 1) never formed a large enough public to possible be a new centre of institutional authority, like the possibility of #MeToo.

These findings suggest that, in the case of palm oil, no new centre of institutional authority has emerged. Different users have been in the position to interpret information in a different way than as given by the state, but the interpretations have not been all in the same direction. The different calls for action present have not lead to a new centre of institutional authority nor is Twitter seen like that by the policymakers interviewed.

6.1.3. Access is granted to all citizens

When access is granted to all citizens, anyone with willingness to join can enter. Additionally, the argument brought in should be free of ideology and power, with the common interest of good opinion formation as the only interest (Kuitenbrouwer, 2018). The opinions expressed on social media, and especially Twitter, do not cover the whole public debate but might can be seen as activated public opinion. This activated public opinion can anticipate trends in public opinion (Andrea and Fedra, 2016). To see if access is granted to all citizens (all meaning all on Twitter, so leaving out those with no internet or no Twitter account), different actors are grouped and mapped.

Mapping different groups of actors

This section builds further on the actors identified in section 6.1.1. It is important to see out of what interest the users are participating. Is the interest based on ideology of power and thus strategic, or is the main interest the common interest of good opinion formation? The first step is mapping who is all participating. The top twenty of users who have been retweeted the most, are mentioned by others the most or have tweeted the most are analyzed. After taken doubles out, there are 41 users left. See appendix C for the 41 users and their categories. Figure 6.9 shows the division into six categories, which are not mutually exclusive nor collective exhaustive but based on the data:

1. Business (for profit, like Nestle)
2. Government (like the Malaysian Ministry of Primary Industries)
3. Individuals
4. Media
5. NGO (like Greenpeace)
6. nonprofit (like the RSPO)

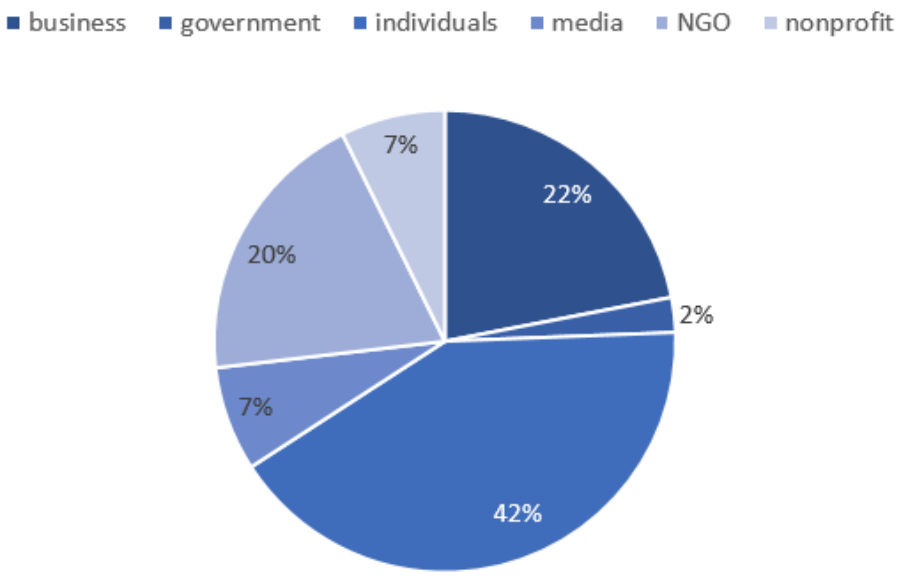


Figure 6.9: Division based on the 20 most mentioned users, 20 most retweeted users in the database and 20 users of the most frequent tweets in database (leading to n=42 when taken duplicates out)

It can be argued that businesses- with 22% of the selected group- not necessarily participate out of the common interest of good opinion formation. NGOs -with 20% of the selected group- often participate out of an ideology. See figure 6.2c, where the NGO Greenpeace UK argues against multinationals cutting down forest. The arguments of NGOs can be argued to be out of ideology to save the forest. As Greenpeace describes on their website, they “exists because this fragile earth deserves a voice” (Greenpeace, 2018a). This is even more clearly an ideology and an ideology is, according to Habermas, not a valid base to participate neither.

Beside what their motivation to participate, the high percentage of individuals -42% of the selected group- can be interpreted to grant access to all. Not all these individuals are interest free, however. Indra Nooyi (until recently the CEO of PepsiCo) and Mike Hudema (activist at Greenpeace Canada) are examples individuals who might post in the interest of their organization, like Indra Nooyi in this tweet from July 11:

> "My mom once asked me what I would do to change the world. My answer's been the same since becoming CEO @PepsiCo: Lead a company that's a force for good. Our sustainability report tracks the good we're doing around the world thanks to our amazing associates"

This tweet can be classified as a strategic tweet in the interest of PepsiCo, to show how good they do around the world. To conclude, it is very hard to make a distinction between tweets in the interest of all and in the interest of good opinion formation, or tweets in the interest of an ideology or out of power. In conclusion, access is granted to all instead of only to all citizens. Many participating in the debate are businesses and NGOs, who's interest can be argued to be either ideological, out of power or strategic, which all influence public opinion.

Conclusion characteristic 3: access is granted to all citizens

Access is granted to all means to all with reason and willingness to join. The argument brought in should be free of ideology and power, with the common interest of good opinion formation as the only interest (Kuitenbrouwer, 2018). It is difficult to define if all those with willingness to participate can actually participate, because we only see those who are participating and not those who are not. We know those without internet access, enough cultural capital and new media literacy are excluded.

The users participating in the debate on palm oil on Twitter, in this research, can be divided into six categories: business, government, individuals, media, NGO and nonprofit. Individuals have the largest share and it can be argued that their argument is private without being strategic. A multinational like PepsiCo is less likely to have non-strategic arguments. Many individuals, however, post in name of an organization. The previous page highlights the example of the individual account of the CEO of PepsiCo, making it very difficult to make a definite statement about the interest of users.

When comparing these findings back to the coffeehouse, anyone with internet access, enough cultural capital and new media literacy can enter. Maybe that is not very different from the 18th century, where access is meant to be granted to all but excluded all women (also half the world population). But times have changed drastically and global capitalism can be argued to have shifted those participating in public debate from individuals to multinationals. To conclude, these findings suggest that multinationals, NGOs and individuals who are active either at multinationals or NGOs play a significant role in the debate on palm oil. Their arguments are not always free of ideology and power and also not have the common interest of good opinion formation.

6.1.4. Those forming public opinion are presupposed reasonable

Habermas has argued for the best argument instead of the argument made by the one with the highest status as part of the rational-critical debate. Rational is there where voluntas (desires) transform into ratio (reason). People are presupposed reasonable. Reasonable means having reason, and reason is a statement offered in explanation or justification (Merriam-Webster, 2018g). These statements should, according to Habermas, not be based on an ideology or power (Kuitenbrouwer, 2018). Critical stands for the possibility to express critical judgment. This section looks into the argumentation of the most frequented retweeted tweets, to see if they can be seen as reasonable. Figures 6.2 and 6.3 show the most retweeted tweets. Since the three tweets by Badan Pengelola Dana Perkebunan Kelapa Sawit are only high in the ranking due to bots and are already deleted by Twitter, they are not analyzed in-depth. The other seven are analyzed. Table 6.4 shows the arguments in a tweet, if a video, image or link is used and additionally, an automated package is used for the polarity of a tweet, which can also be seen in the table. Polarity is showing if a tweet is negative (-1) or positive (1). See appendix A for an explanation on how the polarity of tweets is defined.

Observations based on the data

The argumentation of the most frequented retweeted tweets is checked for, to see their reason (the explanation and justification used). The results can be seen in table 6.4. What can be seen from this table is that all tweets are about deforestation and all of them use an image or video. Six out of seven tweets are classified as more positive than negative. The tweet by Mike Hudema is classified as slightly negative. It is indeed not framed in a positive way. The tweets by Peter Egan and Stephen Fry show an clear contrast. Where their message is the same (buy land) Peter Egan tweets very positively and optimistically while Stephen Fry is a bit less optimistic by mentioning how worrying the situation is. This leads to different polarity scores (0.90 versus 0.00). The positive polarities show how messages that get retweeted often do not only contain complains and worries, but by positive turns and (in)direct calls for action.

Tweet	**Argument**	**Image/ Video/ link**	**Polarity**	**Followers**
Tweet 3 (figure 6.2c)	Deforestation, endangered orangutans	Video	0.10	17.993
Tweet 5 (figure 6.2e)	Deforestation, endangered orangutans	Video	0.11	467.167
Tweet 6 (figure 6.2f)	Deforestation	Image	-0.17	78.759
Tweet 7 (figure 6.2g)	Deforestation, endangered animals	Image	0.00	12.748.457
Tweet 8 (figure 6.2h)	Deforestation, endangered animals	Image	0.90	48.131
Tweet 9 (figure 6.3a)	Deforestation, endangered animals	Image	0.029	84.566
Tweet 10 (figure 6.3b)	Deforestation, endangered orangutans	Video	0.85	1.700.871

Table 6.4: Top 7 retweets left with their argument, media used and polarity

Conclusion characteristic 4: those forming public opinion are presupposed reasonable

The 'best argument' and not the argument made by the one with the highest status plays an important role in the Habermasian public sphere: may the best argument win. Those with an argument that translated voluntas (desires) into ratios (reason) can be argued to be reasonable. Additionally, these arguments should not be based on an ideology or on power (Kuitenbrouwer, 2018). It is very difficult to define what the best argument is because it is very subjective to define what is best. It is also difficult to define the highest status.

Additionally, the area between voluntas and ratio is grey. All tweets desire to stop deforestation and use different reasons (endangered animals or specifically orangutans). Also, all tweets express judgment about the current situation in the rain forest and can therefore be called critical. When assuming a correlation between followers and status, table 6.4 shows the number of followers to indicate that none of these users has less than 17.993 followers.

These finding suggest that is very difficult to decide if the users on Twitter are reasonable. Reason -a statement for explanation or justification- is selectively used by users on Twitter. I can therefore not conclude that the debate is rational. The debate can be characterized as critical, the second part of the rational-critical debate, since many judgments are expressed.

6.2. Construct 2: Twitter causes a fragmentation of the public

Argued is how social media can fragmentize political discourse (Barber, 1998; Papacharissi, 2002). The term public sphericules is used to describe the small social fragments, moving around thematic debates, with the same characteristics as public spheres. In other words, public sphericules are used to describe smaller public spheres in which something approach public opinion on a specific topic is formed. This construct consists of two characteristics, to be elaborated on in the next sections.

1. The public on Twitter is fragmented (section 6.2.1)

   Analyzing the hashtags and word frequencies to see what main topics are and if these are fragmented.

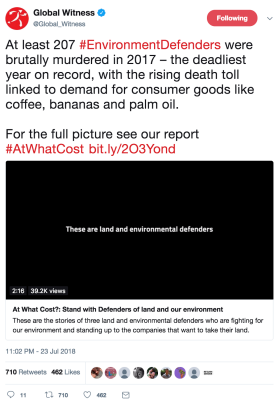
2. Twitter causes this fragmentation (section 6.2.2)

   Analyzing tweets to see if they contain images, snippets of information and interpreted data, which can all be argued to add to a fragmentation of the public.

6.2.1. The public on Twitter is fragmented

Different public sphericules have different topics at their centre. An important realization to keep in mind is how different sphericules fragmentize the public sphere. The public sphere is, in brief, a transcending space as an abstract forum for dialogue and public opinion. The public sphere can be argued to consist of many sphericules. Multiple checks are conducted. The first check is based on a count of the most frequent hashtags. The use of hashtags can be argued to give insights in what topics are being discussed. The people discussing these form the public around the hashtags. The public sphericule is the space where that specific publics forms something approaching public opinion on that specific hashtag.

See table 6.5 for an overview and figure 5.6 for a visualization in a word cloud. As can be seen, #palmoil is with a frequency of 18.395 by far the most frequent hashtag in the database. The next two, #sawitbaik and #biodieselhematdevisa are caused by the bots, who have been deleted from Twitter now. So even though they pop up high in the ranking, they have achieved their position in an unfair way. The hashtags contain a lot of words directly relating to deforestation (orangutans, orangutan, envirionmentaldefenders, deforestation, worldorangutanday, elephants, rainforest, orangnutanday, wildlife, tigers, extinct). These are 11 out of 20 hahstags (and when excluding sawitbaik and biodieselhematdevisa even 11 out of 17). The hashtags are in line with the most frequent retweets, where all seven are about deforestation. The different animals (orangutan, elephant and tigers) can be argued to be different sphericules when not mentioned in the same tweet. Additionally, the hashtag envirionmentdefenders is based on a campaign by Global Witness and can be seen as a sphericule in itself. See figure 6.10a for the tweet.



(a) Tweet by Global Witness (retrieved August 30, 2018)

(b) Tweet containing Indofood, PepsiCo, Nestle and RSPOtweets (retrieved September 17, 2018)

Figure 6.10: Tweets with topics that can be argued to be sphericules

Word frequency count

A second check is based on the word frequency count. When different key words are frequently used, this can point towards a fragmented public sphere. The fragmented public sphere, in turn, makes up a public talking about many different topics. Research shows how on average only 20% of the tweets contains a hashtag (Orellana-Rodriguez and Keane, 2018). In this dataset, 37.5% of the tweets contain a hashtag. This is almost double the average, but still less than half the tweets. Therefore, the word frequency is also listed in table 6.5 and visualized in a wordcloud in figure 5.5. The most frequent words are oil and palm, with over 46.500 mentions. bpdp_sawit (for user Badan Pengelola Dana Perkebunan Kelapa Sawit) and #sawitbaik have not fairly achieved their position as an influencers so are excluded from further analysis. Besides, three other @ to address another user are included. Indofood, PepsiCo and RPSOtweets are in 2000 tweets mentioned together. The message is similar to the tweet in figure 6.10b. Even though these words and names do not directly include deforestation, the messages behind it do. So also here deforestation is an (indirect) frequently discussed topic.

Hashtag	**Frequency**	**Word**	**Frequency**
palmoil	18.395	oil	46.615
sawitbaik	6.535	palm	46.576
biodieselhematdevisa	1.964	#palmoil	18.418
orangutans	1.786	destroyed	7.923
plantation	1.461	aceite	7.765
orangutan	1.221	palma	7.656
environmentdefenders	1.176	orangutans	6.860
deforestation	1.065	@bpdp_sawit	6.715
worldorangutanday	952	#sawitbaik	6.535
elephants	808	way	5.712
indonesia	722	@rspotweets	5.476
rainforest	651	forest	4.944
orangutanday	642	indonesia	4.828
wildlife	585	@indofood	4.544
rspo	584	make	4.418
tigers	519	years	4.372
extinct	486	@pepsico	4.260
malaysia	483	half	4.125
palmolie	435	company	3.731
rt	421	plantations	3.730

Table 6.5: Most frequent hashtags- and words with their frequencies

Topic Modelling

Topic modelling is used additionally to word- and hashtag count. Word- and hashtag counts were used to get a general idea of the different topics discussed and the publics forming the sphericules. To see if no topics are missed, the technique of topic modelling is used. Different tweets are grouped based on different topics to get a deeper inside beyond the most frequent topics. Figure 6.11 shows the result on a scatter plot. Simply put, one dot means the exact same tweet (in other words, retweets). A cluster of dots means similar wording but not exactly the same. Many of the dots and clusters identified are as discussed above. The most remarkable additional insight is 'Africa', which is mentioned 3.292 times. See appendix A for information on the topic visualization technique used.

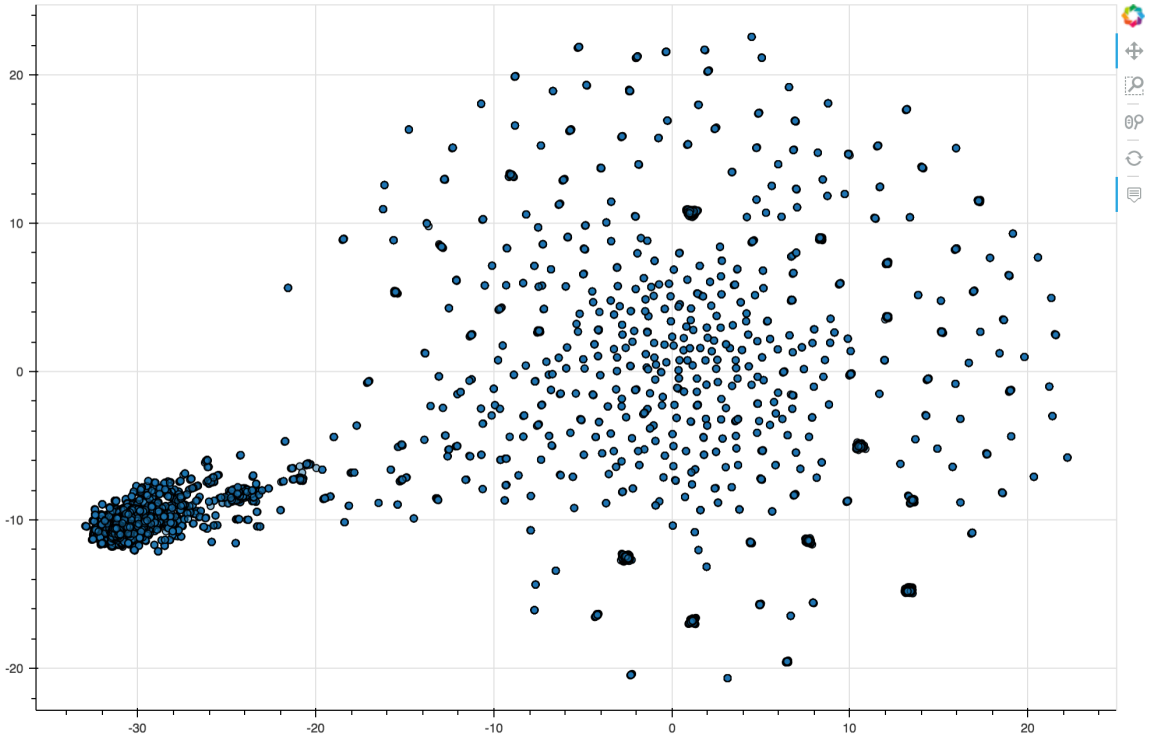


Figure 6.11: Different topics by themes on a scatter plot

Overlap between the different sphericules

Fragmentation of the public sphere does not have to lead to mutually exclusive sphericules. Separate fragments can be overlapping. The previous analysis has lead to main sphericules, the largest fragments in the public sphere. These are shown in table 6.6. An example of possible overlap is the orangutan and Indonesia. The orangutan can live in Indonesia and there is the overlap.

The sphericules in 6.6 are organized in three main categories: deforestation, places and organizations. Additionally, these are organized in the large sphericules before specializing into smaller sphericules. Organizations can also be seen as influencers in sections 6.1.1, but as the tweets show, many topics circle around certain organizations and are therefore seen as separate sphericules.

Categories	**Large sphericules**	**Smaller sphericules**
Deforestation	Effects on wildlife	Orangutan
		Elephant
		Tiger
	Effects on people	Murder
		Exploitation of its workers
Geographical location	Countries	Indonesia
		Malaysia
	Continents	Africa
Organisations	nonprofit	RSPO
		Greenpeace
	for-profit	Indofood
		PepsiCo
		Nestle

Table 6.6: Categories and sphericules: the last column, smaller sphericules, are explicitly mentioned on Twitter

Figure 6.12 shows how often topics are mentioned in combination (only the most specific level of topics are taken, since these are the ones mostly explicitly mentioned). All the corners of the figure mention the topics, just like the colored lines with shapes. To clarify, when looking at Nestle on the top left corner, we see an orange line coming up high. The legend tells us this line stands for PepsiCo, meaning that the combination of Nestle and PepsiCo has appeared approximately 4.000 times. RSPO (the purple line) has appeared frequently in combination with Indofood, PepsiCo and Nestle (see figure 6.10b for the message in a tweet). Orangutan (red line) has appeared clearly in relation with Indonesia, the tiger, Greenpeace, PepsiCo and RSPO.

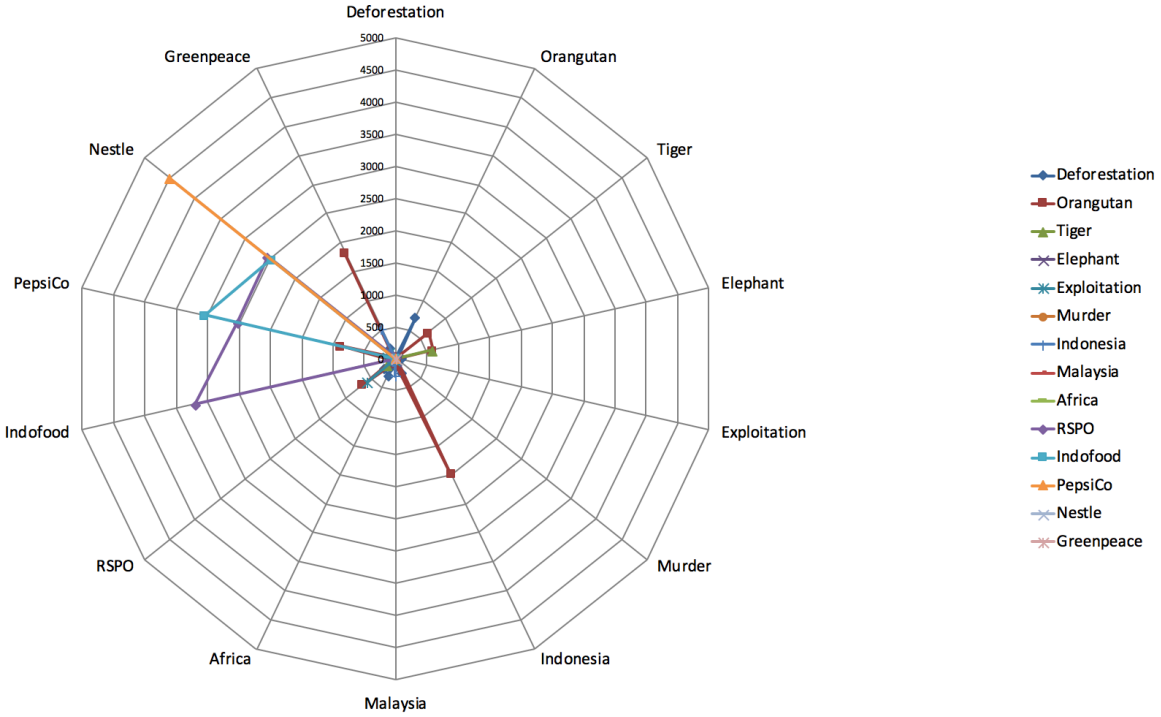


Figure 6.12: Coherence between the different topics

Conclusion characteristic 1: the public on Twitter is fragmented

Argued is how discourse on social media can be fragmentized and the term used for these fragments are sphericules. A fragmented discourse leads to a fragmented public. The sphericules have the same above mentioned characteristics of the public sphere but are smaller. Word and hashtag frequency counts show a clear general direction of public opinion expressed on Twitter. Deforestation is the main topic discussed, inside which are different public sphericules such as the orangutan, exploitation of plantation workers, Indonesia and RSPO. These topics are often overlapping. Even though they seem independent, when mapping them in relation to each other it shows how they often overlap. The overlap can be clarified even more by looking at the implicit topics that are not mentioned. Deforestation does not show clearly on figure 6.12 because it is not always mentioned explicitly, but is is often there implicitly. Besides the fact that sphericules are smaller than public spheres, they are argued to have the same characteristics. Checking all the sphericules for the four characteristics of construct 1 falls outside the scope of this thesis, but it can be argued that because they all occur on Twitter on the same main theme (palm oil), the analysis of characteristics is the same as in section 6.1.

These findings suggest that, depending on what scale one looks, the public can be seen as fragmented. There are different sphericules but these cannot be seen as completely separate and independent. The sphericules with a public large enough to be taken into account in figure 6.12 show that the different sphericules are not very different and, as can be argued in line with Habermas, the overlap can counter fragmentation. It can be concluded the public

on Twitter in the case of palm oil on the highest level is not very fragmented but on the lower levels, different overlapping fragments are present.

6.2.2. Twitter causes this fragmentation

To see if the possible fragmentation from the previous section is actually caused by Twitter, the top ten retweets are analyzed based on three characteristics by Barber (1998). Since three out of these ten retweets are not supposed to be there (due to the bots), only seven are analyzed. See table 6.7 for the analysis. They are analyzed based on the following characteristics (as explained in section 6.2).

1. Their bias towards images over text
2. Their partiality to raw data rather than informed knowledge
3. Their inclination to audience-segmentation rather than to a single, integrated community of users/viewers

Their bias towards images over text

The first analysis is based on the bias towards images over text since an image might be worth more than a thousand words. As can be seen in the column 'image' in table 6.7, four out of seven have a video and three out of seven an image. These images can be argued to be in line with the bias towards images over text. Argued is how pictures are not conducive to thinking (Barber, 1998), let alone videos. They all sketch a one-sided view, as engaging as possible, to grab one's attention in the little time there is. The short movies in four out of seven tweets are spoon-feeding information. There is no need (or possibility) to fill in blanks, one of the characteristics of McLuhan's hot media (Chandler and Munday, 2011).

Their partiality to raw data rather than informed knowledge

Raw data versus informed knowledge does not seem to be true in this case. All the data is very interpreted. The only data that can be argued to be raw is based on the information on the Tarsier (tweet 6). It can be argued that the story comes across a lot stronger with interpreted data than with raw data. Additionally, one might need more characters than 280 or a 1:30 minute video to make a point with raw data. Barber's warning of a data overload does not happen and can be argued to be prevented by a specific interpretation of the data. A point of discussion is if interpreted knowledge is informed knowledge. Informed is described as "based on possession of information" (Merriam-Webster, 2018d). Interpreted knowledge is certainly based on possession of information, but it is not certain on what information.

Their inclination to audience-segmentation rather than to a single, integrated community of users/viewers

All tweets tell parts of the story, snippets of information. Palm oil is, as discussed in section 2.6, a wicked problem and it therefore is impossible to shed lights on all sides in 280 characters. As I've learned from the interviews with policymakers on palm oil, a boycott on palm oil will not have the desired effect. If we do that as the Netherlands, we give away all market share and therewith our market power. Not all these tweets argue for a complete boycott, but it is not very clear what *exactly* they want. Greenpeace (tweet 1 and 7) argues to "drop dirty palm oil" but does not explain what dirty palm oil is. By looking a bit deeper into the website of Greenpeace, they explain how dirty palm oil is "palm oil from suppliers that trash rain forests" (Greenpeace, 2018b). This definition is still ambiguous. Besides, it is not very probable that everyone who sees the tweets goes out and investigates what exactly they mean with dirty palm oil. This ambiguity might be created on purpose, to leave space for the individual to come up with an interpretation that suits their ideas. If this is the case, there seems to be a lot of common ground on the surface (i.e., many retweets to drop dirty palm oil) but when going under the surface, it is not clear if the all the individuals making up the public have the same opinion on what exactly is meant with dirty palm oil.

Conclusion characteristic 2: Twitter causes this fragmentation

Multiple analyses are conducted to see if the argued fragmentation of the public sphere is caused by Twitter. There are multiple (overlapping) public sphericules visible in the Twitter

sphere on palm oil. To see if the emergence of the different public sphericules is caused by Twitter, most frequent retweets have been analyzed by three characteristics from Barber (1998). This is solely a first step and no ultimate conclusion can be done if fragmentation is caused by Twitter.

These findings suggest, however, that there is a bias towards images and videos over text. There is not much raw data but more interpreted data, which is not the same as informed knowledge. Snippets of ambiguous information are given. Not-specifying certain words, like dirty palm oil, leaves them open to the interpretation of the audience. The combination of images, interpreted data and snippets of ambiguous information hint towards McLuhan's hot medium, where the information is spoon-fed and one believes whatever they see in front of them. There is no reason to go around and try to fill the gaps, since what is presented seems to be the whole story. It cannot be stated that Twitter causes fragmentation but certain features do add to the quick attention grabbing of users into certain directions and therewith influence public opinion, possibly into different fragments.

When looking back to section 6.1.4, people in the public sphere and in public sphericules are presupposed reasonable. This means they use reason: statements offered for explanation and justification. The findings in this section suggest that reason is loosely interpreted. Arguments made are backed up with interpreted data and snippets of ambiguous information. They do give explanations and justification, but only very narrow and in the direction favorable to the user. These findings makes it not certain that the users are reasonable.

Different tweets are analyzed based on these different features. These features can be combined with the theory of McLuhan on hot and cool media (see section 4.1.2). Hot media are high in sensory data and ask for less sensory involvement. They are 'spoon-feeding' the content. Cool media need more sensory involvement and participation is key. Tweets which contain multiple features are pointing towards hot media, since only a small community of users posts most of the tweets, whereas other are taking information in like warm cookies.

Tweet	Image	Raw Data	Part of story
Tweet 3 (figure 6.2c)	Movie	The data is interpreted, even though the statement of “the 25 orangutans we loose every day” can be seen as raw.	The movie of 1:30 minute is merely meant as an illustration to rise awareness and does not aim to tell the whole story.
Tweet 5 (figure 6.2e)	Movie	The data is interpreted. “Over half Borneo’s orangutans have perished since 1999” and “less than 10 percent of companies in Borneo produce sustainable palm oil.”	The movie of 1:21 minute is meant to tell the story of palm oil in Borneo. It sheds light on the deforestation and consequences for the Orangutan, but not on any other consequences.
Tweet 6 (figure 6.2f)	Photo	The data is interpreted. “One palm oil company has destroyed an area of forest half the size of Paris”.	The tweet is meant to rise awareness. Little information is given, however. Not about which company and where the forest is destroyed.
Tweet 7 (figure 6.2g)	Movie	The data is interpreted. “What is happening to the Indonesian jungle is deeply worrying”.	The movie of 1:15 minute is meant to rise awareness and as a call for action to help buying back a plantation. Not much information is given about what they mean with ’deeply worrying’.
Tweet 8 (figure 6.2h)	Photo	The data is interpreted. “Buy the land and replant forest to create a safe haven for orangutans, tigers and elephants”.	The tweet is meant to rise awareness and as a call for action to help buying back a plantation. Not much information is given about why the animals need the ’safe haven’.
Tweet 9 (figure 6.3a)	Photo	The data is raw and interpreted. Raw is the information about the tarsier (the first part of the tweet) and interpreted is the data about the endangered species.	The tweet is meant to tell the story of palm oil. It sheds light on the habitat destruction and shows the example of a tarsier.
Tweet 10 (figure 6.3b)	Movie	The data is interpreted, even though the statement of “the 25 orangutans we loose every day” can be seen as raw.	The movie of 1:30 minute is merely meant as an illustration to rise awareness and does not aim to tell the whole story.

Table 6.7: Most frequent retweets analyzed based on three characteristics by Barber (1998)

6.3. Construct 3: Twitter creates space for issue publics

We live in a time with access to lots of information and lots of media fighting for our scarce attention. Due to the many distractions we have shorter attention spans which give rise to different shorter-lived topics entering our mind. These can be called issue publics. Issue publics have three elements: they emerge, exist for varying duration and than eventually dissolve. This section consists of two characteristics which are elaborated on in the next sections.

1. Twitter functionalities (retweets and hashtags) create space for issue publics (section 6.3.1)

   Analyzing different peaks (when many tweets were posted in a small time slot) to see if these are caused by retweets and hashtags.

2. Certain issues push each other away in a time where our attention is scarce (section 6.3.2)

   Analyzing the peaks in more depth to see if certain topics push other topics away.

Both characteristics are based on the tweets plotted against the time. See figure 6.13. It is important to emphasize the main difference between public sphericules and issue publics. Issue publics are more dynamic and shorter-lived. There are also similarities: both have the characteristics of a public sphere and both take place around a certain topic.

6.3.1. Twitter functionalities (retweets ans hashtags) create space for issue publics

This section firstly identifies issue publics and secondly identifies their causes. Certain issues are identified that cause a public to talk about them. Figure 6.13 gives an impression where these so-called issue publics might occur. The peaks, surrounded by a green square, are identified and investigated what retweets and/or hashtags cause these peaks. See table 6.8 for an overview of the three peaks with causes, existing time, disappearance and if they can be classified as issue publics. Two out of three peaks have a clear existence, whereas the third one is less clear.

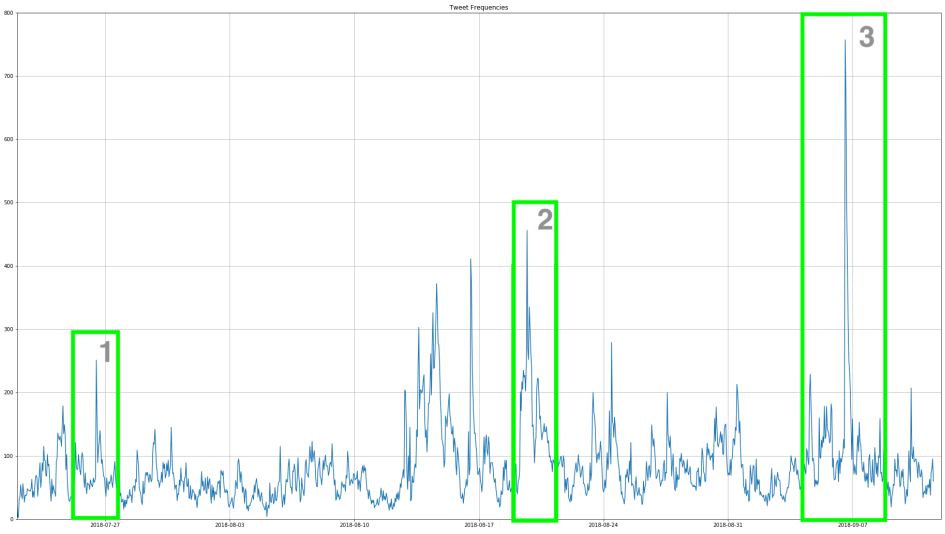
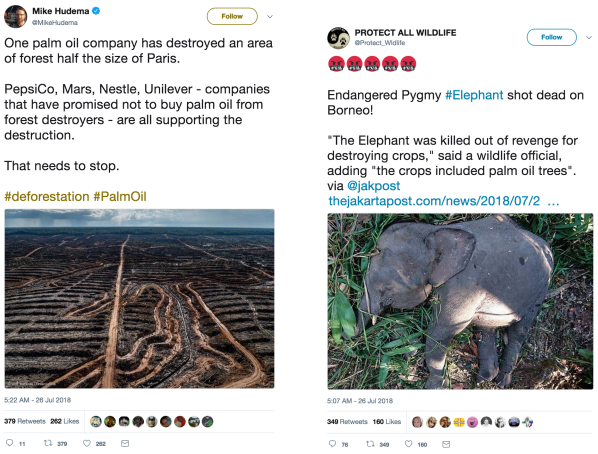


Figure 6.13: Time series of all the tweets from July 23 - September 11 with 3 peaks highlighted in a green square

The first peak is caused by two tweets that have been retweeted frequently. One tweet is from the Canadian Greenpeace activist Mike Hudema (figure 6.14a). The second tweet is is by an organization called Protect All Wildlife, which present themselves as an advocate for wildlife (figure 6.14b). The popularity of the tweets have a clear starting and ending point (as can be seen in the peaks) and so do their publics. The second peak is caused by the issue of world orangutanday, celebrated annually on August 19. BBC Earth has been active in tweeting about this day and also hashtags in different languages like Spanish were popular. After approximately 30 hours, the issue passed and the public fell apart. The last peak is not caused by a retweet or a hashtag. The cause is unclear. Table 6.8 gives possible explanation such as the hyperactive spam user in figure 6.15a, but no single explanation justifies this peak.

There are more peaks visible in figure 6.13 but these three peaks can be argued to be representative for the different known and unknown causes and elements of issue publics. See appendix F for a more elaborate investigation on the possible causes for the last peak. What can be concluded from these three peaks is how retweets and hashtags *create space* for issue publics, but also other (un)known reasons cause these.



(a) Tweet by Mike Hudema (retrieved September 18, 2018)

(b) Tweet by Protect All Wildlife (retrieved September 18, 2018)

Figure 6.14: Two tweets that have been retweeted over 200 times and therewith cause the first peak

	Peak 1	**Peak 2**	**Peak 3**
Date	July 26, 2018	August 19, 2018	September 6, 2018
Cause	Two tweets, one by Mike Hudema and one by Protect All Wildlife. See figure 6.14	World Orangutan Day, celebrated annually on August 19. The most frequently retweeted tweet over this peak was by BBC Earth, raising awareness for the 'perishing Borneon orangutans'. The most frequent hashtags are, besides the usual # palmoil: - # worldorangutanday - # orangutanday - # internationalorangutanday - #orangutans	Unknown, there is no clear hashtag or retweet causing this peak. Possible explanations (see appendix F for elaborations). - 38% of the tweets mention RSPOtweets - Certain accounts have tweeted a lot of similar messages (up to 9/minute). See figure 6.15 for examples. The official limit is 100 tweets per hour (Twitter, 2018a), broken down in smaller sections. With 9 tweets/minute (as infigure 6.15a), the user is well above that. - Certain tweets do not exist anymore. We can see for a specific user called Freja_Petersen that her tweets do not exist anymore. The url says "Sorry, that page doesn't exist!"
Existing time (of peak)	Approximately 5 hours	Approximately 30 hours	Approximately 7 hours
Disappearance (back to 0)	After 24 hours the retweet is close to 0		
Issue publics	The two retweets issue publics by satisfying all characteristics (separately)	World Orangutan Day is an example of issue publics by satisfying all characteristics	

Table 6.8: Different peaks with data, cause, existing time, disappearance and if they fulfill characteristics of issue publics

Conclusion characteristic 1: Twitter functionalities (retweets ans hashtags) create space for issue publics

We live in a time with access to lots of information and lots of media fighting for our scarce attention. Argued is how we have shorter attention spans due to the many distractions, giving rise to issue publics. Issue publics are publics that form around an issue and are shorter-lived and more dynamic than public sphericules and public spheres. Retweets and hashtags are argued to cause these issues and thus the publics around them. Two out of three peaks analyzed have clear retweets and hashtags as their cause. One is caused by two frequently retweeted tweeted tweets (one by Mike Hudema and one by Safe all Wildlife), and the other peak is caused by #worldorangutanday. Both issues have gathered are public around them that has retweeted and tweeted with the specific hashtag frequently enough to cause a serious peak in the data. Twitter can, in these cases, be argued to function like an echo chamber where users are active only retweeting similar ideas. The third peak analyzed, however, does not have a clear cause for this peak. Figure 6.15 shows some peculiarities in the tweets, such as hyperactive accounts tweeting 9 times a minute and therewith exceeding the Twitter limit. Exceeding the limit is possible since the Twitter limit is expressed per day, 'broken down in smaller segments'. It is undefined how small the segments are broken down into, but clearly larger than per minute. To conclude, it can be argued that retweets and hashtags indeed create the space for and often create issue publics.

(a) Nine tweets containing the same message and tweeted within the same minute (retrieved September 18, 2018)

(b) Four tweets mentioning RSPOtweets- they are similar but different and tweeted in a short period of time (retrieved September 18, 2018)

Figure 6.15: Multiple tweets mentioning RSPOtweets- they are similar but different and tweeted in a short period of time (retrieved September 18, 2018)

6.3.2. Certain issues push others away in a time where our attention is scarce

This section looks into the attention economy by analyzing if popular tweets pushing other tweets away. This analysis is based on the same peaks as in section 6.3.1. Unfortunately Twitters trending algorithm is not transparent and therefore it is unclear what the effect is of making topics trending (or even if they have been, for some users). The following sections describe peak 1 and 2, which have been classified as issue publics. Since peak 3 does not have a clear cause, it cannot be seen if this tweet pushes other tweets away.

Peak 1: July 26, 2018

Figure 6.16 shows two graphs of time series between 07/25 and 07/27. It shows the amount of tweets between 07/25 and 07/27 including a peak caused by Mike Hudema and Protect All Wildlife (top line) and excluding the peak (bottom line). The graphs clearly show how -when the popular tweets taken out- the peak disappears but the level does not drop below normal. This shows that these two tweets and their retweets did not lead to the pushing away of other tweets.

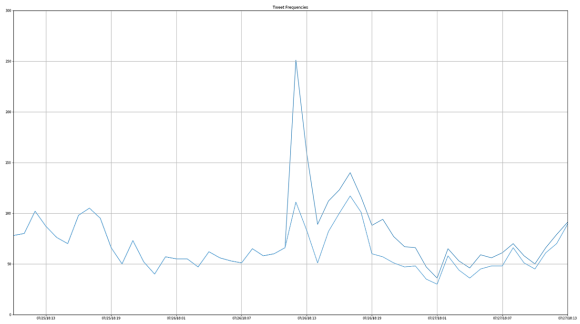


Figure 6.16: Time series of the amount of tweets between 07/25 and 07/27 including a peak caused by Mike Hudema and Protect All Wildlife (top line) and excluding the peak (bottom line)

Peak 2: August 19, 2018

Figure 6.17 shows three times series on top of each other. The top line shows all the tweets over these days. The middle line shows all the tweets containing orangutan and the bottom one shows all the tweets containing one of the above mentioned hashtags for World Orangutan Day. These lines show how, like in peak 1, a popular theme does not push away other tweets. There are still above average tweets that day.

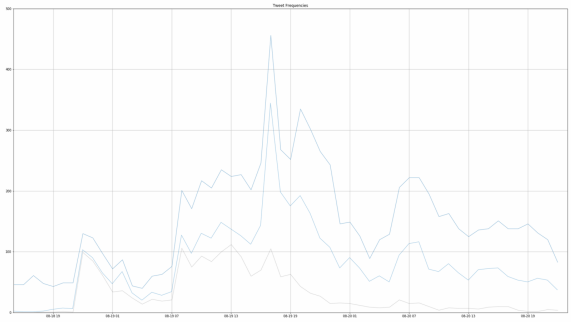


Figure 6.17: Three different time series of World Orangutan Day (August 19), from 09/18 to 09/20

Conclusion characteristic 2: certain issues push others away in a time where our attention is scarce
It is difficult to draw conclusions with regards to the attention economy because conclusions should be drawn over a wider scope instead of just within the palm oil debate. Gathering tweets on other topics falls outside the scope of this thesis. The first two peaks show how, within the palm oil debate, issues are not pushed away. They are either stable or amplified by the discussion, giving the whole discussion on palm oil a boost.

7

Characteristics of the debate

To answer the fourth sub-question, *What can be generalized about public opinion on Twitter from the case study of palm moil?*, conclusion are drawn based on the previous sub-question. Figure 7.1 shows the selection of the framework from chapter 3 that accompanies this chapter highlighted in green. Phase G is highlighted. Here, conclusions from the case study of palm oil are generalized. The structure of the previous chapters- based on concepts and constructs- is released and overarching characteristics are defined. But before releasing the concept-construct structure, the next section sketches an overview of how the different constructs come together to see where the conclusions come from in the sections thereafter.

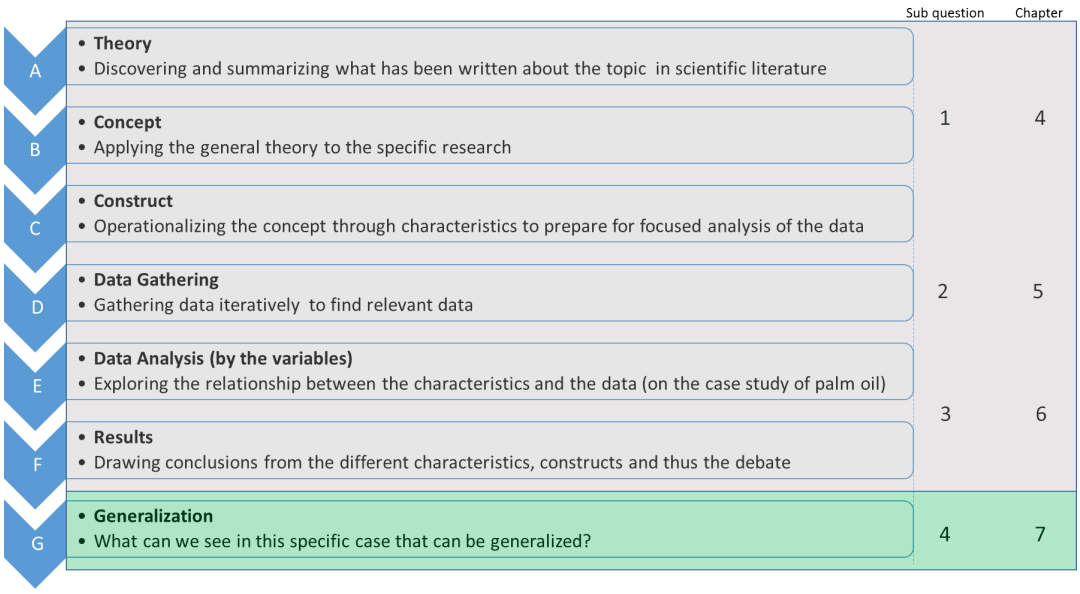


Figure 7.1: Framework introduced in chapter 3 with phase G highlighted

After the general overview in the next section, the different types of users are described in section 7.1. To be a user on Twitter, one needs access to the internet, enough cultural capital and new media literacy. Cultural capital is mostly gained through an exclusive education system, excluding certain groups like in the eighteenth century (Docherty, 2015). New media literacy involves different skills that are crucial for living in the participatory (digital) society (Koc and Barut, 2016). These skills include reading and writing media and critical thinking (Koc and Barut, 2016). The need for cultural capital, new media literacy and internet access do not make Twitter automatically a public sphere in which all can participate.

The users can be divided into humans and bots and additionally into legitimate and spam users. The different influencers are described in section 7.2. Influencers are the glue of Twitter networks with their large share of total information diffusion. One does not have to tweet about a certain topic to be an influencer due to other users talking about them. Additionally, policymakers look at influential people and organizations in real life and follow them on Twitter, where they can be argued to function as an activated public opinion. An activated public can anticipate trends in wider public opinion (Andrea and Fedra, 2016).

The different (types of) arguments in tweets are described in section 7.3. The twitter functionalities of retweets and hashtags play an important role in spreading information, just like calls for action and the use of videos and images. A tweet can not shine light on all sides of a wicked problem in 280 characters. To convince the reader of the message, many tweets tell only a part of the (much more complex) story and interpret data in a way convenient for their message. The different topics and their coherence is described in section 7.4. Topics on Twitter are often not mutually exclusive but can be overlapping. The size and duration vary. Lastly, topics can influence the popularity of the bigger theme they are part of.

Overview of the different constructs

Figure 7.2 shows an overview of how the different constructs come together. All separate aspects are elaborated on in the next sections. This overview can be helpful to put the constructs into perspective. The data in this chapter (on which figure 7.2 is based) is an attempt to explore the relationship between the literature (with the concepts and constructs) and the tweets gathered. A discussion, to see if Twitter actually can be seen as a a public sphere and if public sphericules and issue publics exists, is conducted in chapter 9.

The squares in figure 7.2 can be seen as matryoshka dolls, with the object-within-similar-object relationship. Public spheres, sphericules and issue publics are all based on the same characteristics (see section 4.2.1). The outside square is internet as a public sphere. As can be seen, not all actors are included. The three icons on the left are actors with no internet access, not enough cultural capital or no new media literacy. They cannot be part of the debate. The square inside is twitter as a public sphere. Inside the Twitter square are multiple smaller squares (public sphericules). The public sphericule investigated in this thesis is palm oil. One layer deeper is deforestation taken as an example of a public sphericule, with inside the orangutan and sustainable palm oil against deforestation. World orangutan day was an event on August 19 and caused a peak in tweets, that disappeared as quick as it emerged. Therefore, world orangutan day can be seen as issue around which a public formed and disappeared. RSPO is a public sphericule inside the sustainable palm oil against deforestation sphericule.

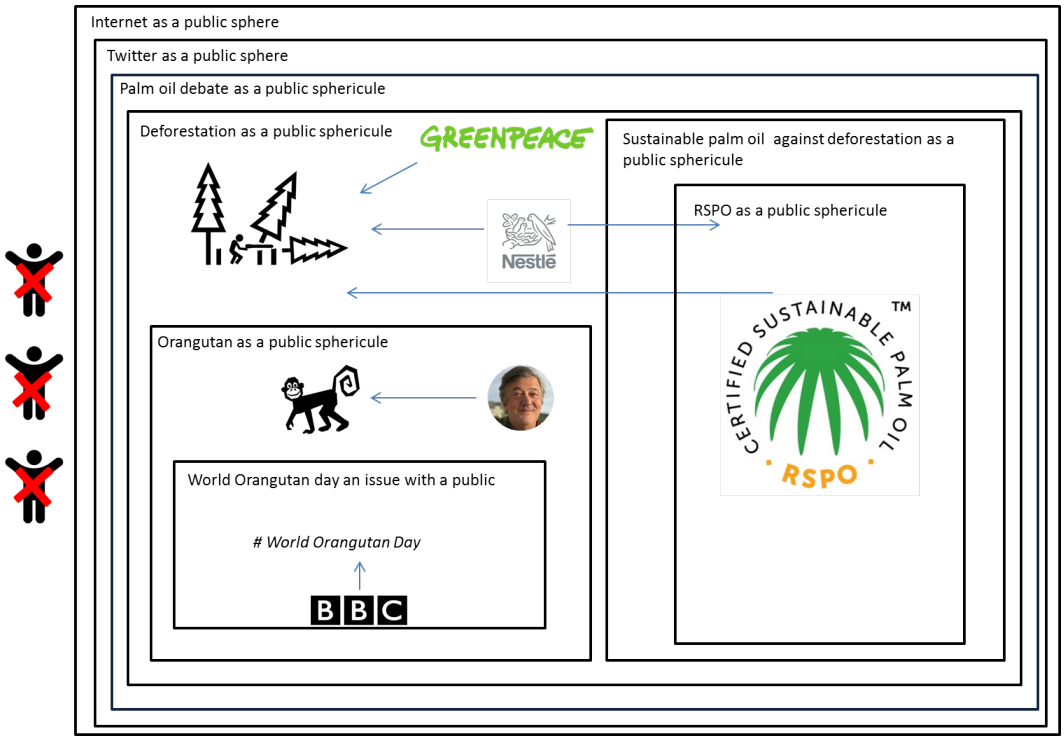


Figure 7.2: Overview from Twitter as a possible public sphere with its public sphericules and issue publics. An example of included actors is shown in the square and excluded actors on the side.

The actors in figure 7.2 are a small selections of those participating in the public sphericule of palm oil. See section 7.1 for a more complete view. Greenpeace, BBC and RSPO are nonprofit actors, whereas Nestle is for-profit. Stephen Fry has played an influential role in the debate on palm oil, especially with regards to the orangutan. See section 7.2 for other influential actors and generalizations on influencers. Inside the public sphericule of deforestation is the Round table of Sustainable Palm oil shown as a separate sphericule *and* actor, influencing the deforestation public sphericule. *How* different users try to influence public opinion can be seen in section 7.3 and the relation between different topics is elaborated upon in section 7.4.

7.1. The different types of users

There are 336 million monthly active Twitter users worldwide (Statista, 2018b). Anyone can create an account and participate in the online sphere of Twitter, provided that the entity has enough cultural capital, new media literacy and internet access. Additionally, different types of organizations like (non-)governmental organizations, (non-)profit organizations and different media organizations are on Twitter.

The different types of Twitter users are defined as human or bots, with the sub-definition of legitimate and spam (see figure 4.4). Argued is how on Twitter 15% of the active users are bots. These do not all have to be malicious bots as can be seen in the figure. The above mentioned organizations can use social bots for legitimate purposes such as web care. The behaviour can, on the other hand, be violating Twitters term of use (Twitter, 2018d) and be classified as spam.

7.1.1. The different types of users on palm oil

There are 65.417 unique accounts in the database on palm oil, which accumulates to approximately 65% of unique users in all the tweets gathered. This section describes the different users and the next section elaborates on spam users. Based on a selection of users (mostly mentioned, mostly retweeted and mostly tweeted), six different categories are present. This list is not mutually exclusive nor collective exhaustive but solely based on the users analyzed as described in section 6.1.3. These categories are businesses (22%), individuals (42%), NGOs (20%), nonprofit (7%), media (7%) and government (2%). These categories do not align directly with private or public interest. Private interest in this case means narrating a strategic argument. An example is the ex-CEO of PepsiCo, who tweets from her private account but in strategic and thus private interest. Another example is the influential user Mike Hudema, who is besides an individual an activist for Greenpeace.

Spam users

Besides the above mentioned categories there is the category of spam users. Two types of spam have been identified in the data:

1. **Fake accounts:** retweeting specific tweets and mentioning certain usernames by accounts that do not seem legit and are deleted by Twitter. See the figures 6.5a for the accounts that have been deleted all with 0 followers, figure 6.6 for the tweet that has been retweeted many times by fake accounts and figure 6.7 for the user- and screennames of the accounts that have been deleted.
2. **Hyperactive automated users:** many tweets containing the same message in different wording by one user, violating the official Twitter limit of 100 tweets an hour (Twitter, 2018a). See figure 6.15 for two examples: one for nine tweets in one minute, and one for four tweets in one minute. Exceeding the limit is possible since the Twitter limit is expressed per day, 'broken down in smaller segments'. It is undefined how small the segments are broken down into, but clearly larger than per minute.

7.1.2. General characteristics of the different types of users

Four characteristics can be defined based on the case study on palm oil for the general Twitter sphere on the different types of users.

1. Only those with internet access, enough cultural capital and new media literacy can participate on Twitter.

   The exact effect on public debate of those lacking internet access, cultural capital and new media literature is not clear. I cannot measure what is not there and therefore, this thesis has only looked at those that are participating. Twitter gives all its users a medium to express their opinion. This does not necessarily mean that they have any influence on the discourse, as will be elaborated on in section 7.2.

2. Twitter users can be divided into humans and bots and further into legitimate and spam users.

The difference between humans and bots is how bots post autonomous. When they post moderately and within the Twitter usage policy, they do not have to be malicious. An example of a legitimate bot is a bot used for web care by an organization on Twitter, to have automatic and autonomous replies to users with questions. This has been seen in this thesis by PepsiCo, tweeting to all users with questions the same reply.

3. Legitimate users exist in at least exist 6 categories: businesses, individuals, media, NGO, nonprofit, governmental.

   It is not possible to say anything about the percentages in general and neither about the different interests of different parties. It can be argued how some pursue a strategic interest (like businesses) and others a public interest (like governmental organizations). The lines between individuals and organizations are blurred sometimes, like in the case of the CEO of PepsiCo, tweeting from her personal account about the mission of PepsiCo.

4. Spam users exist in at least two categories: fake accounts and hyperactive automated accounts.

   Spam users often produce irrelevant content and violate Twitter's term of use. An example from my data is a user who tweeted 9 times in 1 minute with a similar message to boycott an Indonesian palm oil company. Twitter has deleted 70 million fake accounts in May and June to decrease the flow of misinformation on the platform (Timberg, 2018). Both types of users influence public opinion on Twitter but it is difficult to estimate to which extent, since Twitter does not enclose information on their spam users.

7.1.3. Interpretations on the types of users

This section interprets the descriptive characteristics from section 7.1.2 in a normative way. According to Habermas, the public sphere should be free of power and ideology. The only common interest should be the interest to create good opinion formation (Kuitenbrouwer, 2018) and it can be argued that a strategic interest (such as profit making) does not have good opinion creation as the main interest. Many of the users on Twitter are organizations or representatives from organizations. Besides, many accounts are spam (either bots or legitimate), also not with the interest to create good opinion formation but out of strategic interest.

The users who all bring in their public opinions lead up to public opinion on a certain subject. Whenever the users that are not excluded (due to lack of internet access, cultural capital or new media literacy), many participate from a strategic interest instead of out of interest of good opinion formation. It is also difficult to argue how all the organizations participating do not enter with power and ideologies. Therefore, the users participating on Twitter do not fulfill the characteristics of the users as Habermas described them. I question the quality of public debate due to its strategic, organizational and spam users on Twitter and therefore the public opinion following from that debate.

7.2. The meaning of influence and the different influencers

There is no unanimous agreement on when exactly a Twitter user becomes an influencer. Influencers on Twitter are in this thesis defined based on different resources like the number of followers, following, tweets, replied to, retweets, mentions and favorited. These influencers are, with their large share in the total information diffusion, described as the glue of Twitter networks (Patel, 2014). Conversation would not exist without these influencers and the whole Twitter sphere is argued to fall apart.

7.2.1. The meaning of influence and the different influencers on palm oil

This section describes influencers (1) based on the data, (2) through spam and (3) according to the interviewed policymakers. There are multiple influencers found in this data based on the above mentioned resources. They can be divided into different categories:

1. Businesses (multinationals in the fast moving consumer goods (35%) and palm oil companies (25%))

2. Individuals (10%)
3. Media (10%)
4. NGO (10%)
5. nonprofit (10%)

As can be concluded when comparing this list with the list of users in section 7.1.1, many different types of users can classify as an influencers. The only category missing from section 7.1.1 is the governmental organization. The role of individuals as influencers is much smaller than the amount of individuals in the selection of the data (42% vs 10%) and the role of companies much bigger (60% vs 22%). This difference is partly due to the fact that one can become an influencers on a certain topic *without* tweeting about that topic but just being tweeted to or about. This is more often the case with businesses than with individuals, since they are more well-known by a broader public.

Influence through spam

The different types of spam influence the users. As can be seen from the first case in section 7.1.1, the Indonesian Investment Fund bpdp_sawit seemed like a very influential users in the tweets gathered with many user mentions and retweets. When looking at the influence *now*, with most of the accounts that gave rise to the status of influencer deleted by Twitter, bpdp_sawit does not have much influence anymore.

Influencers according to policymakers

The policymakers from the ministry of social affairs and employment do not use social media data to measure public opinion. They use Twitter data to follow influential people from the Hague (politicians, journalist and unions). These users can not be seen as the whole public but can be argued to represent an activated public opinion, anticipating trends in the wider public opinion. The Amsterdam municipality, on the other hand, monitors everything that happens on social media by an easy-to-use-interface. This does not mean the monitoring directly influences policy making but creating awareness under policymakers is a first step in using the data for policy making.

7.2.2. General characteristics on influencers

Four characteristics can be defined based on the case study on palm oil for the general Twitter sphere on influencers.

1. Influencers are the glue of Twitter networks with their large share of total information diffusion.

   There is not a list of criteria agreed upon in literature on the definition of influencers, but a large share of information diffusion is achieved when many other users engage with one's tweets (either by just looking/reading or by liking or retweeting). One can become an influencers in different ways, for example by having a lot of followers, by being retweeted a lot, by being mentioned a lot or by using hashtags and other rhetorical devices possible in tweets.

2. Influencers exist in at least five categories: businesses, individuals, media, NGO and nonprofit.

   The number of followers and the type of user do not determine if one is an influencer. With only 20.000 followers one can be the most retweeted users and individuals as well as businesses can gain an influential role. It can be argued to be easier for businesses with more brand awareness to become an influencers, but nothing is impossible in the age of Twitter. Individual influencers are often linked to an organization.

3. One can become an influencers on a certain topic without tweeting on that topic.

   This can happen in multiple ways, for example when other influential users tweet about you. Imagine someone with over a million followers mentioning you: this will automatically increase exposure to your profile. Another way is when something has happened outside the Twitter sphere that is brought into it, such as a scandal discovered by a main media outlet.

4. Tweets of influencers of public opinion in real life are used by policymakers as an activated public opinion.

   Policymakers use the tweets by politicians, journalists and unions to get an idea of the activated public opinion. These tweets might lead to a self-fulfilling prophecy, where policymakers assume the public thinks in a certain direction (based on the activated public opinion) and act upon that, leading to policies that make the public act in a certain way.

7.2.3. Interpretations on influencers

This section interprets the descriptive characteristics from section 7.2.2 in a normative way. Argued is how the ordinary public in the coffeehouses of the 18th century did not have a say but how communication was carried out by an elite, the bourgeoisie. The role of the bourgeoisie made it, even though Habermas wanted it to be, not a very egalitarian place. The same goes for Twitter. Not everyone can gain an influential role in the network and many with influence are organizations, who, as described in section 7.1.3, participate from a strategic interest. I think the major role of influencers in the debate on Twitter leads to a distorted public opinion.

7.3. The different types of arguments in tweets

Tweets can contain only 280 characters so there is no space for an elaborate description of one's opinion with light on all sides. Certain Twitter functionalities (retweets and hashtags) play an important role in the diffusion of information throughout the network. Retweets are argued to often happen based on the users imagined audience and hashtags are used to kick-start or join a discussion of interest. These functionalities do not belong to a single argument but they do belong to a type of argumentation.

Users on Twitter make use of videos, images and links for many reasons. Some of these are to emphasize or clarify messages, shock others, grab attention or for advertisement. Some of the tweets contain a call for action (direct, semi-direct or indirect), for example to sign a petition or not to buy something anymore. They try to convince one by telling part of a story and data interpreted in a certain way to suit their message.

7.3.1. The different types of arguments in tweets on palm oil

The database with over 100.000 tweets gathered shows multiple arguments and types of arguments. The content of the arguments is discussed in section 7.4. The following three sections elaborate on (1) retweets and hashtags, (2) calls for actions and (3) telling part of the story and interpreted data, both to suit the story and (4) the use of images and videos.

Retweets and Hashtags

Of all the tweets gathered, 37% contains at least one hashtag and 62% are retweets. Hashtags create debates around issues like palm oil in general or on issues more specific like world orangutan day. Hashtags can emerge spontaneously from within the Twitter community, such as #Orangutan.

Others are pre-planned, like the hashtag by RSPOtweets #RT16 to acquire awareness for the next RSPO event. Twitter's trend feature can amplify certain hashtags. The feature, as described by Twitter, tailors trends by an algorithm just for you, based on your interest (Mauskopf, 2012). Once a story achieves enough attention to be upgraded to be trending, it attracts even more attention. Morozov argues how Twitter's trend feature makes Twitter an engine creating realities instead of a camera reflecting realities (Morozov, 2013).

People retweet based on the imagination, or perception, in which people imagine their social existence. These retweets can therefore create a public much larger than the initial group of users that have seen the tweet. Retweeting is an active form of engaging with a tweet, compared to the more passive forms of simply looking at a tweet. Retweeting and hashtags cannot be seen separately since many retweets contain hashtags.

Call for action

Multiple calls for actions are possible in tweets. A call for action can be very direct, semi direct or indirect. More information on calls for action is in section 6.1.2. Examples of direct calls are the ability to buy a piece of land to replant forest or to sign a petition. An example of a semi direct call is to tell big brands to drop dirty palm oil since not many people will go out and actually tell the big brands. Lastly, a shout out for awareness can be seen as an indirect call for action. Out of the seven most frequently retweeted tweets, four ask for direct action by either a signing a petition or giving a donation.

Telling parts of a story and data interpreted in a certain way

Most users want to optimize the message in one tweet. By telling only the snippet of the story relevant for your message and interpreting the data in a certain way, also relevant for your message, you can create a strong and convincing message. The data is in most of most frequently retweeted tweets certainly interpreted and not containing raw data: that would leave too much open for the interpretation of the reader. The part of the story that is communicated is pretended to be the whole story. Especially when using movies, see the next section.

The use of images and videos

All of the most frequently retweeted tweets contain an image or a video, possibly with an extra URL (linking to more information, articles, videos, etc.). The photos are grabbing one's attention and visualizing the story in the tweet. The videos pretend to explain the whole story around palm oil in a bit over a minute. Spoon-feeding information leaves little space for interpretation or to fill in any blanks.

7.3.2. General characteristics on argumentation

Four characteristics can be defined based on the case study on palm oil for the general Twitter sphere on argumentation.

1. The twitter functionalities of retweets and hashtags play an important role in spreading information on Twitter

   Even though not all tweets contain a hashtag (varying from 20% on average (Orellana-Rodriguez and Keane, 2018) to 37% in the data on palm oil), the hashtag as a clickable hyperlink is a strong mechanism to amplify certain keywords. The same goes for retweets. When a single tweeted is retweeted thousands of times, the reach is automatically much larger than it was before the retweets.

2. Tweets can contain multiple different calls for action

   Direct calls for action try to convince people to act right away. Semi-direct calls for actions do not make people act right away but does give direction to how to act in general. Indirect calls for action gives readers something to think about without expecting any action.

3. Tweets often tell part of a story and interpreted data

   Many tweets try to convince us from something and often, the whole story on a wicked problem, is not very grasping and difficult to explain in 280 characters. Therefore, tweets often only tell part of the story, information snippets, relevant for convincing you and backed up by data interpreted in a certain way.

4. Tweets often use images and videos to strengthen their message

   An image is worth more than a thousand words says it all. Additionally videos are used to spoon-feed one information and leave no space for interpretation. These images and videos also attract and keep one's attention, since we are living in a time where different parties online are vying for our attention.

7.3.3. Interpretations on argumentation

This section interprets the descriptive characteristics from section 7.3.2 in a normative way. A tweet can only contain 280 characters and these make it difficult to put out a very nuanced debate. These 280 characters are backed up by images and videos to strengthen the message but usually not to add any nuance. The same effect goes out from the snippets of information and interpreted data. Data and technologies are not neutral, whereas many people think it is. The use of data gives unfair credit to a tweet. I think all the above mentioned tools for argumentation (retweets, hashtags, calls for action, snippets of a story, interpreted data, images and videos) are extremely useful framing tools and those who use them smartly have the most convincing tweet. There is nothing wrong with using all these tools when the ordinary audience is aware of the the use of these tools as tricks to simply convince them, and not to tell the whole story. But I think that the audience is not aware of these tools, and therefore I do not know how fair the process of public opinion formation is when these tools are used without awareness.

7.4. The different topics and their coherence

We are living in the so called attention economy, where we have many distractions fighting for our attention and limitations on how many issues we can actually discuss in public debate. As described by Dewey and Rogers (2012), publics form around issues. The amount of issues we can form a public around is limited, since the quality of public debate decreases when there are too many issues involved. This will lead to people knowing lots of things, without having in-depth knowledge and therefore lacking the possibility to come up with reasonable arguments in public debate. Additionally is argued how we have shorter attention spans than we used to have. This combination leads to many short-lived topics entering our mind. These topics do not all have to be separate but can be overlapping. Argued is how, when more people take interest in more issues, the overlap of the issues counters fragmentation (Habermas, 2006). So different topics do not have to cause a fragmented public.

7.4.1. The different topics and their coherence on palm oil

The debate of palm oil contains different topics. Some are large (like palm oil in general) and some are smaller (like world orangutan day). This section (1) elaborates on the different topics and (2) puts these findings in the light of the attention economy. The topics in the palm oil debate are not mutually exclusive but often overlapping and/or belonging together. 65% of the most frequent hashtags are related to deforestation, just like 100% of the most frequently retweeted tweets. They are not all the same or on exactly the same subject but ranges from an animation movie on the orangutan, to an image of a tarsier and a link to buy land to replant forest. Depending on what level one looks, the debate can be argued to be fragmented on the lowest level (a tarsier and an orangutan are not the same). But on a higher level it is not fragmented, since the common denominator is deforestation.

Attention Economy

In a time where our attention and the amount of issues we can have a debate on in public debate are scarcities, argued is how issues push each other away for attention. The opposite turned out to be true in the case study on palm oil. When a certain issue went viral, such as world orangutan day, the whole debate on palm oil got a boost. It seems to raise awareness of the issue of palm oil in general. What happens to other issues (e.g. climate change) has not been investigated.

7.4.2. General characteristics on topics

Three characteristics can be defined based on the case study on palm oil for the general Twitter sphere on topics.

1. Topics on Twitter are often not mutually exclusive but can be overlapping.

   Within the topic of palm oil fall multiple different topics from which some are overlapping, like deforestation and the orangutan. It can be argued that this is the case all over

Twitter. The topics are, besides not mutually exclusive, also not collectively exhaustive. Not all topics or sides of a debate are discussed, for example not all scientific findings are published on Twitter.

2. The size and duration of topics varies.

   It can be argued how smaller topics have a shorter duration than the longer topics and are engaged with by fewer people. This has to be looked into further. Twitter is also able to influence the duration and size of the different issues by deciding what tweets to put in your stream and what topics to make trending. An elaborate discussion on this can be found in chapter 9.

3. Topics can influence the popularity of the bigger theme they are part of.

   When a certain topic gets a lot of attention on Twitter, that might work as an accelerator for other topics around. It can be even be argued that when something goes viral on Twitter more people engage with Twitter as a platform and therefore also discuss other topics.

7.4.3. Interpretations on topics

This section interprets the descriptive characteristics from section 7.4.2 in a normative way. It is important to note that the debate on Twitter is not the whole debate. The offline conversations and online conversations on other platforms cannot be forgotten. Topics discussed on Twitter are neither mutually exclusive nor collectively exhaustive, making it difficult for the audience to get an overview of the actual topic. Certain topics gaining popularity influence the popularity of others.

This is all not very different from the debate in coffeehouses. Public opinion approached in these places could never cover all sides of a topic. Certain topics gaining popularity would also here influence what is being discussed more. There are, in my opinion, two main differences between Twitter as a public sphere and the coffeehouse as a public sphere. The first difference is the increased supply of topics fighting for our attention in the Twitter sphere. I think a combination of the above mentioned interpretations -on the strategic users, the distorting influencers and the tricks used to convince us- makes the arena for public debate more extreme and outspoken. This, in turn, leads to public opinion formation in a very different way from how it has ever been formed before.

The second difference is based on way public debate happens in the online public sphere. The online debate can be argued to be quicker and therewith less profound than offline debate. It can be argued that this makes the participants less reasonable, since their arguments do not have to be backed up by reason. When looking at the public around issues by Dewey, the public is that which forms itself around an issue and consists of anyone who is effected by or concerned with the issue (Dewey and Rogers, 2012). The increased ease in which people can get concerned with issues due to the accessibility of information and possible participation online does not automatically lead to better public debate. This thesis has not investigated if it leads to worse public debate, but in my opinion, the one-sided view of palm oil by lots of people who are not experts but just want to have a say on many different topics they feel ‘concerned by’ does not lead to the highest quality of public debate possible.

8

Conclusion

The previous chapters have shown insight in opinions on Twitter, based on the case study of palm oil. The research went from general to specific and back to general (see the hour-glass figure 3.1). This chapter connects the research objective as introduced in chapter 1 to the final outcomes and therewith bringing the whole thesis together. Section 8.1 is a short recap of the problem causing the reason for this thesis. Section 8.2 attempts to answer the sub research questions and the main research question. Sections 8.3 and 8.4 mention the scientific contribution and societal relevance of this thesis.

8.1. Short Recap

The world faces many challenges today, from climate change to geopolitical tensions and from food security to artificial intelligence. These issues of the 21st century concern a broad range of societal and environmental issues and can be classified as wicked problems (Duckett et al., 2016) or grand challenges (Grandis and Efstathiou, 2016). The term *wicked problems* was first coined by Rittel and Webber (1973). Wicked problems have different characteristics, including incomplete, contradictory, and changing requirements that are often difficult to recognize. Additionally, there is no indisputable public good nor single definition of equity in society, meaning there is no optimal solution to these problems possible (Rittel and Webber, 1973). *Grand challenges* do not have a single unambiguous definition, but they indicate the great problems of our age. Many of the characteristics of wicked problems and grand challenges are the same.

Many of these challenges are incorrigible and extremely challenging for policymakers. A better understanding of public opinion better can be helpful for policy makers in dealing with challenges. Ultimately, policy is made for the public, also in the case of wicked problems and grand challenges. Some policy makers already take public opinion data into account when making policy: research has shown the correlation between public opinion and policy response (Barabas, 2011; Monroe and Alan, 1998; Page and Shapiro, 1983; Pierson, 2002; Stimson et al., 1995). Public opinion formation has changed over the last centuries due to the differences in life styles and standards. Before the invention of the telephone, the telegram and the television, all interactions were face to face. The Internet was invented almost thirty years ago and has argued to drastically change the way we shape and share opinions. The internet today, with ample amount of data available, can be a useful source to analyze public opinion.

There is a lot of meaningful public opinion data available on social media, but policymakers hardly use this data. They often do not know how to use it, do not see the added value or do not know what is all being discussed. Even though the opinions expressed on social media most probably do not represent the whole public, they can represent an 'activated public opinion'. This, in turn, can anticipate trends in the public opinion (Andrea and Fedra, 2016). Once the policymakers have extracted the public opinion, it can be used in all phases of the policy cycle, from the formation to the implementation of public policy.

This thesis focuses on public opinion data expressed on Twitter. More than 8000 tweets are posted every second (Stats, 2018). These tweets create a lot of data from which it is difficult -especially for policymakers who do not have a background in big data- to separate noise from signal. It is important to thoroughly understand public opinion and the large amount of data available on social media before it can be used in policy making. By analyzing tweets on the case study of palm oil, Twitter is characterized. The red line of research is based on the German philosopher Jürgen Habermas and his book *The Structural Transformation of the Public Sphere*. The public sphere should not be seen as a market nor a coffeehouse, but as a transcending space as an abstract forum for dialogue and public opinion, for vivid debates on multiple levels in society. The theory on the Habermasian public sphere is applied to Twitter.

8.2. Answering the Research Questions

The main research question in this thesis is the following:

How can public opinion on Twitter be characterized, on the case study of palm oil?

To answer the main research question, the four sub research questions are answered first. These answers are summarizing, whereas the answer to the main research question in section 8.2.5 contains the main conclusion. The first sub-question, *what are possible characteristics of public opinions on Twitter?* is answered through literature research. The second sub-question, *what are relevant tweets regarding palm oil?* is answered through data gathering. The third sub-question, *what characterizes public opinion on palm oil on Twitter?* is answered through a combination of the organized literature from sub-question 1 and the tweets gathered from sub-question 2. The tweets are analyzed based on the literature to explore a relationship between the philosophy and the empiricism. The fourth sub-question, *What can be generalized about public opinion on Twitter from the case study of palm oil?* is answered through an interpretation and generalization of the third sub-question. Before policymakers use social media data, it is important they understand how public opinion is shaped on social media and why it is the way it is. This understanding is necessary before the data subtracted is used for policy making.

8.2.1. What are possible characteristics of public opinions on Twitter?

Three concepts are developed to apply general theory to the specific research of public opinion on Twitter. These concepts are operationalized into constructs. It is important to emphasize the ideal-typical character of these operationalizations and characteristics. Assumptions and simplifications are made here to prepare the theory to be linked to the empirical data.

1. Twitter is an incarnation of the Habermasian public sphere

   Concept: As defined by Habermas himself, the public sphere is a “realm of our social life, in which something approaching public opinion can be formed” (Habermas et al., 1974, 49). It is a sphere of people, who from their private interest and space, enter and therewith shape a public. Those forming public opinion are presupposed reasonable, making the rational-critical debate a key feature. Here, may the best argument win and not the one made by the man of highest status. Twitter is argued to be an incarnation of the Habermasian public sphere due to fewer central nodes, fewer gatekeepers and fewer agenda setters, whereas on the other hand is argued how many lack internet access, cultural capital and new media literacy and are therewith excluded.

   Construct: Different characteristics are defined to operationalize the concept.

   - Social inequalities are bracketed

     Showing the different users participating and especially looking at those with an influential role to see if everyone is seen as equal and no differences are made between different users.

- Public Sphere is a new centre of institutional authority
  Interviewing policymakers to see what they use the data for and if they see the public organizing into a counterweight to the traditional institutional authority. Additionally, tweets are analyzed to see if they contain a call for action.
- Access is granted to all citizens
  Mapping different groups of users to find out what categories of users are active. Additionally checking if these users participate out of public or strategic interest.
- Those forming public opinion are presupposed reasonable
  Analyzing tweets to see what argumentation is used, if additional material (such as a video or image is used) and if they are classified as positive or negative.

2. Twitter causes a fragmentation of the public

   Concept: With regards to fragmentation, the term public sphericules is used. These are small fragments with the same above mentioned characteristics as public spheres. Since social media is increasingly used to follow the news by instant updates, small thematic debates are born. When the thematic debates become smaller, so do the sphericules.

   Participation is key on Twitter and the existence of the platform depends on its users. Argued is how this active engagement makes social media more influential than traditional media. On the contrary, only a small community of users posts the most tweets, where all the others are passively taking in information (Kittle, 2011). One does not need to actively go around to fill in any blanks and one does not have to have the feeling they are missing out. This is due to the feeling that the information that reaches one is informed knowledge and thus one is informed.

   Construct: Different characteristics are defined to operationalize the concept.

   - The public on Twitter is fragmented
     Analyzing the hashtags and word frequencies to see what main topics are and if these are fragmented.
   - Twitter causes this fragmentation
     Analyzing tweets to see if they contain images, snippets of information and interpreted data, which can all be argued to add to a fragmentation of the public.

3. Twitter creates space for issue publics

   Concept: According to Dewey, the public forms itself around an issue (Dewey and Rogers, 2012). When we base issue publics on the notion of publics forming around issues by Dewey, it can be argued that publics can only form around a limited number of issues. Especially when we want to achieve coordinated action on these issues. We have to align and coordinate the issues between individuals, and to align, we have to decide upon issues to discuss. When only a limited number of issues is discussed in the public space, the quality of public debate can increase since issues are discussed more in-depth with more elaborate arguments. Additional to this argument of social organization, there is the argument of the psychological limitations of our brain. There are only limited issues we can get involved in. These two arguments are the base for issue publics. Issue publics form around shorter-lived topics and are more dynamic than the public sphericules and public spheres. Certain Twitter functionalities are argued to have created space for issue publics, such as retweets and hashtags. Retweets often show a common interest around an issue and therewith create a public around the issue. Hashtags are clickable hyperlinks that make it easy to form and find an issue, and join the public by using the hashtag. Twitter is argued to play a role in the different issues and topics that enter the users' time line by nontransparent algorithms. Argued is even how Twitter is an engine creating realities instead of a camera reflecting them (Morozov, 2013).

   Construct: Different characteristics are defined to operationalize the concept.

- Twitter functionalities (retweets and hashtags) create space for issue publics
  Analyzing different peaks (when many tweets were posted in a small time slot) to see if these are caused by retweets and hashtags.
- Certain issues push each other away in a time where our attention is scarce
  Analyzing the peaks in more depth to see if certain topics push other topics away.

8.2.2. What are relevant tweets regarding palm oil?

The keywords used to gather relevant tweets are palm oil, oil palm, palm oil, palm oil plantation, rspo, mpob, palmolie, minyak kelapa sawit and aceite de palma. The hashtags are #palmoil, #oilpalm, #rspo, #RefinedPalmOil and #palmolie. These words and hashtags lead to a total of 103.500 tweets gathered in 50 days. are 65.417 unique accounts in these tweets. 75% of these accounts is in English. The word clouds in figure 5.5 and 5.6 give a first impression of the most frequent words and hashtags in the tweets gathered. The words in the word cloud shown most prominently are mentioned the most.

8.2.3. What characterizes public opinion on palm oil on Twitter?

This sub-question is answered through the lens of the previously defined constructs and the different conclusions per constructs are given. Construct 1 concerns the incarnation of the Habermasian public sphere, construct 2 the fragmentation of the public and construct 3 issue publics.

Construct 1 - characteristic 1: Social inequalities are bracketed

Bracketing social inequalities means that the public sphere is of and for everyone, not just for the privileged ones in society. To show who is all participating and who all has a say, the reply network (who has replied to who) and influencers are mapped. The network shows how there are four key actors. Three out of four are multinationals, namely PepsiCo, Nestle, and Kitkat. The other actor is RSPOtweets, which has has a very active role and gained its status by replying more than it has been replied to. The way they have gained their position is different from Kitkat and Nestle, which have hardly participated in the debate on palm oil.

Influencers are based on their central role in the reply network, mentioned the most by other users in this database and are retweeted the most in this database. Here, some fake accounts have had some influence in the debate. Twitter has deleted these accounts but since the data collection is real time, they are captured before they were deleted. These have tried to make the Indonesian investment fund Badan Pengelola Dana Perkebunan Kelapa Sawit show up prominently in the debate but Twitter did not let them. Besides the above mentioned multinationals, the palm oil company Indofood, the NGO Greenpeace UK, the Greenpeace activist Mike Hudema and the media channel BBC earth have influential roles in the debate of palm oil and therewith influence public opinion.

Construct 1 - characteristic 2: public sphere is a new centre of institutional authority

Public sphere is a new centre of institutional authority means it is more than a new place to talk. It is a place sovereign in relation to the state, meaning that it is not part of the state. Social media can be used to organize a completely different public, like the well known #MeToo-discussion. This discussion can be argued to have influenced public opinion on sexual harassment in the work place. In only becomes a new centre of institutional authority when consequences are attached to the public organized but influencing public opinion is certainly a first step. To find out if Twitter is a possible new centre of institutional authority, policymakers are interviewed to see what they use social media data for. Additionally, the tweets are checked for calls for actions.

Interviews are conducted at the ministry of social affairs and employment (SZW) and at the Amsterdam municipality. Both have very different polices: SZW uses social media mostly to see what politicians, journalists and unions post, to anticipate trends in public opinion. The Amsterdam municipality uses social media to monitor what is going on in the city. The monitoring does not directly influence policy making but indirectly and in the long term, it might. The data shows different calls for action such as calls to sign a petition and to donate

money to buy a piece of the rain forest. It has not been investigated what the actual effect of these calls is. The debate on palm oil has in the monitored time (from July 23 - September 1) never formed a large enough centre of institutional authority to gain status and potential to be one like the #MeToo.

Construct 1 - characteristic 3: access is granted to all

Access is granted to all means that anyone with reason and willingness to join is welcome to join the debate. It builds further on the first characteristic but only looks at the groups to which access is granted on Twitter. People who lack internet access, cultural capital or new media literacy do not have access. The users with access can be divided into six categories: business, government, individuals, media, NGO and nonprofit. Individuals have the largest share and it can be argued that their argument is private without being strategic. A multinational like PepsiCo is less likely to have non-strategic arguments. Many individuals, however, post in name of an organization like PepsiCo, which makes it very difficult to make a definite statement about the interest of users.

Construct 1 - characteristic 4: those forming public opinions are presupposed reasonable

Habermas based his public sphere on the 'best argument' and not the argument made by the one with the highest status. It is very difficult to define what the best argument is, but this characteristic has looked at different tweets to see what their argumentation is, if a video/image is used and how negative or positive the tweet is. All tweets analyzed discuss the topic of deforestation, all of them contain a video or an image and most tweets analyzed are positively framed.

Construct 2 - characteristic 1: the public on Twitter is fragmented

Argued is how discourse on social media can be fragmentized and the term used for these fragments are sphericules. They have the same above mentioned characteristics of the public sphere but are smaller. Word and hashtag frequency counts show a clear general direction of public opinion expressed on Twitter. Deforestation is the main topic discussed, inside which are different public sphericules such as the orangutan, exploitation of plantation workers, Indonesia and RSPO. These topics are sometimes overlapping, like the orangutan and the elephant. Some users (like RSPO) can be an influential user and a public sphericule, since part of the debate is concerning their organization. It can be concluded the public on Twitter in the case of palm oil on the highest level is not very fragmented but on the lower levels, different overlapping fragments are present. It can be seen as the square-like matroyska dolls in figure 8.1, where the different topics have topics inside.

Construct 2 - characteristic 2: Twitter causes this fragmentation

Argued is how the fragmentation of the discourse on social media is caused by Twitter. Features of Tweets that can cause fragmentation are the use of snippets of information, the interpretation of data and the use of images or videos. They often tell snippets of information instead of the whole story. 280 characters is very short to tell a complex and complete story and users use that gratefully in their advantage. The same goes for data which is never neutral but can be interpreted to use in one's own benefit. Lastly, tweets often strengthen their convincing power by using images and videos, since a picture is worth more than a thousand words. It cannot be stated that Twitter causes this fragmentation but certain features do add to the quick attention grabbing of users into certain directions and therewith influence public opinion, possibly into different fragments.

Construct 3 - characteristic 1: Twitter functionalities (retweets and hashtags) create space for issue publics

We live in a time with access to lots of information and lots of media fighting for our scarce attention. Argued is how we have shorter attention spans due to the many distractions, giving rise to issue publics. Issue publics are publics that form around an issue and are shorter-lived and more dynamic than public sphericules and public spheres. Retweets and hashtags are argued to cause these issues and thus the publics around them. Two out of

three peaks analyzed have clear retweets and hashtags as their cause. One is caused by two frequently retweeted tweeted tweets (one by Mike Hudema and one by Safe all Wildlife), and the other peak is caused by #worldorangutanday. Both issues have gathered are public around them that has retweeted and tweeted with the specific hashtag frequently enough to cause a serious peak in the data. The third peak analyzed, however, does not have a clear cause. It can be concluded that retweets and hashtags indeed create space and often create issue publics

Construct 3 - characteristic 2: certain issues push others away in a time where our attention is scarce
Within the sphere of palm oil, different issues do not push each other away but do the opposite: when a topic gains a lot of attention and a large public shapes around it, the overarching topic of palm oil also gains more attention. This characteristic is not analyzed based on different topics to see what this effect would be, which might be opposite.

8.2.4. What can be generalized about public opinion on Twitter from the case study of palm oil?

To generalize public opinion on Twitter, the concepts from this thesis are shown in figure 8.1. This figure from chapter 7 is briefly revised, before different generalizations are made on users (types of users and influencers), the argumentation in tweets and the different topics. The squares in figure 8.1 can be seen as matryoshka dolls, with the object-within-similar-object relationship. Public spheres, sphericules and issue publics are all based on the same characteristics: social inequalities are bracketed; public sphere is a new centre of institutional authority; access is granted to all citizens; and those forming public opinion are presupposed reasonable. The outside square is internet as a public sphere. As can be seen, not all actors are included. The three icons on the left are actors with no internet access, not enough cultural capital or no new media literacy. They cannot be part of the debate. The square inside is twitter as a public sphere. Inside the Twitter square are multiple smaller squares (public sphericules). The public sphericule investigated in this thesis is palm oil. One layer deeper is deforestation taken as an example of a public sphericule, with inside the orangutan and sustainable palm oil against deforestation. World orangutan day was an event on August 19 and caused a peak in tweets, that disappeared as quick as it emerged. Therefore, world orangutan day can be seen as issue around which a public formed and disappeared. RSPO is a public sphericule inside the sustainable palm oil against deforestation sphericule.

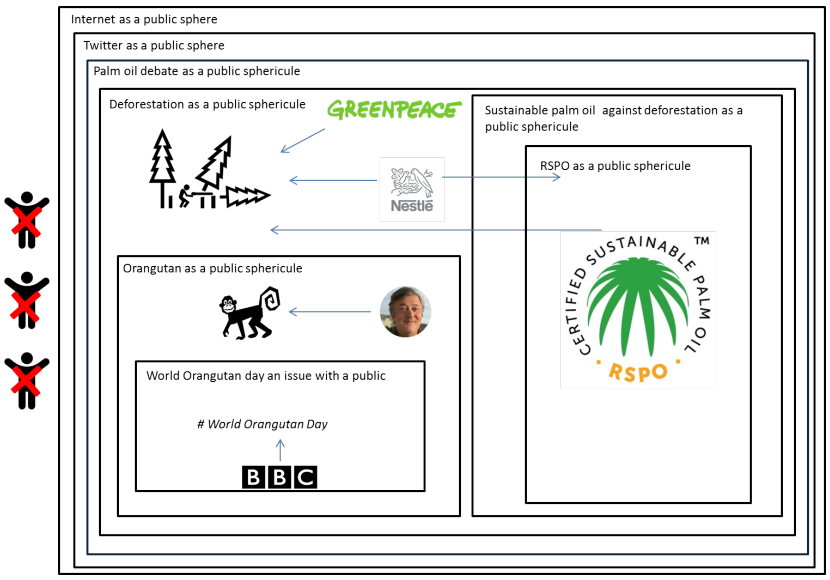


Figure 8.1: Overview from Twitter as a possible public sphere with its public sphericules and issue publics. An example of included actors is shown in the square and excluded actors on the side.

Users

With regards to the different types of users, the most important conclusion is that one can only participate with internet access, enough cultural capital and new media literacy. Besides humans fulfilling these requirements, there are bots active. Bots do not have to be malicious and human do not have to be right: both groups can be divided into legitimate and spam users. Spam users exists as fake accounts and hyperactive automated accounts.

Influencers

Influencers are argued to be the glue of Twitter networks with their large share of total information diffusion. One can become an influencer on a certain topic without tweeting on that topic but just being tweeted about frequently. The influence does not stop at the virtual borders of the Twitter community but influence policymakers and journalist to spread their power further into the offline world. Many of the influencers are part organizations or part of organizations and it can be argued they often participate from strategic interest.

Argumentation

The twitter functionalities of retweets and hashtags play an important role in spreading information on Twitter. Tweets can contain multiple calls for actions. They can be direct (e.g. signing a petition), semi-direct (e.g. telling you to not buy something in the future) or indirect (e.g. to give one food for thought). They often tell snippets of information instead of the whole story. 280 characters is very short to tell a complex and complete story and users use that gratefully in their advantage. The same goes for data which is never neutral but can be interpreted to use in one's own benefit. Lastly, tweets often strengthen their convincing power by using images and videos, since a picture is worth more than a thousand words.

Topics

Topics on Twitter are often not mutually exclusive nor collectively exhaustive. Some of these topics are overlapping or inside each other, like matryoshka dolls. They all vary in size and duration and different topics can influence the popularity of the bigger theme they are part of. The influence of different topics on each other can even spill over into the real world when traditional media pick it up.

8.2.5. Main Research Question

The four sub-questions lead to the answer of the main research question: *How can public opinion on Twitter be characterized, on the case study of palm oil?* The main answer to the research question is the following. All characteristics will be explained separately.

> The findings from this research suggest that Twitter in its current form does not form the optimal environment for forming something that is approaching public opinion due to five characteristics: the exclusion of more than half the world population (1), the domination by a few users (often with strategic interest) (2), the pollution through spam (3), Twitter's role (4) and the characteristics of tweets (5).

Before explaining the characteristics separately, it is important to first note how *the* public opinion is an utopia. Never in our history has public opinion represented *all* public opinions in society. According to Habermas, by rational-critical debate, public opinion can be approached *without* representing all public opinions in the beginning of the debate.

Some places in history have been spheres which give the rise to the possibility to create something approaching public opinion. Twitter has the potential to be a public sphere, "in which something approaching public opinion can be formed" (Habermas et al., 1974, 49). But Twitter might not form the optimal environment for forming something that is approaching public opinion.

1. The exclusion on Twitter

The exclusion of more than half the world population from social media is at least based on the lack of internet access, new media literacy and cultural capital. Less than 5% of the world population is on Twitter, which is a lot less than half the world population. There are many

more factors why some cannot or do not want to participate. Anyone with reason and the willingness to join was argued to be welcome to join the debate in the Habermasian public sphere. The internet has for some made it more accessible to join public debate, and for some more inaccessible. Even though social media is still not for everyone, there has never been a medium before that connected so many people around the world. Twitters reach is broader and crosses more borders than any other possible incarnation of the Habermasian public sphere ever had or did before.

2a. The domination on Twitter

The dominating key players are users who get the most attention from other users, policymakers and journalists. This partly depends on the topic and partly transcends topics. The findings from this research suggest that organizations like Greenpeace, BBC Earth and RSPOtweets have an above average role in steering public opinion. It is difficult to find a user that is saying something completely different than the defined influencers are saying. See section 6.1.1 for who are defined as influencers and how they are defined. These few key players play an above average role in guiding the thinking of the public, and this role is not limited to Twitter. Their thoughts can lead to a spill over of topics discussed on Twitter to topics discussed outside Twitter. And visa versa: influential actors outside Twitter often automatically have a higher status on Twitter and therewith have more influence. Policy makers have indicated in the interviews in section 6.1.2 to mostly read tweets of the users influential outside Twitter, such as politicians and prominent journalists. The role of influential players turns Twitter into a public sphere similar to the Habermasian public sphere, which was an arena of the bourgeoisie. The white, middle class male took an above-average role in guiding the thinking of the public. My findings suggest that on Twitter it is not the bourgeoisie but (non-governmental) organizations and media that have an above-average role in guiding the thinking of the public.

2b. The strategic interest on Twitter

The tweets analyzed show how the dominant key players often act with a strategic interest in mind. The public sphere should, according to Habermas, be free of power and ideology and the only common interest should be the interest in good opinion formation. It can be argued that strategic interest does not have good opinion creation as their main interest. It is hard to distinguish private interest from strategic interest, and as was the case in the Habermasian public sphere, all citizens are private persons bringing in private arguments. But when the arguments brought in are only beneficial for a specific organization, the question is how much these opinions (should) add to public opinion. To clarify, many actors in the debate on palm oil are multinationals like PepsiCo and Nestle. By creating a Twitter account and interacting with other users, these companies are given a face and therewith blur the lines between the private and the public sphere. Their interest is, however, still to benefit the company. See section 6.1.4 for an elaboration on the strategic interest of certain players. When you would, hypothetically, not take any organization into account (which would drastically decrease the number of users), there still is the problem left of individuals representing organizations. An example is the previous CEO of PepsiCo, tweeting from her personal account strategic messages about the good intentions of PepsiCo for our planet. See section 6.1.4 for an exemplary tweet. Times have changed since the Habermasian public sphere is developed, and in our current global capitalistic world, many companies and their spokespeople play a big role in influencing our public opinion. Leaving them out will leave us with an even less representative public opinion, since they have become part of public opinion formation. We have to be aware of an increasing amount of strategic actors present on Twitter as a public sphere, which dominate public debate and therewith influence public opinion.

3. Spam on Twitter

The pollution through spam is argued to play a serious role in influencing public opinion on Twitter. It is unclear how big the role is, but based on the palm oil debate, fake accounts and hyperactive automated users are active. Fake accounts retweet specific tweets and mention certain usernames, to make these accounts seem much more popular than they actually are.

An example of fake accounts in the case of palm oil happened by over 2000 retweets for an Indonesian palm oil investment fund. The accounts that retweeted the investment funds' tweets were deleted not much after they retweeted by Twitter, since Twitter marked all their accounts as fake. See section 6.1.1 for the tweets and an elaboration and the substantiation that these accounts were fake. Hyperactive automated users tweet many similar messages in a very short period of time and subsequently violate the official Twitter limit. An example of hyperactive automated users is a certain user who tweeted nine times in one minute - with nine slightly different tweets, all with the message to boycott an Indonesian palm oil company- and therewith exceeded the official Twitter limit of less than two tweets per minute (Twitter, 2018a). See section 6.3.2 for an elaboration on the hyperactive automated users. The exact effects of fake accounts and hyperactive automated users is unclear but it can be argued, as can be seen from the examples, that they cause a stream of misinformation and therewith influence public opinion.

4. Twitter's role

Twitters algorithms possibly play a role in public opinion on the platform. Technologies are never neutral and McLuhans 'the medium is the message' can be argued to be applicable. The medium influences how the message is perceived. In the case of Twitter, it does not only influence the perception but also what is communicated. Twitter, together with other internet giants like Google and Facebook, can be argued to be the gatekeeper of the internet. They control what ideas and opinions are being shared. Due to their nontransparent algorithms it is very difficult to decipher the exact influence these powerful entities have on public opinion. I cannot prove what Twitter is doing, and if and how they have influenced the debate on palm oil, since the algorithm is nontransparent. Examples like Morozov' accusation of censorship in their "trending algorithm" shows the influence they can have. See section 4.1.3 for an elaboration on the trending algorithm, and section 9.4 for an in-depth discussion on Twitter.

5. The characteristics of tweets

Tweets are often retweets, contain hashtags, calls for actions, snippets of information, interpreted data and images or videos. Retweets can be argued to be echo chambers where users only retweet others with similar ideas. Hashtags can amplify any discussion and calls for action are frequently found on Twitter. Since it is much more direct to ask people to retweet your tweet, sign a petition or ask for a donations, these are the calls of action that frequently occur. Of the most frequently retweeted tweets in my database, 57% contained a direct call of action by either signing a petition or asking for a donation. Snippets of information spoon-feed information while pretending they feed the whole story. However, it is difficult to capture a complicated story in less than 280 characters. To reinforce their message, interpreted data is added. Data is, like a technology, never neutral but it can be used to influence the public of something by using it the right way. Of the most frequently retweeted tweets, 86% uses interpreted data to get their point across. Lastly, a picture is worth more than a thousand words explains why many Tweets show photos of dying orangutans, bulldozers cutting down rain forest or cute animation videos. Of the most frequently retweeted tweets, 57% contains a picture and 43% contains a video. See section 6.2.2 for an in-depth analyses of the characteristics per tweet. To conclude, tweets are not just 280 characters. They are often carefully considered combinations of hashtags, calls for action, snippets of information, interpreted data and images or videos. These characteristics all help in convincing the reader of the point the user is trying to make.

8.3. Scientific Contribution

In this section the additions to the knowledge base are recapitulated. A literature review has been conducted to discover what has been written about public opinion and Twitter. Chapter 2 concludes with a knowledge gap based on existing literature. Public opinion, in this thesis, is based on the line of reasoning by Habermas and the focus in social media is on Twitter. The main scientific contribution aims to contribute to the main knowledge gap and is the following:

Explore a relation between the conceptual Habermasian public sphere on Twitter and results based on actual tweets.

The novelty of this thesis and therewith the main contribution to scientific literature is the elaborate discussion and exploration, based on actual tweets, of characteristics of Twitter as an Habermasian public sphere. This contribution to the knowledge gap is based on the framework I have developed in section 3.2. The framework can be seen in figure 8.2 and is itself an addition to literature. The framework can be used in other research, to combine philosophy with data science. Stages A-C, marked in darker blue, are showing the philosophical investigation on public opinion on Twitter. This combines traditional literature on public opinion, based on the Habermasian public sphere, with more recent literature on social media as a public sphere. There are many more studies focusing on public opinion formed based on traditional media than on public opinion formed on new media, so this thesis gives insight in Twitter as a public sphere. Based on this theory, concepts and constructs are developed.

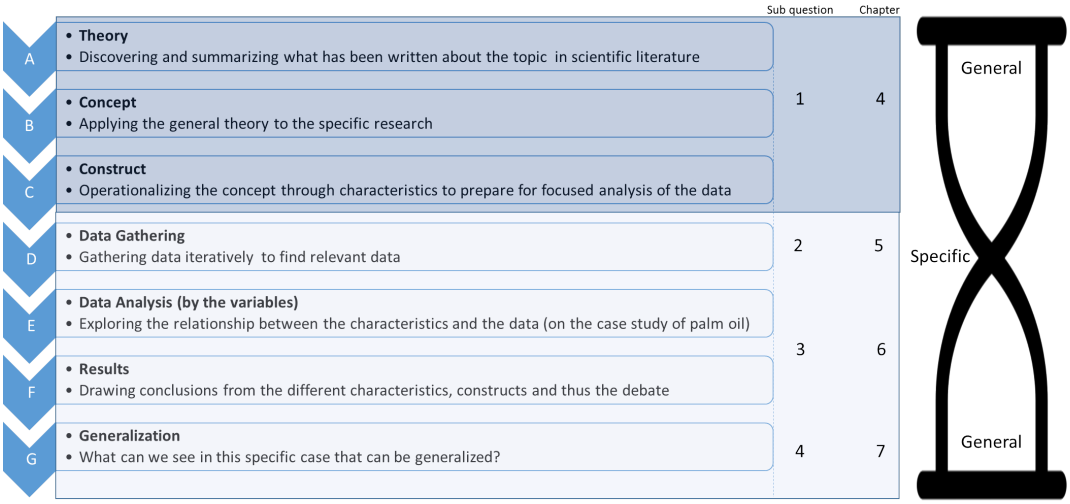


Figure 8.2: Guiding framework in this thesis with steps A-C a philosophical investigation on the Habermasian public sphere and D-G a combination with actual tweets.

Stages D-G consists of an empirical investigation on actual tweets. This empirical investigation is based on the case study of palm oil. The data generated (more than 100.000 tweets) are used to actually see the behaviour by Twitter users and link this back to the Habermasian public sphere. This combination created empirical insight in Twitter as a public sphere instead of only theoretical insight. The insights created are only a starting point. See section 9.6 for recommendations for further research. To conclude, the main contribution to scientific literature is the elaborate discussion and exploration, based on actual tweets, of characteristics of Twitter as an Habermasian public sphere.

8.4. Societal Relevance

This thesis has defined characteristics of public opinions based on Twitter and therewith, this thesis has societal relevance in three areas. Firstly, it creates understanding of public opinion expressed on Twitter. This is not only relevant to inform policymakers, but also for Twitter users in general (individuals as well as organizations). The answer to the main research question shows how Twitter in its current form is not the optimal environment for forming public opinion and I think people should be aware of this. If policymakers take these limitations in mind, social media data on public opinion can be used in different stages of the policy cycle. As can be seen in figure 8.3, social media data can be used in the agenda setting stage to put issues on the agenda that are frequently discussed on social media.

In the policy formulation stage, preferences on different policy alternatives can be analyzed using social media data. In the policy implementation phase, the level of satisfaction can be measured based on expression on social media. Lastly, in the policy evaluation stage, feedback expressed on social media on a specific policy can be used as an alternative feedback model. Not everyone likes direct feedback so analyzing feedback expressed as an opinion on social media can be a valuable addition.

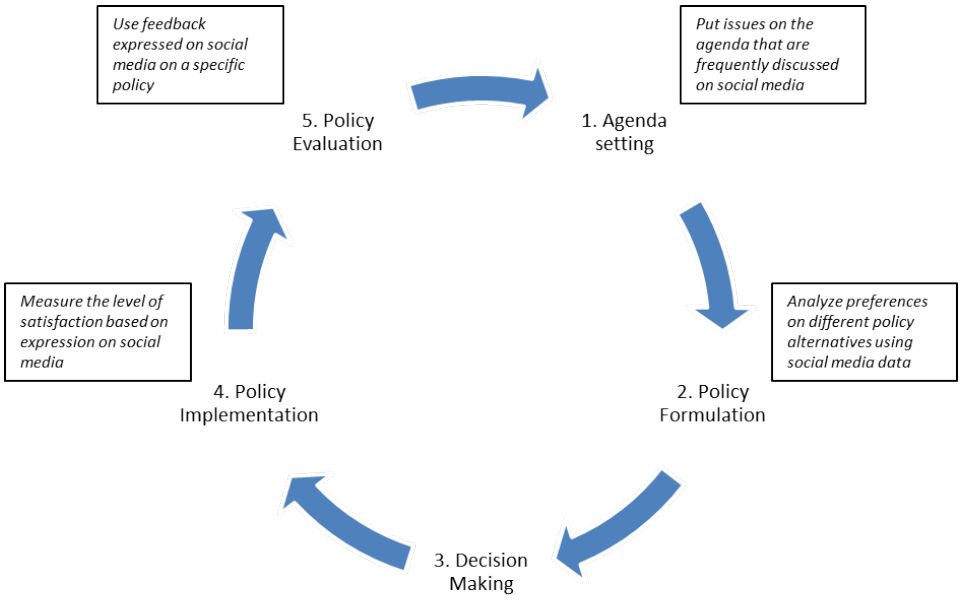


Figure 8.3: Five stages of the policy cycle as described by Howlett and Giest (2015). Adapted to include social media data.

Secondly, policymakers can use this thesis and the Python script to find out what public opinions are on Twitter on any wicked problem. Imagine you are a policy maker at the ministry of Health, Welfare and Sport, and you want to encourage Dutch citizens to eat less meat. You can use social media data to find out if there are aspects of vegetarianism that are already discussed and preferred on social media (in the policy formulation stage of figure 8.3). If you are implementing a policy such as reducing tax on meat substitutes, you can measure the level of satisfaction expressed on social media (in the policy implementation stage of figure 8.3). You can do an in-depth investigation on the users that are dominating the debate on Twitter and therewith influence public opinion. Lastly, you can gather feedback based on what users on social media are expressing about different policies (in the policy evaluation stage of figure 8.3). It is important to keep looking out for different types of spam (such as fake accounts and hyperactive automated users) influencing the debate.

Thirdly, even though palm oil was only used as a case study, insights in palm oil are created. Even though not the whole public is on Twitter, when seeing the users as an activating public opinion, the main online debate mainly expresses its worries about deforestation due to palm oil. Certain actors (multinationals, individuals and NGOs) stand out are are argued to have a lot of influence. Other actors, who can be argue to have a big play in the palm oil sector in real life like Unilever, have managed to stay off radar. Additionally, certain fake accounts have influenced the debate. It can be interesting to find out more about their motives.

9

Discussion and reflection

The conclusions of the previous chapter and the process of this thesis are discussed and reflected upon in this chapter. In this thesis, over 100.000 tweets on the topic of palm oil are analyzed to create insight in public opinion on Twitter. However, this thesis has its limitations. Figure 9.1 shows the figure from chapter 5 with validation specifications. This chapter describes firstly the validation as mentioned in figure 9.1 and secondly reflects critically upon the findings, Twitter and the process.

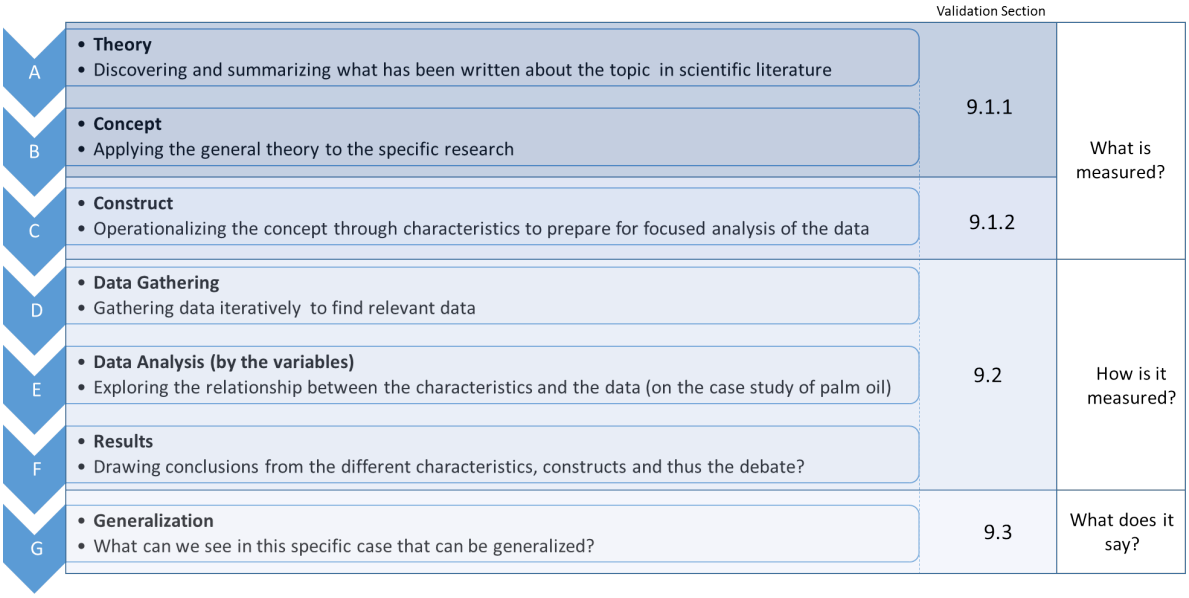


Figure 9.1: Framework adapted for validation

Section 9.1 looks critically at the notion of public sphere in this thesis and at the validity of the different constructs. It concerns steps A, B and C of the framework in figure 9.1 and is therewith looking at *what is measured.* Habermas is chosen as the main line of reasoning and three concepts and constructs are based on his theory. Other thinkers could have been chosen, and even though I believe many roads lead to Rome, the Habermasian public sphere have been of added value when analyzing public opinion on Twitter. The concepts and constructs using the Habermasian public sphere functioned as a lens to look at Twitter and came up with general characteristics

Section 9.2 discusses critical assumptions in this thesis. It concerns steps D and E and F of the framework in figure 9.1 and is therewith looking at *how it is measured* . After all the operationalization from concepts into constructs, there are again many different assumptions made on the most specific level. This section discusses all of these per construct. Examples

are the selection criteria for influencer or the amount of tweets analyzed in specific cases. Most of these assumptions are made because of limited resources.

Section discusses three things: the representativeness of public opinion on palm oil on Twitter versus public opinion on palm oil in real life; the general characteristics defined in chapter 7 to see if they say something about public opinion on Twitter in general; and it briefly mentions limitations of the Twitter API. It concerns steps F and G of the framework in figure 9.1 and therewith looks at *what these measurements say*. Over 100.000 tweets are gathered over 50 days on the topic of palm oil. All topics on Twitter are unique and there is no topic that is the perfect example. The topic of palm oil included human and bot accounts, legitimate and spam accounts, has many subtopics and included tweets going viral. Therefore, I believe that the topic of palm oil and the generalized characteristics are representative for public opinion on Twitter.The last four sections are not part of the framework in figure 9.1 but can be seen as overarching and reflecting.

Section 9.4 describes the limitations to Twitter on four different levels: the limitations due to the anonymity possible on Twitter, possible consequences of Twitters privacy policy, the possible responsibility to take for Twitter and the in-or exclusiveness of Twitter. The question if Twitter can be used is discussed before so this section merely discusses what the limitations are given that Twitter is used as a medium.

Section 9.5 discusses the practical implementation of the script to analyze the tweets. Even though the script is easy to use, some enhancements can be done to improve the user-friendliness of the scripts. Before these enhancements are done, it needs to be investigated if policymakers are interested in using this script and policymakers need to be made aware of how public opinion is shaped on social media.

Section 9.6 describes possible directions of further research as based on the validation and reflection. Research is recommended in three ways. The first direction is towards philosophers and consists mainly of suggestion for more research towards the public sphere in the 21st century. The second direction is towards data scientists and addresses certain limitations in the way the data is gathered and analyzed. The third direction is towards policy research to address the challenges for policymakers when using social media data.

Section 9.7 contains my personal reflection, firstly on the future of social media and Twitter and secondly on the process of this thesis. The internet has turned 29 this year and Twitter has turned 12. Both are not mature technologies yet. I turned 25 this year and still have a lot to learn; I also learned a lot over the last 6 months of writing this thesis.

9.1. Limitations of this research: what is and what is not being measured?

The main line of reasoning is based on the Habermasian public sphere. The theory of the Habermasian public sphere is developed based on 18th century Europe and this section discusses this theory in the light of the 21st century. Twitter plays a major role in this thesis but does not go uncriticized. Section 9.1.1 looks at the analysis of Twitter through the lens of the Habermasian public sphere. Section 9.1.2 carries out construct validity on the three main constructs: *Twitter is a incarnation of the Habermasian public sphere*, *Twitter causes a fragmentation of the public* and *Twitter creates space for issue publics*. Section 9.3 looks at how representative this case study is on Twitter for public opinion.

9.1.1. Twitter as a public sphere

Before exploring the advantages and disadvantages of looking at Twitter through the lens of the Habermasian public sphere, it is important to note that the *The Structural Transformation of the Public Sphere* does not go uncriticized. Section 4.1.1 elaborated on different critiques of the Habermasian public sphere. Some critiques have led to the different concepts, so in that, they are taken into account. For example is the critique by Butsch (2016), who argued that the public is too diverse to be caught in one public sphere. Therefore, concept and construct 2 are developed: Twitter causes a fragmentation of the public. Another outspoken critic of *The Structural Transformation of the Public Sphere* is Nancy Fraser (Calhoun, 1992). Many of her critiques on the Habermasian public sphere are very applicable to Twitter. Firstly, she

argued how the ordinary people is only able to watch the ”events unfolding on this ”virtual stage of mediated communication” (Bruns and Highfield, 2016, 56). This means that the ordinary people do not have a say and therefore, the public sphere is not a very egalitarian place. Twitter can also be argued to not be very egalitarian, since most users are merely seeing event unfolding. Secondly, Fraser argued how not everyone was welcome in the coffeehouse (as an illustration for the public sphere). Women were excluded. This is not very different from Twitter, where also more than half of the world population is excluded due to a lack of internet access, new media literacy and cultural capital. To conclude this part and continue with the differences between the 18th and 21st century, the Habermasian public sphere *including* all its critiques has been a good lens to look at public opinion on Twitter.

There are many differences to be made when comparing the 18th with the 21st century. The main differences analyzed here are based on the internet and are divided into the age before the internet (18th century) and the age with the internet (21st century). Arguments to be made in favor of still applying the theory of the Habermasian public sphere are based on the idea that times have always changed. Argued can be how capitalism has created mass commercial culture that leaves no space for the public sphere (Papacharissi, 2002), so it is not the internet that drastically changes the public sphere. Additionally, even though everyone (with new media literacy, cultural capital and internet access) can participate in online debate, the discourse is still dominated by a few (Papacharissi, 2002).

Arguments to be made against the application of the Habermasian public sphere because the times are radically different are based on the believe that the internet is drastically different from everything we have ever seen before.

> “Such a hope represents exactly what Marshall McLuhan used to call ’rear-view mirror’ thinking: the assumption that a new medium is merely an extension or amplification of an older one; that an automobile, for example, is only a fast horse, or an electric light a powerful candle. To make such a mistake in the matter at hand is to misconstrue entirely how television redefines the meaning of public dis- course. Television does not extend or amplify literate culture. It attacks it” (Postman, 2006, 83-84)

As the television ’attacks’ literature culture, imagine what social media does. Online communication is different and argued to not be based on a rational, focused conversation but more about venting emotion and expressing hasty emotions (Papacharissi, 2002). It can be argued that behaviour rooted in reason, the most important assumption for Habermas, is simply not the case on the internet. The internet is a place where one can shout louder, but that does not mean that others are listening nor you are making a difference. Offline, popular newspapers allow on their front page multiple stakeholders to express their reasons and arguments and therewith shows publics behind the different interpretations. In the age of the internet, in the words of Morozov, “Google and Facebook unilaterally decide what counts both as issues of importance and as the right ways to campaign on them” (Morozov, 2013, 476). Even though Morozovs’ statement is extreme, there is no doubt that the power in the world is drastically shifting.

There is something to say for both sides. I believe that the thoroughly researched theory on the public sphere works well as a lens to look at Twitter. That said, times *have drastically changed*. I think that the online public sphere is drastically different from the traditional public sphere. And to go one step further, the online public sphere *strongly influences* the offline public sphere when issues debated online spill over in the real world. The reach of Twitter is bigger and crossing more borders than any other possible incarnation of the public sphere that I know of. Therefore, I believe the Habermasian theory of the public sphere works well as a lens to analyze Twitter but the different age we are living in should be kept in mind.

9.1.2. Construct validity

The term ’construct validity’ is derived from psychology, where psychological constructs are essentially unobservable (Cronbach and Meehl, 1955). Construct validity “concerns the validity with which constructs are linked to observations and measurements” (Cherryholmes, 1988, 421). This means- in this thesis- the construct validity is looking at what is being measured in the defined constructs, and if that is what I wanted to measure. All three constructs

will be validated separately by looking at what they actually measure. If these constructs give an idea of public opinion is discussed before going into all constructs separately. They are, just like in psychology, essentially unobservable. The difficulty in this thesis is the ambiguity and disagreement in literature on the different keywords. The first big assumption, or simplification, is to use the Habermasian public sphere as place where something approaching public opinion is happening.

Construct 1: Twitter is a incarnation of the Habermasian public sphere

The first construct is divided into four characteristics. This division is in itself already based on assumptions on the simplification of the Habermasian public sphere to try to turn something very conceptual into something more concrete:

1. Social inequalities are bracketed
2. Public Sphere is a new centre of institutional authority
3. Access is granted to all citizens
4. Those forming public opinion are presupposed reasonable

These characteristics- even though part of the construct- are still ambiguous and explained in many different ways in scientific literature. To turn them into measurable characteristics, simplifications are made. The first characteristic is only expressed based on the network and influencers. These are the only things directly measured and do not (directly) express the actual public status, since public status is not something that is directly observable. Different choices might have lead to different outcomes, but due to the limitations of the Twitter API and what people put on Twitter, is it not possible to measure users' public status directly.

Many different interpretations are possible for what is meant by the second characteristic, a new centre of institutional authority. This thesis bases it on interviews and calls for action based on the data. These two measures create a bit of insight in possible centres of institutional authority but are clearly not an opaque approach. Further research is recommend to cover the second characteristic.

Groups of actors are mapped to operationalize the third characteristics, access is granted to all citizens. This measure looks at the different groups of actors based on the most common users in the database and not necessarily if access is granted to all citizens. It can be argued to be a good measure to show different actors with access, but the biggest limitation is that you do not see what is not there. So we do not see who access is not granted to. A measure taken to overcome this is to look at who is all excluded from Twitter based on their lack of new media literacy, cultural capital and internet access.

The most frequently retweeted tweets are analyzed manually to see if the users are acting reasonable. Reason is solely based on their argument, use of media and polarity. More measures could be added, but the primary question that to be asked is what it means to act reasonable. There is no unanimous agreement on this in the literature and therefore it is difficult to come up with different measures that would measure reasonability better. More tweets could be analyzed to have a bigger sample, but this does not solve the definition question.

Construct 2: Twitter causes a fragmentation of the public

The second construct and divided into two characteristics. This, in itself, is already based on assumptions on the simplification of the Habermasian public sphere, to try to turn something very conceptual into something more concrete:

1. The public on Twitter is fragmentized
2. Twitter causes this fragmentation

The first characteristics to check if the public is fragmentized is based on all tweets. Hashtags and words are counted and compared, additional topics are found and the overlap between the different topics is shown. These measure quite directly show fragmentation. The second characteristic is less straight forward and is based on three measures by Barber (1998): bias towards image over text, partiality to raw data rather than informed knowledge and telling snippets of the story rather than the whole story. These separate measures do

not individually lead up to the causality of fragmentation caused by Twitter but can be seen as an approximation.

Construct 3: Twitter creates space for issue publics
The third construct and divided into two characteristics. This, in itself, is already based on assumptions on the simplification of the Habermasian public sphere, to try to turn something very conceptual into something more concrete:

1. Twitter functionalities (retweets and hashtags) create space for issue publics
2. Certain issues push others away in a time where our attention is scarce

The first characteristic is specifically based on two Twitter functionalities that seems the most important functionalities in creating issue publics but others can be argued to also play an important role. These two measure direct forms of attention: indirect forms of attention, like looking at or reading a tweet, are not measured, while these can play a very important role. Nothing has been done to overcome this imperfection, since Twitter does not give any information on other forms of engagement for their tweets than the directly observable ones.

The second characteristic looks at a time series of the different peaks to see if these push other issues or tweets away or amplify the debate. This is a specific interpretation of the attention economy. The biggest limitation is that only the tweets gathered on the topic of palm oil can be compared, and not different tweets on different topics. This might be where the real attention fight is happening and tweets on all topics have to be investigated to see if that is the case.

9.2. Critical assumptions

This section describes the assumptions made in the middle of the hourglass (see figure 3.1), on the most specific level. They follow the constructs and therefore this section is build up of the different constructs. Section 9.2.1 describes construct 1, *Twitter is an incarnation of the Habermasian public sphere*, section 9.2.2 describes construct 2, *Twitter causes a fragmentation of the public* and section 9.2.3 describes construct 3, *Twitter creates space for issue publics*.

9.2.1. Construct 1: Twitter is an incarnation of the Habermasian public sphere

This section looks at the operationalization of the four characteristics on the lowest level. Section 9.1.2 looks at the higher level definitions.

1. **Social inequalities are bracketed**

   Public status is defined based on the network and the influencers. The network is only based on the giant reply network. In network theory, the giant component is “a connected graph that includes a large fraction of all documents in the data set that can be connected (Bollobás and Riordan, 2004)” (Small, 2009, 332). The choice for the giant component is made for practicality reasons: the whole network instead of the giant component would be too big to show and therefore difficult to generate insights from. The giant component is generally accepted, but further research could look at the whole network. Additionally, other types of networks such as followers- or retweet networks can be looked into.

   Influencers are based on:

   (a) Central role in reply network (see figure 6.4)
   (b) Mentioned the most by other users in this database
   (c) Retweeted the most in this database

   There are multiple others ways to define influencers and there is no single agreement in literature. By combining these three measures and the visible overlap, it seems like these are good measures. More users per measure are recommended for further research, since there are now only 4 analyzed.

2. **Public Sphere is a new centre of institutional authority**

   A centre of institutional authority is defined based on interviews with policymakers and the top seven most frequently retweeted tweets. Only two interviews with policymakers have been conducted and since policymakers all make very different policies, these cannot be taken as the blueprint. They do, however, create insight in how *some* policymakers use the data.

   A limitation on the analysis of the top seven most frequently retweeted tweets is the low amount. To partly overcome this limitation, an additional measure is conducted to strengthen the conclusion that there are calls for action. This measure is based on the frequency of the words boycott, stop and petition as direct calls for action in all the tweets.

3. **Access is granted to all citizens**

   Access is checked for different groups based on:

   (a) 20 most mentioned users in this database
   (b) 20 most retweeted users in this database
   (c) 20 users with the most tweets in this database

   The combination of these three groups leads to 60 users from which there are 19 overlapping, so in total 41 unique users. This is a small sample of the total amount of users in the database and it is therefore difficult to generalize the percentages mentioned in figure 6.9 to the whole data set or the whole of Twitter. After gathering the different users they are divided into six categories:

   (a) Business (for profit, like Nestle)
   (b) Government (like the Malaysian Ministry of Primary Industries)
   (c) Individuals
   (d) Media
   (e) NGO
   (f) nonprofit

   Other categories could have been chosen. NGO and nonprofit, for example, could have been added together. The reason they are separated is because the roles in this debate on palm oil are very different. NGOs take a much more activist role (like Greenpeace), whereas nonprofit take a mediating role (like RSPO).

4. **Those forming public opinion are presupposed reasonable** This characteristic is, like characteristics 2, based on the seven most frequently retweeted tweets and thus comes with the same limitations. Additionally, the automated polarity package is used. This package is not very transparent and therefore it is difficult to value the results of this package.

9.2.2. Construct 2: Twitter causes a fragmentation of the public

This section looks at the operationalization of the two characteristics on the lowest level. Section 9.1.2 looks at the higher level definitions.

1. **The public on Twitter is fragementized**

   For the conducted word frequency and hashtag count, all tweets are taken into consideration. Only 37,5% of all tweets contains a hashtag so it is questionable how telling the hashtag count is. Hashtags are frequently used so it can be argued that, even though the percentage of use is not very high, they are telling. The overlap between the topics is based on my own interpretation and it can be recommended for further research to validate these topics with someone who looks at the same data with a fresh eye, to see if they draw the same conclusions.

2. **Twitter causes this fragmentation**

   Besides the definition-assumptions, this analysis is based on the most frequently retweeted tweets and therefore comes with the same limitations as mentioned in section 9.2.1.

9.2.3. Construct 3: Twitter creates space for issue publics

This section looks at the operationalization of the two characteristics on the lowest level. Section 9.1.2 looks at the higher level definitions.

1. **Twitter functionalities (retweets and hashtags) create space for issue publics**

   This characteristic is based on three specific peaks. More peaks are visible but due to limited resources, only these three are analyzed. Further research could analyze more peaks and look further into the existence of the third peak (see figure 6.13), since this peak is not completely defined yet.

2. **Certain issues push others away in a time where our attention is scarce**

   This characteristic is based on two (out of three) peaks that are defined in the previous section. Even though these two show clear results, the quantity is not very large. For further research it is recommended to look into different peaks.

9.3. Representiveness and generalizability of the case study

This section firstly looks into the representativeness of public opinion on palm oil on Twitter versus public opinion on palm oil in real life. Secondly, it looks into the general characteristics defined in chapter 7 to see if they say something about public opinion on Twitter in general. Thirdly, it briefly mentions limitations of the Twitter API. It is very difficult to find out if common opinions on Twitter on palm oil are representative for public opinion on palm oil in real life. A representative survey could be conducted, but Twitter does not stop at geographical borders and it would be very challenging to send a survey to all countries where people have tweeted from. Based on this thesis, it can be argued that because parties like the BBC are involved in expressing and spreading opinions, this might influence people outside Twitter. It is very difficult to define what came first in this specific case: general attention on Twitter for palm oil, or the reportage by BBC on palm oil. This traditional chicken and egg problem does not lend itself to one cause and one consequence, but it can be concluded that there is a correlation (undefined how large) to what happens on Twitter and what happens in the real world.

General characteristics

The general characteristics derived from the palm oil debate can show insight in general public opinion on Twitter. More than 100.000 tweets and 50 days is an amount and duration to derive general insights. All topics on Twitter are different and there is no topic that is the perfect example. The topic of palm oil included human and bot accounts, legitimate and spam accounts, has many subtopics and included tweets going viral. Additionally, influential users defining the debate on palm oil are not only tweeting on palm oil but on many different topics. Therefore, I believe that this topic and the generalized characteristics are representative for public opinion on Twitter.

Limitations of the Twitter API

All tweets are collected using the Twitter API. API stands for Application Programming Interface, and the Twitter API used in this research searches against a sampling of recent Tweets published in the past 7 days (Twitter, 2018c). They mention further how the standard search API is focused on relevance and not completeness. The main drawback is the lack of documentation on what and how much users get and therewith if the sampled data can be used as a valid representation of all Twitter activity (Morstatter et al., 2013). When you want to use all tweets in further research, Twitter recommends paid version of the API. What the effect of this sampling is unknown. It can be argued that to show insight in the debate, not more tweets are needed. But since we don't know is all in the tweets we don't have, it is difficult to guarantee the effect of sampling is not major.

9.4. Limitations of Twitter

This section describes the limitations of Twitter in different areas. Section 9.4.1 describes the limitations due to the anonymity possible on Twitter. Section 9.4.2 describes possible

consequences of Twitters policy. Section 9.4.3 describes the possible responsibility to take for Twitter. Section 9.4.4 describes the in-or exclusiveness of Twitter. Before diving into the different section, a note on McLuhan needs to be made.

McLuhan, mentioned earlier in this thesis regarding hot and cold media, has also written extensively on the principle that the medium is the message. This means that the medium influences how the message is perceived (Blackman, 2017). In the case of Twitter, it does not only influence the perception but also what is communicated. Twitter, together with other internet giants like Google and Facebook, can be argued to be the gatekeeper of the internet. They control what ideas and opinions are being shared. Due to their nontransparent algorithms it is very difficult to know what influence this exactly has on public opinion.

9.4.1. Anonymity of Twitter users

In the real world anonymity is not done. On the internet anonymity is accepted. Some users on Twitter are not posting under their real name. These anonymous/pseudonym accounts are very often legitimate users (Oltulu et al., 2018). Twitter asks for an email address and a phone number when creating an account, both which can easily be created (Bright, 2017). Due to the possibility to tweet to anyone, immediate feedback is possible, challenging users. But when comparing Twitter to the 18th century coffeehouses, it is much easier to disagree on the internet than to walk into a coffeehouses where the criticism is linked directly to you as a person (and not to your pseudonym). Argued is how that makes online communication not based on a rational, focused conversation but more about venting emotion and expressing hasty emotion (Papacharissi, 2002). Papacharissi argues also how anonymity and the absence of face-to-face interaction limits the possible to asses the impact of our words. It can give an empowering feeling but might have a limited power to actually make impact.

The anonymous possibilities on Twitter also have given rise to bots. Any human can make multiple (automated) accounts. Twitter has deleted 70 million fake accounts in May and June to decrease the flow of misinformation on the platform (Timberg, 2018). It is unclear what the effects of these are on public opinion, but it can be argued that it does not work in favor of the legitimacy of the platform when an unknown but decent percentage of the accounts are fake.

9.4.2. Privacy on Twitter

Governmental organizations are using information from social networking sites to monitor citizens. The municipality of Amsterdam recently gained attention by analyzing the friend lists on Facebook of 126 so called 'nuisance generating' youth. The Dutch Data Protection Authority judged this analysis went too far and infringes on their privacy (van Lonkhuyzen and Stokmans, 2018). Accenture urges law enforcement agencies to "Tap the Power of Social Media to Drive Better Policing Outcomes" (Accenture, 2018). A start-up from Virginia describes its system as "a social media surveillance solution providing real-time monitoring of Twitter, Facebook, Google groups, and many other communities where users express themselves freely" (Morozov, 2013, 456). There are multiple comments about this start-up. Obviously, is is not clear what is being monitored and where it stops. It is also not clear what is done with this information.

Governmental organizations do not easily have access to private communications and less obvious actions, like clicking links. The big tech companies like Google, Twitter and Facebook surely do and therewith their power for mass surveillance is much larger. The recent EU General Data Protection Regulation (GDPR) that came into force in May 2018 enforces privacy by design and it is not yet clear if this limits their power. Privacy by design (PbD) is explained by (Romanou, 2018, 100) "as the implementation of several privacy principles directly into the design specifications of the technological systems, in a way that privacy rules will be embedded in the operation and management of the processing of the data." This means that PbD approaches privacy reversed from the way it has been approached, namely embedded in the technology to process data. The privacy of the user is therewith safeguarded, without limiting the capacities of technologies. PbD enforced in the GDPR is, in my opinion, a big step in the right direction.

9.4.3. Responsibility of Twitter

There is no consensus on what responsibility Twitter should take on all different aspects. This section will discuss their responsibility regarding free speech, fake accounts and topics. Regarding free speech, opinions are very divided. Twitter started off as a platform for free speech, but it is unclear if hate speech is free speech. Additionally, the laws about free speech differ per country. Twitter mentions the following on their website regarding freedom of expression:

> "Freedom of expression means little if voices are silenced because people are afraid to speak up. We do not tolerate behavior that harasses, intimidates, or uses fear to silence another person's voice" (Twitter, 2018b).

This means that Twitter can decide who to silence. According to some, it becomes dangerous when social media platforms have the right to dictate what speech is or is not hateful (VPRO, 2018). These decisions made by twitter possibly lead to the silencing of important conversations on certain platforms, or a possible move to other platforms with less regulations, moderation and control such as 4chan. Twitter is not very transparent about what behaviour they have not tolerated and therefore it is difficult to judge what exactly they are doing regarding free speech.

Fake Accounts

Regarding fake accounts, Twitter is taking action by deleting many fake accounts (Timberg, 2018). But there are many fake accounts left and these influence Twitter users in multiple ways. Firstly, evidence by MIT shows how social bots play a major role in fake news (MIT, 2017). Secondly, many users 'buy' influence by buying followers, for which fake accounts are created (Confessore and Dance, 2018). There is no consensus on what responsibility Twitter should take with regards to fake accounts. There are computational challenges like the identification of fake accounts and governance challenges like the definition of fake accounts and the policy on what to do with them.

Twitter's choice for topics

Regarding topics, Morozov and other accuse Twitter of censorship by defining what topics are trending. He additionally questions if we want to move towards a system that hijacks our attention through nudges:

> "If citizens come to care about Bosnia or Rwanda or Syria not because they believe in the importance of humanitarian intervention or deliberately seek out news about those lands but because some combination of nudges and algorithms has made such caring all but inevitable, this seems like a tacit acknowledgment that deliberation and morality no longer command any respect in our political life and that now it all boils down to Skinnerian experimentation as to what combination of incentives—not arguments!—yields the desired action" (Morozov, 2013, 477)

The biggest problem, in my opinion, is that we do now know what the underlying nudges and algorithms are- let alone the combination that is being used. Whereas Habermas' focused was on the rational-critical debate, Twitter currently only focuses on the combination of incentives for the desired action. They make smart use of the 'curated' timeline, meaning Twitter decides what one sees based on their nontransparent algorithms (Romano, 2018). It is possible to change ones 'curated' timeline to a chronological timeline, but this process is not straight forward.

9.4.4. Inclusiveness of Twitter

This thesis has mentioned from the very beginning how only people with internet access, new media literacy and enough cultural capital can have access to Twitter. 45% of the world is already excluded by a lack of internet access (Group, 2018) and it is difficult to say how many others are excluded by a lack of new media literacy or cultural capital. Less than 5% of the world population is actually on Twitter (Statista, 2018a). Additionally, as mentioned

before, not everyone who speaks out gets heard. This is confirmed by the interviews with policymakers who say to only use Twitter for the opinions of certain users like politicians, unions and journalists.

9.5. Practical implementation

Before the script can be used, we need to have the discussion if public opinion expressed on social media is public opinion we want to use in our policy making. It can be argued how it is too dominated by certain organization, fake accounts or by the influence of Twitter itself. The medium Twitter is the message. The medium influences largely how the message is perceived, and Twitter definitely has a a seat at the table- no one know how big this seat is, but it is there. When you want to use the script after these considerations, below are some comments.

For the script used to analyze the tweets on palm oil to be used on different topics some enhancements can be done to improve the user-friendliness of the script. Even though it is relatively quick and easy to use, the policymaker or other person wanting to use the script needs some knowledge of Python. When Python knowledge is present, it can easily create insight in the different users, influencers argumentation and topics.

The script can be turned into a more user-friendly dashboard where you only have to fill in the keywords. In the current situation, the scripts that gathers the tweets needs to be uploaded to a server. One of the two policymakers interviewed has expressed to not be interested in an easy script to measure public opinion on a certain topic because he does not believe the public on Twitter is representative enough for the Dutch public. Due to limited resources no more time is put in creating a user-friendly dashboard but this can easily be added later.

9.6. Future Research

No research is perfect and neither is this. This work should be seen as a starting point for further research. The suggestions in this section are based on the discussion earlier in this chapter and give direction in three ways: towards philosopher, towards data scientist and on policy research. Philosophers can take different angles for further research. Firstly, another line of reasoning can be taken to see if the generalized characteristics are drastically different from the ones defined now. A suggested direction is John Dewey. Dewey has been used in this research to elaborate on issue publics, but Dewey's 1927 *The Public and its Problems* can take up a more central role in defining the forming of the public in another research. Dewey wrote his book as an answer to an anti-democratic attitude prevalent among some political commentators at the time. These commentators argued that the increased complexity of society and the rapid rate of technological progression made politics too complicated for the ordinary public; modern government should instead rely on the rule of experts in a technocracy. The public, it was though, was generally too uniformed or uninterested to be the basis of decision making. Against this disbelieve in democracy, Dewey rethought the concept of the public by placing the issue central. The public is that which forms itself around an issue, and consists of anyone who is effected by or concerned with the issue. This formation of public is, for Dewey, the basis for democracy. *The Public and its Problems* can be applied to Twitter in the same way this thesis has applied the Habermasian public sphere on the framework guiding this thesis (see among others figure 9.1). Dewey's theory can be combined with actual tweets to see what Twitter does to democracy, and therewith defines characteristics of Twitter. Secondly, when staying with Habermas, different concepts and constructs can be defined, with different characteristics. There are many possibilities here. One suggestion is based on the transformation of the public sphere. This thesis did not include anything on the actual *transformation*, which was argued by Habermas to transform from a bourgeois public sphere to a sphere dominated by capital-driven mass media.

I have six concrete recommendations for data scientists. Firstly and most importantly, use the version of the Twitter API that gathers all tweets and not just a sample. Secondly, look further into the tweets that Twitter deletes soon after they are posted to get an idea of the level of censorship. Thirdly, all my analysis can be conducted on more tweets, especially

when I have only looked at the most frequently retweeted tweets. Fourthly, the network is in this thesis only analyzed by way of the giant component of the reply network. Recommended is to look at the whole reply network and at different networks such as followers- or retweet networks to create better insight in who is all participating. Fifthly, the increasing knowledge of machine learning techniques creates possibilities to apply more advanced topic models on the tweets, to see if other topics pop up than the ones defined now. Lastly, not all peaks in tweets gathered over 50 days have a single clear cause. Further research can analyze into these peaks deeper to see what combination of factors causes peaks and what we can learn from that. Regarding the peaks, it is also recommended to cross check different topics. In the palm oil debate, topics do not push each other away, but this might be different across Twitter as a whole.

There is a lot of work to do for policy researchers. Social media is growing as a larger share of the world population is on social media. Policies are not up to date and based on the idea that of the early 2000s that the internet is a sovereign body that cannot be regulated. As can be seen on the case of Twitter, it is unclear if and to what extent Twitter is censoring. The previous above mentioned recommendation for data scientists, to get insight in Twitters censorship, can be a starting point. To conclude, these three lines of recommendations -for philosopher, data scientists and policy researchers- can be used to guide further research.

9.7. A reflection on the future of Twitter and the process of this thesis

This section contains my personal reflection, firstly on the future of social media and Twitter (section 9.7.1) and secondly on the process of this thesis (section 9.7.2). Argued is how the internet is broken and the future of the future direction of internet is not certain. The reflection of the process describes the key findings about myself over the last six months. Some characteristics of me, are interestingly enough, similar to some characteristics of Twitter.

9.7.1. The future of Twitter and the internet

The internet has turned 29 this year and for the first time in its life, more than half the world population is connected (Berners-Lee, 2018). But argued is how the internet is broken (Stikker, 2018). As the founder of the internet, Tim Berners-Lee, states:

> "The threats to the web today are real – from misinformation and questionable political advertising to a loss of control over our personal data. But I remain committed to making sure the web is a free, open, creative space – for everyone" (Berners-Lee, 2018)

Before the web becomes a free, open and creative space again, things will have to change. The power is currently in the hands of a few dominant platforms who function as gatekeepers. They decide what is being discussed. According to Jeroen van den Hoven, professor of Ethics and Technology at Delft University of Technology, the idea to decentralize power fits in our new political ideology. People are starting to understand we cannot continue going the way we go, with the few gatekeepers who sell our personal data as their business model (Verhagen, 2018). Times have changed drastically and the law has stayed behind. Where traditional media like newspapers have more than 50% market share, they are reported to the competition authority. Google, for example, has over 90% market share (Desjardins, 2018) and their market share is not acted upon by the competition authority. Google, and others, have incredible amounts of data about all of us and are turning this big data into big money.

Orwell's 1984, written in 1949. warned for Big Brother. The Big Brother of that time was the government as an authoritarian regime, carrying out the 'thought police' against individualism. To protect the beginning of the internet from the state, it was forbidden for Big Brother. Now, almost thirty years later, these companies can be best describes as Big Business and Big Brother in one (Kuitenbrouwer, 2018). New initiatives to counter Big Business and Big Brother in one are emerging, based on decentralized power and giving back the control over our personal data. The big question is if initiatives like these are going to give the power back to the people, or if the current gatekeepers keep their power. Another possibility

is a state controlled medium such as the Chinese alternative for Twitter called Weibo. State controlled mediums raise different challenges, not leading to Big Brother and Big Business in one, but just as Orwell feared, Big Brother.

Whatever the future of the internet may bring, the ethical and governance choices are going to be much more challenging than the technology behind them. The internet in its current form, as an incarnation of the Habermasian public sphere- open, democratic and egalitarian- has failed. But it is not too late to repair the internet. Let's make the web a better place.

9.7.2. The process of this thesis

This process has been a true adventure. Everyone around me warned me how it was going to be the most terrible period out of my life. But I have truly enjoyed the ride. Putting everything I have learned into practice and more has been challenging and rewarding. This section firstly describes my challenges during this thesis and how I have learned from them. Secondly, it describes my strengths, how these have been useful and how I have improved them.

Challenges

Three main challenges are identified: my lack of nuance, my directness and my enthusiasm to start new things. Firstly with regards to my lack of nuance, I recognize myself in the way Twitter works. I am used to saying what it is and not sugarcoat anything and neither are tweets sugarcoating (since they only have 280 characters). This has been challenging sometimes, in making blunt statements, that are not fully thought out nor explained. The world is not as simple as I would like to present it sometimes. I have learned that sugarcoating sometimes is better and makes your message stronger. Additionally, I have learned how some simplifications are too simplified and simply not reflect the truth anymore.

Secondly, my directness. I tend to not waste any words on what I am going to do or what I have done. By doing so, it is difficult for a reader to understand my (complex) train of thoughts sometimes. This takes away power of statements and texts. A way to deal with this challenge has been to try to explain statements without just dropping them like little bombs waiting to explode.

Thirdly, my enthusiasm to start new things. The beginning of my thesis felt like running around on a playground. I could do many things a little bit without having to finish up anything. This has changed over the last six months, since I had many loose ends. I also realize that many of the projects I have started take a lot of work to finish, much more than I thought when I started. The only way to deal with this challenge is work hard to finish up what I have started, and make decisions on what not to finish and leave out. Additionally, I have not started new analyses later on in the thesis since there was too much to finish first. These three challenges have kept the process interesting. I never had the feeling it was too much and always had the feeling I was growing a lot as a professional.

Strengths

Three main strengths are identified: my ability to focus, my ability to structure and my broad interest. Firstly, my focus has been the reason I have achieved so much. Even though I enjoy starting up different things more than finishing them, you have got to do what you have got to do and therefore, you just do it. This has helped me to get a lot done over the last six months. Secondly, I love structure. I have been considering putting it at my challenges since I always thought I was more structured than I have been in this thesis. I love to write in bullet points or to count in a text. I have realized, however, that with very complex philosophical content, it is difficult to simplify things fully even though I wish the whole world could be broken down in clear bullet points. Since this is not the case, I think that, within reason, my crave for structure has been of great help. Thirdly, my broad interest is partly reflected in the last weakness (my enthusiasm to start new things) but is more encompassing. Not many people enjoy philosophy and data analysis as much as I do- especially in combination. I have loved spending days in Python and have loved spending days in analyzing Habermas' *Structural Transformation of the Public Sphere*. These three strengths have made the process

a true pleasure. I am very excited for what the future has to bring and am definitely going to try to create a future in which I can combine more philosophers with more data!

Bibliography

Accenture (2018). *Tap the Power of Social Media to Drive Better Policing Outcomes*. https://www.accenture.com/us-en/insight-tap-power-social-media-drive-better-policing-outcomes [Accessed: October 2, 2018].

Allport, G. (1935). *Handbook of Social Psychology*. Clark University Press, Worcester, MA.

Alp, Z. Z. and Şule Gündüz Öğüdücü (2018). Identifying topical influencers on twitter based on user behavior and network topology. *Knowledge-Based Systems*, 141:211 – 221.

Anderson, B. (1991). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso, London.

Andrea, C. and Fedra, N. (2016). The "social side" of public policy: Monitoring online public opinion and its mobilization during the policy cycle. *Policy & Internet*, 8(2):131–147.

Atchley, P. and Lane, S. (2014). Chapter four - cognition in the attention economy. volume 61 of *Psychology of Learning and Motivation*, pages 133 – 177. Academic Press.

Bailey, A. A., Bonifield, C. M., and Arias, A. (2018). Social media use by young latin american consumers: An exploration. *Journal of Retailing and Consumer Services*, 43:10 – 19.

Ballard, T. (2012). 1 - what does the internet have to do with my library? In Ballard, T., editor, *Google This!*, Chandos Information Professional Series, pages 1 – 10. Chandos Publishing.

Barabas, J. (2011). Degrees of democracy: Politics, public opinion, and policy. *Public Opinion Quarterly*, 75(1):192–195.

Barber, B. (1998). *Which Technology and Which Democracy?* http://web.mit.edu/m-i-t/articles/barber.html [Accessed: August 20, 2018].

Berners-Lee, T. (2018). *The web can be weaponised – and we can't count on big tech to stop it*. https://www.theguardian.com/commentisfree/2018/mar/12/tim-berners-lee-web-weapon-regulation-open-letter [Accessed: August 20, 2018].

Blackman, C. (2017). Telecommunications policy: The medium is the message. *Telecommunications Policy*, 41(10):847 – 852. Celebrating 40 Years of Telecommunications Policy – A Retrospective and Prospective View.

Bollobás, B. and Riordan, O. (2004). The phase transition and connectedness in uniformly grown random graphs. In Leonardi, S., editor, *Algorithms and Models for the Web-Graph*, pages 1–18, Berlin, Heidelberg. Springer Berlin Heidelberg.

BPDP (2018). *President to Sign Regulation on Expanding Mandatory B20 Soon*. https://www.bpdp.or.id/en/energi/presiden-segera-tandatangani-perpres-b20/ [Accessed: September 11, 2018].

Braatz, K. (2011). *Friedrich Nietzsche - Eine Studie zur Theorie der Öffentlichen Meinung*. Monographien und Texte zur Nietzsche-Forschung. De Gruyter.

Bright (2017). *Zo gebruik je Twitter anoniem*. https://www.bright.nl/tips/artikel/3934116/zo-gebruik-je-twitter-anoniem [Accessed: October 8, 2018].

Broekhans, B. (2018). *Understanding international grand challenges[Powerpoint slides]*. https://brightspace.tudelft.nl/d2l/le/content/36247/viewContent/551067/View [Accessed: July 18, 2018].

Bruns, A. and Burgess, J. E. (2011). The use of twitter hashtags in the formation of ad hoc publics. In *6th European Consortium for Political Research General Conference*, University of Iceland, Reykjavik.

Bruns, A., Enli, G., Skogerbo, E., Larsson, A., and Christensen, C. (2015). *The Routledge Companion to Social Media and Politics*. Routledge Media and Cultural Studies Companions. Taylor & Francis.

Bruns, A. and Highfield, T. (2016). Is habermas on twitter? social media and the public sphere. In Bruns, A., Enli, G., Skogerbø, E., Larsson, A. O., and Christensen, C., editors, *The Routledge Companion to Social Media and Politics*, pages 56–73. Routledge, New York.

Butsch, R. (2016). *Media and Public Spheres*. Palgrave Macmillan UK.

Calder, B. J., Phillips, L. W., and Tybout, A. M. (1982). The concept of external validity. *Journal of Consumer Research*, 9(3):240–244.

Calhoun, C. (1992). *Habermas and the Public Sphere*. Studies in Contemporary German. Paperbackshop UK Import.

Carter, Jr., R. E. (1956). Personal influence: The part played by people in the flow of mass communications. by elihu katz and paul f. lazarsfeld. glencoe, illinois: The free press, 1955. 420 pp. *Social Forces*, 34(4):383.

Ceron, A. (2015). Internet, news, and political trust: The difference between social media and online media outlets. *Journal of Computer-Mediated Communication*, 20(5):487–503.

Ceron, A. and Negri, F. (2015). Public policy and social media: How sentiment analysis can support policy-makers across the policy cycle. *Rivista Italiana di Politiche Pubbliche, Rivista quadrimestrale*, (3/2015):309–338.

Chandler, D. and Munday, R. (2011). Hot and cool media.

Chen, L., Wu, X., and Li, M. (2018). Formation and fragmentation within a networked public sphere: Social media debates on traditional chinese medicine. *Telematics and Informatics*.

Cherryholmes, C. H. (1988). Construct validity and the discourses of research. *American Journal of Education*, 96(3):421–457.

Chong, M. (2016). Sentiment analysis and topic extraction of the twitter network of #prayforparis. *Proceedings of the Association for Information Science and Technology*, 53(1):1–4.

Colleoni, E., Rozza, A., and Arvidsson, A. (2014). Echo chamber or public sphere? predicting political orientation and measuring political homophily in twitter using big data. *Journal of Communication*, 64(2):317–332.

Confessore, N. and Dance, G. (2018). *Battling Fake Accounts, Twitter to Slash Millions of Followers*. https://www.nytimes.com/2018/07/11/technology/twitter-fake-followers.html [Accessed: August 20, 2018].

Cronbach and Meehl (1955). Construct validity in psychological test. 52:281–302.

Cunningham, S. (2001). Popular media as public 'sphericules' for diasporic communities. *International Journal of Cultural Studies*, 4(2):131–147.

Cvetojevic, S. and Hochmair, H. H. (2018). Analyzing the spread of tweets in response to paris attacks. *Computers, Environment and Urban Systems*, 71:14 – 26.

Dahlberg, L. (2007). Rethinking the fragmentation of the cyberpublic: from consensus to contestation. *New Media & Society*, 9(5):827–847.

Dahlgren, P. (2009). *Media and Political Engagement: Citizens, Communication and Democracy*. Communication, Society and Politics. Cambridge University Press.

Dentoni, D. and Bitzer, V. (2015). The role(s) of universities in dealing with global wicked problems through multi-stakeholder initiatives. *Journal of Cleaner Production*, 106:68 – 78. Bridges for a more sustainable future: Joining Environmental Management for Sustainable Universities (EMSU) and the European Roundtable for Sustainable Consumption and Production (ERSCP) conferences.

Desjardins, J. (2018). *How Google retains more than 90% of market share*. https://www.businessinsider.com/how-google-retains-more-than-90-of-market-share-2018-4?international=true&r=US&IR=T [Accessed: October 8, 2018].

Dewey, J. and Rogers, M. (2012). *The Public and Its Problems: An Essay in Political Inquiry*. Pennsylvania State University Press.

Docherty, S. (2015). The contemporary global public sphere as reincarnation of habermas' bourgeois society. *Inquiries Journal/Student Pulse*, 7(02):1–2.

Duckett, D., Feliciano, D., Martin-Ortega, J., and Munoz-Rojas, J. (2016). Tackling wicked environmental problems: The discourse and its influence on praxis in scotland. *Landscape and Urban Planning*, 154:44 – 56. Special Issue Working with wicked problems in socio-ecological systems: More awareness, greater acceptance, and better adaptation.

Elmer, G., Langlois, G., and McKelvey, F. (2012). *The Permanent Campaign: New Media, New Politics*. Digital formations. Peter Lang.

Francalanci, C. and Hussain, A. (2017). Influence-based twitter browsing with navigtweet. *Information Systems*, 64:119 – 131.

Fraser, N. (1992). *Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy*.

Galloway, C. (2017). Blink and they're gone: Pr and the battle for attention. *Public Relations Review*, 43(5):969 – 977.

Giaxoglou, K. (2018). jesuischarlie? hashtags as narrative resources in contexts of ecstatic sharing. *Discourse, Context Media*, 22:13 – 20. Discourse of Social Tagging.

Grandis, G. D. and Efstathiou, S. (2016). Introduction—grand challenges and small steps. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 56:39 – 47.

Granville, K. (2018). *Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens*. https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html [Accessed: October 31, 2018].

Gray, R. (2017). *Lies, propoganda and fake news: a challenge for our age*. http://www.bbc.com/future/story/20170301-lies-propaganda-and-fake-news-a-grand-challenge-of-our-age [Accessed: October 20, 2018].

Greenpeace (2018a). *About Us*. https://www.greenpeace.org/international/about/ [Accessed: September 12, 2018].

Greenpeace (2018b). *Tell big companies to drop dirty palm oil*. https://www.greenpeace.org/international/act/take-action-indonesia-forests-companies-dirty-palm-oil/ [Accessed: September 17, 2018].

Group, M. M. (2018). *Internet Usage Stats*. https://www.internetworldstats.com/stats.htm [Accessed: August 20, 2018].

Gruzd, A., Wellman, B., and Takhteyev, Y. (2011). Imagining twitter as an imagined community. *American Behavioral Scientist*, 55(10):1294–1318.

Habermas, J. (2006). Political communication in media society: Does democracy still enjoy an epistemic dimension? the impact of normative theory on empirical research. *Communication Theory*, 16(4):411–426.

Habermas, J., Habermas, J., Burger, T., and Lawrence, F. (1991). *The Structural Transformation of the Public Sphere: An Inquiry Into a Category of Bourgeois Society*. Studies in Contemporary German. MIT Press.

Habermas, J., Lennox, S., and Lennox, F. (1974). The public sphere: An encyclopedia article (1964). *New German Critique*, (3):49–55.

Hansen, S. B., Padfield, R., Syayuti, K., Evers, S., Zakariah, Z., and Mastura, S. (2015). Trends in global palm oil sustainability research. *Journal of Cleaner Production*, 100:140 – 149.

Hayes, T. J. (2017). mynypd: Transforming twitter into a public place for protest. *Computers and Composition*, 43:118 – 134.

Henson, I. E. (2012). 1 - a brief history of the oil palm. In Lai, O.-M., , Tan, C.-P., , and Akoh, C. C., editors, *Palm Oil*, pages 1 – 29. AOCS Press.

Hoskins, A. (2013). Death of a single medium. *Media, War & Conflict*, 6(1):3–6.

Howlett, M. and Giest, S. (2015). Policy cycle. In Wright, J. D., editor, *International Encyclopedia of the Social Behavioral Sciences (Second Edition)*, pages 288 – 292. Elsevier, Oxford, second edition edition.

Hutchinson, A. (2016). *Here's Why Twitter is so Important, to Everyone*. https://www.socialmediatoday.com/social-networks/heres-why-twitter-so-important-everyone [Accessed: July 25, 2018].

Ingargiola, A. (2015). *What is the Jupyter Notebook?* https://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/what_is_jupyter.html [Accessed: July 31, 2018].

Inuwa-Dutse, I., Liptrott, M., and Korkontzelos, I. (2018). Detection of spam-posting accounts on twitter. *Neurocomputing*.

Jackson, L. and Valentine, G. (2014). Emotion and politics in a mediated public sphere: Questioning democracy, responsibility and ethics in a computer mediated world. *Geoforum*, 52:193 – 202.

Java, A., Song, X., Finin, T., and Tseng, B. (2007). Why we twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, WebKDD/SNA-KDD '07, pages 56–65, New York, NY, USA. ACM.

Kaplan, A. M. and Haenlein, M. (2010). Users of the world, unite! the challenges and opportunities of social media. *Business Horizons*, 53(1):59 – 68.

Kasmani, M. F., Sabran, R., and Ramle, N. (2014). Can twitter be an effective platform for political discourse in malaysia? a study of pru13. *Procedia - Social and Behavioral Sciences*, 155:348 – 355. The International Conference on Communication and Media 2014 (i-COME'14) - Communication, Empowerment and Governance: The 21st Century Enigma.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., and Silvestre, B. S. (2011). Social media? get serious! understanding the functional building blocks of social media. *Business Horizons*, 54(3):241 – 251. SPECIAL ISSUE: SOCIAL MEDIA.

Kittle, M. (2011). *Twitter according to McLuhan*. http://www.meagankittleautry.com/2011/01/25/twitter-according-to-mcluhan/r/ [Accessed: September 19, 2018].

Koc, M. and Barut, E. (2016). Development and validation of new media literacy scale (nmls) for university students. *Computers in Human Behavior*, 63:834 – 843.

Kudugunta, S. and Ferrara, E. (2018). Deep neural networks for bot detection. *Information Sciences*, 467:312 – 322.

Kuitenbrouwer, J. (2018). *Als we het internet vrij laten, neemt het ons gevangen*. https://www.nrc.nl/nieuws/2018/10/05/als-we-het-internet-vrij-laten-neemt-het-ons-gevangen-a224738 [Accessed: October 8, 2018].

Kwak, H., Lee, C., Park, H., and Moon, S. (2010). What is twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 591–600, New York, NY, USA. ACM.

Lehti, L. and Kallio, J. (2017). Participation in an online social policy discussion: Arguments in focus. *Discourse, Context & Media*, 19:58 – 65. The Digital Agora of Social Media.

Li, X., Wang, Y., Zhang, A., Li, C., Chi, J., and Ouyang, J. (2018a). Filtering out the noise in short text topic modeling. *Information Sciences*, 456:83 – 96.

Li, X., Zhang, A., Li, C., Ouyang, J., and Cai, Y. (2018b). Exploring coherent topics by topic modeling with term weighting. *Information Processing Management*.

Lo, S. L., Chiong, R., and Cornforth, D. (2016). Ranking of high-value social audiences on twitter. *Decision Support Systems*, 85:34 – 48.

Masani, M. Y. A., Izawati, A. M. D., Rasid, O. A., and Parveez, G. K. A. (2018). Biotechnology of oil palm: Current status of oil palm genetic transformation. *Biocatalysis and Agricultural Biotechnology*, 15:335 – 347.

Mauskopf, S. (2012). *Tailored Trends bring you closer*. https://blog.twitter.com/official/en_us/a/2012/tailored-trends-bring-you-closer.html [Accessed: September 11, 2018].

McClarren, R. G. (2018). Chapter 1 - getting started in python. In McClarren, R. G., editor, *Computational Nuclear Engineering and Radiological Science Using Python*, pages 3 – 19. Academic Press.

Merriam-Webster (2018a). *Authority*. https://www.merriam-webster.com/dictionary/authority [Accessed: August 5, 2018].

Merriam-Webster (2018b). *Characterize*. https://www.merriam-webster.com/dictionary/Characterize [Accessed: July 30, 2018].

Merriam-Webster (2018c). *Fragmentation*. https://www.merriam-webster.com/dictionary/polarization [Accessed: July 25, 2018].

Merriam-Webster (2018d). *Informed*. https://www.merriam-webster.com/dictionary/informed [Accessed: September 17, 2018].

Merriam-Webster (2018e). *Institution*. https://www.merriam-webster.com/dictionary/Institution [Accessed: August 1, 2018].

Merriam-Webster (2018f). *Microblogging*. https://www.merriam-webster.com/dictionary/Microblogging [Accessed: October 10, 2018].

Merriam-Webster (2018g). *Reason*. https://www.merriam-webster.com/dictionary/Reason [Accessed: October 6, 2018].

Merriam-Webster (2018h). *Status*. https://www.merriam-webster.com/dictionary/Status [Accessed: August 10, 2018].

MIT (2017). *First Evidence That Social Bots Play a Major Role in Spreading Fake News*. https://www.technologyreview.com/s/608561/first-evidence-that-social-bots-play-a-major-role-in-spreading-fake-news/ [Accessed: October 2, 2018].

Monroe and Alan (1998). Public opinion and public policy, 1980-1993. *Public Opinion Quarterly*, 62(1):6–28.

Morozov, E. (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism*. PublicAffairs.

Morstatter, F., Pfeffer, J., Liu, H., and Carley, K. M. (2013). Is the sample good enough? comparing data from twitter's streaming API with twitter's firehose. *CoRR*, abs/1306.5204.

Negt, O., Kluge, A., and Labanyi, P. (1988). "the public sphere and experience": Selections. *October*, 46:60–82.

O'Brien, F. A., Meadows, M., and Griffiths, S. (2017). Serialisation and the use of twitter: Keeping the conversation alive in public policy scenario projects. *Technological Forecasting and Social Change*, 124:26 – 40.

Oltulu, P., Mannan, A. A. S. R., and Gardner, J. M. (2018). Effective use of twitter and facebook in pathology practice. *Human Pathology*, 73:128 – 143.

Orellana-Rodriguez, C. and Keane, M. T. (2018). Attention to news and its dissemination on twitter: A survey. *Computer Science Review*, 29:74 – 94.

O'Hallarn, B., Shapiro, S. L., Wittkower, D., Ridinger, L., and Hambrick, M. E. (2018). A model for the generation of public sphere-like activity in sport-themed twitter hashtags. *Sport Management Review*.

Padfield, R. (2010). *Characterize*. https://transform.iema.net/article/striking-balance-palm-oil-debate [Accessed: August 20, 2018].

Page, B. I. and Shapiro, R. Y. (1983). Effects of public opinion on policy. *American Political Science Review*, 77(1):175–190.

Papacharissi, Z. (2002). The virtual sphere: The internet as a public sphere. *New Media & Society*, 4(1):9–27.

Papacharissi, Z. (2008). *The virtual sphere 2.0*. Routledge.

Parankimalil, J. (2014). *Meaning, Nature and Characteristics of Intelligence*. https://johnparankimalil.wordpress.com/2014/11/17/meaning-nature-and-characteristics-of-intelligence/ [Accessed: July 20, 2018].

Patel, N. (2014). *How to Become a Twitter Influencer: Science and Practice*. https://www.socialmediaexaminer.com/become-powerful-twitter-influencer [Accessed: August 20, 2018].

Pesch, U. (n.d.a). Elusive pubilcs in energy projects: How energy citizenship is restricted by separating local from national publics.

Pesch, U. (n.d.b). Making innovation public: Imaginaries and metamorphoses.

Pierson, P. (2002). Coping with permanent austerity: Welfare state restructuring in affluent democracies. *Revue française de sociologie*, 43(2):369–406.

Postman, N. (2006). *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*. Penguin Books.

Prévost, Robertson, M. E. I. (1930). *Adventures of a man of quality*. London: G. Routledge Sons.

Priyadarshini, P. and Abhilash, P. (2018). Sustainability science and research for attaining un-sdgs. *Journal of Cleaner Production*, 184:609 – 610.

Purnomo, H., Okarda, B., Dewayani, A. A., Ali, M., Achdiawan, R., Kartodihardjo, H., Pacheco, P., and Juniwaty, K. S. (2018). Reducing forest and land fires through good palm oil value chain governance. *Forest Policy and Economics*, 91:94 – 106. Forest policy and governance: Taking stock of multiple environmental social sciences.

Reese, S. D. and Shoemaker, P. J. (2016). A media sociology for the networked public sphere: The hierarchy of influences model. *Mass Communication and Society*, 19(4):389–410.

Rescue, R. (2018). *No World Bank money for palm oil*. https://www.rainforest-rescue.org/petitions/587/no-world-bank-money-for-palm-oil [Accessed: October 10, 2018].

Rittel, H. W. J. and Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4(2):155–169.

Romano, A. (2018). *At long last, Twitter brought back chronological timelines. Here's why they're so beloved.* https://www.vox.com/culture/2018/9/20/17876098/twitter-chronological-timeline-back-finally [Accessed: October 31, 2018].

Romanou, A. (2018). The necessity of the implementation of privacy by design in sectors where data protection concerns arise. *Computer Law Security Review*, 34(1):99 – 110.

Rosenberg, S. W. (2015). Opinion formation, theory of. In Wright, J. D., editor, *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*, pages 243 – 245. Elsevier, Oxford, second edition edition.

RSPO (2018). *About Us*. https://rspo.org/about [Accessed: July 30, 2018].

Ruysschaert, D. and Salles, D. (2014). Towards global voluntary standards: Questioning the effectiveness in attaining conservation goals: The case of the roundtable on sustainable palm oil (rspo). *Ecological Economics*, 107:438 – 446.

Schumacher, A. (2015). *TextBlob Sentiment: Calculating Polarity and Subjectivity*. https://planspace.org/20150607-textblob_sentiment/ [Accessed: November 4, 2018].

Shane-Simpson, C., Manago, A., Gaggi, N., and Gillespie-Lynch, K. (2018). Why do college students prefer facebook, twitter, or instagram? site affordances, tensions between privacy and self-expression, and implications for social capital. *Computers in Human Behavior*, 86:276 – 288.

Signorini, A., Segre, A. M., and Polgreen, P. M. (2011). The use of twitter to track levels of disease activity and public concern in the u.s. during the influenza a h1n1 pandemic. *PLOS ONE*, 6(5):1–10.

Sloria (2014). *en-sentiment.xml*. https://github.com/sloria/TextBlob/blob/eb08c120d364e908646731d60b4e4c6c1712ff63/textblob/en/en-sentiment.xml [Accessed: November 4, 2018].

Small, H. (2009). Critical thresholds for co-citation clusters and emergence of the giant component. *Journal of Informetrics*, 3(4):332 – 340.

Speier, H. (1950). Historical development of public opinion. *American Journal of Sociology*, 55(4):376–388.

Statista (2018a). *Number of monthly active Twitter users in the United States from 1st quarter 2010 to 2nd quarter 2018 (in millions)*. https://www.statista.com/statistics/274564/monthly-active-twitter-users-in-the-united-states/ [Accessed: July 15, 2018].

Statista (2018b). *Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2018 (in millions)*. https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/ [Accessed: July 18, 2018].

Stats, I. L. (2018). *Internet Usage Stats*. http://www.internetlivestats.com/one-second/ [Accessed: October 10, 2018].

Stikker, M. (2018). *Zomergasten*. https://www.vpro.nl/programmas/zomergasten/kijk/afleveringen/2018/marleen-stikker.html [Accessed: October 3, 2018].

Stimson, J. A., Mackuen, M. B., and Erikson, R. S. (1995). Dynamic representation. *American Political Science Review*, 89(3):543–565.

Stokel-Walker, C. (2018). Alt-right's 'twitter' is hate-speech hub. *New Scientist*, 237(3167):15.

Taylor, C., Gaonkar, D., Lee, B., and Kramer, J. (2004). *Modern Social Imaginaries*. Philosophy / Social Theory. Duke University Press.

Timberg, C. (2018). *Twitter is sweeping out fake accounts like never before, putting user growth at risk*. https://www.washingtonpost.com/technology/2018/07/06/twitter-is-sweeping-out-fake-accounts-like-never-before-putting-user-growth-risk/?noredirect=on&utm_term=.3b26079cd8d9 [Accessed: September 20, 2018].

Twitter (2018a). *About Twitter limits*. https://help.twitter.com/en/rules-and-policies/twitter-limits [Accessed: September 20, 2018].

Twitter (2018b). *Hateful Conduct Policy*. https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy [Accessed: October 2, 2018].

Twitter (2018c). *Overview*. https://developer.twitter.com/en/docs/tweets/search/overview.html [Accessed: August 26, 2018].

Twitter (2018d). *The Twitter Rules*. https://help.twitter.com/en/rules-and-policies/twitter-rules [Accessed: August 30, 2018].

Twitter (2018e). *Twitter is ...* https://about.twitter.com/en_us.html [Accessed: August 1, 2018].

UN (2018). *Sustainable Development Goals*. https://www.un.org/sustainabledevelopment/sustainable-development-goals/ [Accessed: July 10, 2018].

Unger, J. B., Urman, R., Cruz, T. B., Majmundar, A., Barrington-Trimis, J., Pentz, M. A., and McConnell, R. (2018). Talking about tobacco on twitter is associated with tobacco product use. *Preventive Medicine*, 114:54 – 56.

van der Noordaa, R. and van de Ven, C. (2018). *Hoe Russische trollen inspelen op westerse angsten*. https://www.groene.nl/artikel/hoe-russische-trollen-inspelen-op-westerse-angsten?utm_source=De+Groene+Amsterdammer&utm_campaign=f3fcef4807-Dagelijks-2018-08-29&utm_medium=email&utm_term=0_853cea572a-f3fcef4807-71151297 [Accessed: September 20, 2018].

van Lonkhuyzen, L. and Stokmans, D. (2018). *Amsterdam in de fout met gegevens hangjongeren*. https://www.nrc.nl/nieuws/2018/05/16/amsterdam-in-de-fout-met-gegevens-hangjongeren-2-a1603291 [Accessed: October 2, 2018].

Varol, O., Ferrara, E., Davis, C., Menczer, F., and Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. pages 280–289.

Verhagen, L. (2018). *Godfather van het World Wide Web opent aanval op Google en Facebook*. https://www.volkskrant.nl/nieuws-achtergrond/godfather-van-het-world-wide-web-opent-aanval-op-google-en-facebook~bb0372a0/ [Accessed: October 3, 2018].

VPRO (2018). *VPRO Tegenlicht: Trollen, Trump Thierry*. https://www.vpro.nl/programmas/tegenlicht/kijk/afleveringen/2017-2018/Trollen--Trump---Thierry.html [Accessed: October 2, 2018].

Warner-Søderholm, G., Bertsch, A., Sawe, E., Lee, D., Wolfe, T., Meyer, J., Engel, J., and Fatilua, U. N. (2018). Who trusts social media? *Computers in Human Behavior*, 81:303 – 315.

Webster, J. (2014). *The Marketplace of Attention: How Audiences Take Shape in a Digital Age*. The MIT Press.

WWF (2018). *Which Everyday Products Contain Palm Oil?* https://www.worldwildlife.org/pages/which-everyday-products-contain-palm-oil [Accessed: July 1, 2018].

Zhu, H. and Hu, B. (2018). Impact of information on public opinion reversal—an agent based model. *Physica A: Statistical Mechanics and its Applications*, 512:578 – 587.

A

Python Packages

Name	Version
pandas	0.22.0
numply	1.14.4
matplotlib	2.2.2
nltk	3.3
sklearn	0.19.1
umap	0.3.6
tweepy	3.6.0
json	2.0.9
textblob	0.15.1

Table A.1: Python packages used

TextBlob

TextBlob is a package for Python to conduct Natural Language Processing (NLP) tasks. The package is based on a lexicon by Sloria (2014). This lexicon includes many different English words with different meaning. The word 'great', for example, appears four times in the lexicon with different meanings: very good; of major significance or importance; relatively large in size or number or extent; remarkable or out of the ordinary in degree or magnitude or effect. All these different meanings have their own polarity score, ranging from 1 for the first one to 0.4 for the last one. TextBlob averages the polarity of the four meanings and comes to the score of 0.8. Other words, such as 'not', multiply polarity by -0.5. Words like very multiply the polarity by 0.2. Words TextBlob does not have in its lexicon are not given a score (Schumacher, 2015). More information can be found on Github: `https://github.com/sloria/TextBlob`.

Topic Mdelling: UMAP

UMAP stands for Uniform Manifold Approximation and Projection. It is a relatively new dimension reduction technique that can be used for visualization. It can be used as a drop in replacement for scikit-learn's t-SNE. More information can be found on Github: `https://github.com/lmcinnes/umap`.

B

Structure of a tweet

When figure B.1 is what we see on Twitter, the text below shows what is gathered. It shows how much information is saved per tweet and how much analysis there can be done.

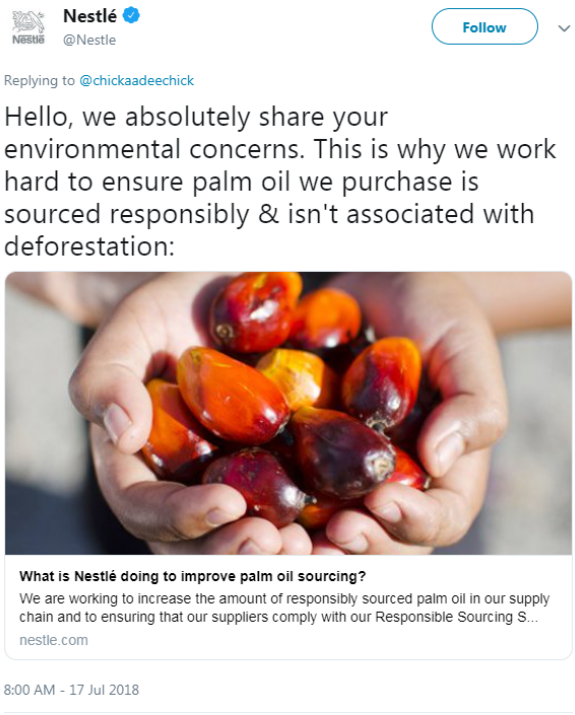


Figure B.1: Example Tweet. Collected July 24


```
{'contributors': None,
 'coordinates': None,
 'created_at': 'Tue Jul 17 15:00:34 +0000 2018',
 'display_text_range': [17, 140],
 'entities': {'hashtags': [],
  'symbols': [],
  'urls': [{'display_url': 'twitter.com/i/web/status.../1',
    'expanded_url': 'https://twitter.com/i/web/status/1019235417198120963',
    'indices': [117, 140],
    'url': 'https://t.co/r23gvmVJ7n'}],
  'user_mentions': [{'id': 157900878,
    'id_str': '157900878',
```

```
    'indices': [0, 16],
    'name': 'Chickaadee Chick',
    'screen_name': 'chickaadeechick'}]},
'extended_tweet': {'display_text_range': [17, 218],
 'entities': {'hashtags': [],
 'symbols': [],
 'urls': [{'display_url': 'bit.ly/2GcQ68n',
    'expanded_url': 'http://bit.ly/2GcQ68n',
    'indices': [195, 218],
    'url': 'https://t.co/Q2Ebeo33ZP'}],
 'user_mentions': [{'id': 157900878,
    'id_str': '157900878',
    'indices': [0, 16],
    'name': 'Chickaadee Chick',
    'screen_name': 'chickaadeechick'}]},
 'full_text': "@chickaadeechick Hello, we absolutely share your environmental
 concerns. This is why we work hard to ensure palm oil we purchase is sourced
 responsibly &amp;amp; isn't associated with deforestation: https://t.co/Q2Ebeo33ZP"},
'favorite_count': 0,
'favorited': False,
'filter_level': 'low',
'geo': None,
'id': 1019235417198120963,
'id_str': '1019235417198120963',
'in_reply_to_screen_name': 'chickaadeechick',
'in_reply_to_status_id': 1019026587264700421,
'in_reply_to_status_id_str': '1019026587264700421',
'in_reply_to_user_id': 157900878,
'in_reply_to_user_id_str': '157900878',
'is_quote_status': False,
'lang': 'en',
'place': None,
'possibly_sensitive': False,
'quote_count': 0,
'reply_count': 0,
'retweet_count': 0,
'retweeted': False,
'source': '<a href="http://www.salesforce.com" rel="nofollow">Salesforce -
Social Studio</a>',
'text': '@chickaadeechick Hello, we absolutely share your environmental
concerns. This is why we work hard to ensure palm ...oi https://t.co/r23gvmVJ7n',
'timestamp_ms': '1531839634595',
'truncated': True,
'user': {'contributors_enabled': False,
 'created_at': 'Fri Mar 06 16:45:17 +0000 2009',
 'default_profile': False,
 'default_profile_image': False,
 'description': 'Enhancing quality of life and contributing to a healthier
 future.',
 'favourites_count': 6707,
 'follow_request_sent': None,
 'followers_count': 219639,
 'following': None,
 'friends_count': 661,
 'geo_enabled': True,
 'id': 23085995,
```

```
'id_str': '23085995',
'is_translator': False,
'lang': 'en',
'listed_count': 1492,
'location': 'Vevey, Switzerland',
'name': 'Nestlé',
'notifications': None,
'profile_background_color': 'C0DEED',
'profile_background_image_url': 'http://abs.twimg.com/images/themes/theme1
/bg.png',
'profile_background_image_url_https': 'https://abs.twimg.com/images/themes
/theme1/bg.png',
'profile_background_tile': False,
'profile_banner_url': 'https://pbs.twimg.com/profile_banners/23085995/
1493799727',
'profile_image_url': 'http://pbs.twimg.com/profile_images/10024725497
73709312/B_17xohH_normal.jpg',
'profile_image_url_https': 'https://pbs.twimg.com/profile_images/1002472549773
709312/B_17xohH_normal.jpg',
'profile_link_color': '0084B4',
'profile_sidebar_border_color': 'FFFFFF',
'profile_sidebar_fill_color': 'DDEEF6',
'profile_text_color': '333333',
'profile_use_background_image': True,
'protected': False,
'screen_name': 'Nestle',
'statuses_count': 11934,
'time_zone': None,
'translator_type': 'none',
'url': 'http://www.nestle.com',
'utc_offset': None,
'verified': True}}
```

C

Streaming, influencers and different users

C.1. Streaming

I have decided to use a virtual server via Digital Ocean. This costs 5 euros per month (memory 1 GB, 1 vCPU, SSD Disk 25 GB and trasfer 1 TB), is based in Amsterdam and works on Ubuntu. It already has Python 3 installed, and I have installed Tweepy using pip. I have uploaded the scripts via SSH, and am using 'Screen' to let the whole project run for two months.

C.2. Influencers

Name		PepsiCo	Nestle	RSPO Tweets	KitKat	bdpd_sawit	indofood	Greenpeace UK	Mike Hudema	BBC Earth
User Level Resources	Following	14876	672	18795	503	47	29	5140	45283	1661
	Followers	327142	223173	22980	386790	1228	2109	175905	78673	466153
	Favorites	532	6902	5670	14415	178	5	5501	15970	15630
	Tweets Total	26236	12406	14180	15783	391	204	35728	26079	24786
	Tweets in database	69	6	345	0	26	0	6	4	2
Tweet Level Resources	Retweeted Total	12413	27620	9735	30855	25988	376	261011	754975	613217
	Favorited Total	29022	27005	19965	20464	7506	187	138969	630485	1221971
	Mentions in database	4260	3235	5476	48	6715	4544	2283	1731	1553

Table C.1: The top ten users defined as influencers and their user- and tweet- level resources. Generated September 12

Category	Users
Business	bpdp_sawit
Individual	altischeah1
NGO	Animal_Watch
Media	BBCEarth
Individual	BellaLack
Individual	BiologistDan
Individual	BlesynAffia
Individual	davidsting414
Individual	DrBirute
Business	DVIINGENIERIA
NGO	extinctsymbol
NGO	foe_us
NGO	Global_Witness
NGO	Greenpeace
NGO	GreenpeaceUK
Business	indofood
Individual	IndraNooyi
Individual	JDobroszczyk
Individual	MikeHudema
Business	milieudefensie
Business	MPOCEurope
Business	mpocHQ
Government	my_mpi
Business	Nestle
Media	NEWSDAYS_HN
Individual	NWarind
NGO	orangulandtrust
Individual	orangutans
NGO	OrangutansSOS
nonprofit	paspi_sawit
Business	PepsiCo
Individual	PeterEgan6
nonprofit	RSPOtweets
Business	sathia_varqa
Media	sawitmedia
Individual	setyourmindfree
Individual	stephenfry
Individual	sunfloweraidil
nonprofit	Tweets4Vegans
Individual	wahyudin_ihsan
Individual	yobeahrs

Table C.2: The different users analyzed and their categorization

D

Interviews with policy makers on the use of Social Media data

Three interviews with social media analysts making policy. The interviews had an open approach, meaning no fixed protocol is used. This allows for in-depth conversations on specific fields of expertise per interviewee. The goal of these interviews is to find out if (and if yes, how) social media data is used to measure public opinion for policy making. For the exploratory interviews, experts with different backgrounds have been selected. Frank van der Linden is media analyst at the Ministry of Social Affairs and Employment and Michael Elbers is social media analyst at the city of Amsterdam.

D.1. Frank van der Linden - Ministry of Social Affairs and Employment

- *Who?* Media Analyst at Ministry of Social Affairs and Employment
- *When?* September 7, 2018, 2.30 - 3.15 pm
- *What?* Does not think we should use Twitter data to analyze public opinion

This interview is conducted in Dutch and therefore the answers are also in Dutch.

Goal
To find out if (and if yes, how) social media data is used to measure public opinion at the Ministry of Social Affairs and Employment.

Nemen jullie in het algemeen publieke opinie data mee?
Ja zeker. Manier waarop verschilt per ministerie.

Hoe?

- Voor sociale zaken waarvan we willen weten wat politici vinden gebruiken we Twitter. Daar spuien ze veel. Veel politici geburiken het omdat er veel journalisten opzitten. Kip-ei. Journalisten zitten erop omdat de politici erop zitten. Ook kijken we naar het sociale middenveld. Zij linken vaak door naar opinieartikelen/artikelen of kondingen stakingen aan. Dit is meer een signalerende functie.
- Sneeuwbal effect: als het ene medium gaat schrijven rolt het verder. Echt wel een prikkel voor politici om het op te pakken.
- Wel als activated public opinie want het is vooral de Haagse top. Politie, journalisten, politici. Groeperingen uit het sociaal middenveld, vakbonden/pensioenfondsen. Mensen die graaag gehoord willen worden door politici. Ons-kent-ons clubje. Journalisten gaan vaak op Twitter kijken wat er speelt en hoe ze daarover schrijven.
- Webcare op Twitter. Als mensen vragen hebben over hun positie, kunenn ze die vragen op Twitter ook stellen.

- Zoeken op #ontslagen? En wat dan, wat willen we daaruit halen? Naast het gebrek aan doel in zo'n zoektoch ook een gebrek aan capaciteit.

Hoe zien jullie Twitter

- Twitter = kroegpraat.
- Gemene deler bij alle ministeries: de publieke opinie kan niet van Twitter gehaald worden doordat er maar een klein gedeelte van de Nederlanders actief is over Twitter.
- Een aantal jaar geleden was het idee dat Twitter de democratiserende macht was. Dat beeld hangt er nog een beetje bij sommigen.
- Frank - communicatie SZW, als mensen van onze beleidsdirecties bezig zijn met nieuwe wetgeving, baseert hij zich niet op wat hij op Twitter ziet. Het lijkt hem niet goed om wetgeving rondom het pensioenstelsel, bijvoorbeeld, te baseren op Twitter.
- Uitgesproken mening is uiteraard aanwezig, en de silent majority dan?
- Hoe bepaal je influence? Kwaliteit van de volgers is erg belangrijk.
- Snel. Mainstream media heeft 10 min-15 min nodig om iets op poten te zetten, Parescope, Twitter meteen informatie.

Wat denkt u dat het effect van opinievorming dmv social media is?
Het kan een voedingsbron zijn (de politici, etc). Hij denkt dat Opiniepagina's in een krant, columns, items in het 8 uur journaal meer impact heeft. Maar de inhoud van die dingen kan wel van Twitter komen.

Wat denkt u dat de invloed van social media in de toekomst wordt?
Ik weet dat slechts een deel van de Nederlanders op Twitter zit en ik bedoel dan ook vooral niet het meten van de algehele publieke opinie (als er al zoiets bestaat). Ik geloof meer dat het kan gaan om de zogenaamde 'activated public opinion', welke trends in de public opinion kan voorspellen.
Hoe bekijken/meten jullie nu publieke opinie?

- Vroeger monitoringsoftware. Vooraf vragen opstellen, dingen die je verwacht tee zien. En dan mis je veel. Daarom vooral publikesonderzoeken met focusgroepen.
- Voor publieke opinie willen we dat vooral zorgvuldig doen. Op de ouderwetse manier. Veel haken en ogen aan het gebruik van Twitter. Duidelijkste haak is dat niet iedereen op Twitter zit dus je weet ook niet hoe representatief is voor de Nederlanders.

Zijn jullie bang Big Brother te worden?
Wij beperken ons tot wat toch al publiekelijk openbaar is. Als we meer willen weten dan maken we er een in-depth face-to-face interview van.

D.2. Michael Elbers - City of Amsterdam

- *Who?* Media Analyst at the city of Amsterdam
- *When?* October 3, 2018, 10.00 - 11.15 am
- *What?* Is already using Twitter data to monitor public opinion but is aware of the recent shifts on Twitter towards the extreme right and bots

This interview is conducted in Dutch and therefore the answers are also in Dutch.

Nemen jullie in het algemeen social media data voor publieke opinie mee? Hoe?

- We gebruiken de tool Coosto om bij te houden wat er gezegd wordt op social media. Dit gaat over Twitter, Facebook, Instagram en op onze website. Het grootste volume gaat veruit via Twitter.
- Elke maandagmorgen hebben we een terugblik op wat er de afgelopen week gebeurd is op social media en of dit aansluit bij onze communicatiedoelen.
- Toch wordt het niet een op een aangenomen als 'publieke opinie'. Veelal wordt gedacht, 'het is maar Twitter', maar bepaalde meningen, van bepaalde invloedrijke gebruikers, zijn niet te missen.
- Wij proberen als gemeente toe te gaan naar een responsieve overheid en daar past social media data absoluut in. Zow illen we weten wat er speelt en die feedback meenemen in verdere ontwikkelen.

Hoe bekijken/meten jullie nu publieke opinie?

- We gebruiken dus social media data maar daarnaast ook tradionale methodieken zoals surveys en bewonersbijeenkomsten.

Waarvoor gebruiken jullie social media data allemaal?

- Wij gebruiken het enkel om te kijken wat mensen vinden. Politie en justitie gebruiken het veel meer.
- Omdat het een vrij nieuw gebied is zijn we nog aan eht werk aan de bewustwording bij ambtenaren wat wel en niet mag. Er zijn regels i.v.m. privacy, maar soms weten wij die ook niet precies. Een groot deel van wat valt in een grijs gebied.

Wat denkt u dat het effect van opinievorming d.m.v. Twitter is?

- Twitter is niet de awarheid.
- De tendens op Twitter is aan het verschuiven. We zien duidelijk een sterke alt right groepering opkomen en twijfelen daardoor aan de betrouwbaarheid.
- Coosto verteld a.d.h.v. alle data wat *trending* is. Dit komt vaak niet overeen met de zogenaamde trending topics van Twitter en deze discrepantie toont aan dat Twitter een vinger in de pap heeft met bepalen wat wij zien.
- Zij maken al een voorselectie en kijken niet meer naar de rauwe data. Wij maken daaroverheen weer een interpretatie. Bij ons wisselt het wie Coosto interpretereert en de resultaten zijn dan ook elke keer iets anders.

E

Interviews with experts on palm oil

Three interviews with experts of the palm oil sector have been conducted. The interviews had an open approach, meaning no fixed protocol is used. This allows for in-depth conversations on specific fields of expertise per interviewee. The goal of these interviews is to get insight in the wicked problem of palm oil and find keywords for the quantitative data gathering. For the exploratory interviews, experts with different backgrounds have been selected. Levien de Legé from the Dutch Palmares Partnership, Thijs Pasman from MVO and Gersom van der Elst from the Dutch Ministry of Agriculture, Nature and Food Quality.

E.1. Levien de Legé - Palmares Partnership

- *Who?* Levien de Legé is for Palamares working to sustainabalize the chain from palm residue in Malaysia
- *When?* July 18, 2018, 11.00 - 11.45 am
- *What?* This problem has to be approached from the demand- and supplyside. The Western world loves palm oil and puts it into everything. You cannot just blame Asian countries for fulfilling this demand.

Wat doen jullie precies met het Palmares Partnership?
PIB - Programma wat het doel heeft om een bepaalde groep NL'se berdrijven in een bepaalde markt te promoten. Heel specifiek, in een land, in een sector. Kan van alles zijn. Palmares: bedrijven met als doel afvalstromen in de palmolie-industrie beter te verwaarden in Maleisië. Van crude palm olie wordt heel veel biomasa verzameld, en daar kan je biogas van maken (voor waterzuivering, vergisting), palletiseermachines (biomassa wordt omgezet in brandstof-pallets), nu wordt het afval gewoon verbrand voor stroom (terug naar plantages, als mest).

Het speelveld van palmolie is zeer complex. Kunt u me kort schetsen hoe u het ziet?
Het feit dat er jungle gepakt wordt is niet goed. Maar twee dingen: in Colombia is geen jungle, gewoon grasveld dus daar past het gewoon een stuk beter. Ten tweede: de jungle is gewoon gekapt, daar moet je wat nuttigs mee doen. Wij kunnen van alles vinden, maar dit hebben wij niet onder controle met onze publieke opinie. Het gebeurt enkel omdat er hier een markt is.

Wie bepaalt, in uw ogen, het debat over palmolie?
Milieuorganisaties, NGOs, daar ontstaan dat soort onderzoeken. Marketing! En niks zo zielig als een oranguntang met een jong kind die in de laatste boom zit. Sterke beeld. RSPO: iedereen probeert zich daaraan te houden. Maar dat voorkomt niet dat er opnieuw jungle gekapt wordt. Bestrijdingsmiddelen, watermanagement, het wordt een hele serieuze industrie. Het hele cowboy gebeuren dat mensen maar vanalles doen dat is in Indonesie zo, maar in Malasie zeker niet.

Kunnen en willen we toe naar een wereld zonder palmolie?

- Vroeger was het ook zo. 40 jaar geleden was er ook geen palmolie. Nu gaat de discussie over het toevoegen van palmolie aan biodiesel, en dat was een zelfde disucssie vroeger toen we bioethanol maakte van mais enzo. Voor energie gaan wij landbouwgewassen gebruiken.
- Primaire levensbehoefte is voedsel, niet energie.
- Je hebt een aanbieder en een klant. Kip en ei: het kappen van regenwouden gebeurt met name voor hout. Daarna voor palmolie. Wat komt er eerst? Wordt het eerst gekapt voor hout en dan landbouwgebied, of andersom? In indonesie is bij wet gergeld dat mensen dat mogen.

Wat is er nodig om volledig over te stappen op duurzame palmolie?

Geen extra regenwoud meer kappen is een mogelijkheid. Organische groei, productiviteitsverbetering. Bedrijven daar hebben RD afdelingen enzo, het is heel serieus allemaal daar. Slimmere landbouw te plegen, betere soorten, meer resistent tegen ziektes. Het is echt een landbouwsector geworden. (gaat allemaal over Malesie). Er zijn nog steeds bandieten die het kappen, maar dat blijft lastig. Corruptie is er toch behoorlijk.

Bij de EO een documentarie over Sarawak, Malesies deel van Borneo, 30 jaar lang een chief minister en die heeft het hele land uitverkocht. Die heeft 32 miljard verdiend, en hij zegt dat hij onschuldig is maar ondertussen... Voor kapvergunning. Nu is 2/3 van het land gekapt. Daar in de plaats zijn wel palmolie plantages gekomen.

Overig

- Publieke opinie wordt vaak gestuurd. Orangoetang is boegbeeld geworden. Het is nooit goed om jungle te kappen. Wij hebben ons land ook helemaal gekapt ooit, hier in Nederland is ook alles aangelegd. Elke boom. 500, 600 jaar geleden hebben wij precies hetzelfde gedaan. Dat zeggen ze in Azie ook: mogen wij niet zelf ons land in richten?
- Overal zijn cerificeringsstystemen voor. RSPO, voor hout ook eentje, dus dat is er allemaal wel. Maar als we ons er allemaal aan zouden houden. Je hebt een supplyer en een demand. Het probleem zit echt bij ons, als het westen. Het is logisch dat je bos kapt voor landbouw, dat gaat ten koste van de natuur. Maar als je dat op een responsible manier doet dan kan dat.
- Mensen die hij in Malesie kent zijn heel serieus, geen boeven, met heel veel onderzoek omdat allemaal te optimaliseren. Het is een serieuze business. Als wij er biodiesel van willen maken.. Wij moesten groene diesel hebben, en dit was de oplossing en daar lopen we nu op af te geven.
- In Malesie is het heel erg tropisch, dus daar groeit ook niet zo veel. Het is veel te heet voor iets anders, maar diversificeren zou goed kunnen. Maar wij nemen als Palmares aan, laten we proberen het zo groen mogelijk te maken. Malesie heeft ook het Parijs-akkoord getekend dus ook een plicht op zich genomen. Ze beseffen wel dat je niet zomaar alles kan dumpen. Daar proberen wij zo veel mogelijk te ondersteunen, door technologie aan te bieden waarmee ze dat afval kunnen verwaarden. 10 procent omzet. Brandstofpallets, biogas, electriciteit, maar er is daar helemaal geen gasafname stysteem. En voor jezelf is het veel te veel. Sommige oliefabrieken maken hun eigen stroom, omdat ze zo ver afgelegen zijn van het electriciteitsnet. De infrastrutctuur is ook niet zo dat het heel makkelijk wat kwijt kan. Daar is veel te halen!
- Wij kunnen onze normen niet opleggen aan andere delen vd wereld. Schoppen werkt averechts.
- Palmares is ook erg veel bezig met de sociale effecten, zoals de energie die opgewekt kan worden met het residu gebruiken voor drinkwater, gas om op te kopen, etc. Lastig omdat dit allemaal private bedrijven zijn die de plantages etc bezitten en het overheidsbelang is. Ipv het gas efficient gebruiken worden er bijv. zonnepanelen geinstalleerd. Wordt ook gekeken naar subsidiemogelijkheden om dit zo aantrekkelijk mogelijk te maken.

E.2. Gersom van der Elst - Dutch Ministry of Agriculture, Nature and Food Quality

- *Who?* Analyst sustainable agrarian raw material chains
- *When?* July 18, 2018, 13.00 - 13.45 am
- *What?* The Netherlands do not want to boycott palm oil because in doing so, we would also give away our marketpower.

This interview is conducted in Dutch and therefore the answers are also in Dutch.

Samenvatting

- Producenten van palmolie zijn erg geconcentreerd. Het speelveld is dan ook niet heel erg complex. Wat het complex maakt zijn de palmoliebedrijven met dochterondernemingen: lees rapport milieudefensie.
- Ontbossing moet in perspectief worden gezien. Zo zorgt soya en vee nog steeds voor veel meer ontbossing dan palmolie.
- Daarnaast is het overstappen van palmolie op andere plantaardige olien helemaal niet per se beter/efficienter. Als je palmolie bomen plant op grasland (wat er nu in Colombia gebeurt) heeft dat zelfs een positieve voetafdruk omdat bomen meer CO2 opnemen dan gras.
- Nederland wil niet helemaal af van palmolie want dan geef je ook helemaal je marktmacht op. Op dit moment zijn de grootste afnemers China (33%) en Indonesie zelf (15%), zelfde als Europa gezamelijk. Nu kan Nederland dus nog macht uitoefenen op zowel productie als consumptie en lijkt er geen beleid gericht te zijn op het helemaal stoppen met het gebruik van palmolie. De overheid zit dus een beetje in een spagaat.
- Nederlandse overheid werkt met samen met IDH/Solidaridad, meer dan met bijv. milieudefensie omdat het beleid van NL niet in lijn is met milieudefensie.

E.3. Thijs Pasmas - The Netherlands Oils and Fats Industry

- *Who?* Working on Ssutainable sourcing in oils and fast at MVO
- *When?* August, 2018, 16.30 - 17.15 am
- *What?*

This interview is conducted in Dutch and therefore the answers are also in Dutch.

Wat doet MVO, de ketenorganisatie voor olien en vetten?
Vertegenwoordigen raffinaderijen (groot in volume, groot in issue) en namens de sector het debat voeren.

Is duurzaam echt duurzaam?
Een van de elementen van het debat. Verduurzaming is een proces van verbetering en nieuwe inzcihten. Voor hem zijn het meer bedrijven die aan boord stappen van dat proces. Bedrijven die initiatieven doen om te verduurzamen, dus het gebeurt echt.

RSPO- standaarden, in de keten gecertificeerd, meest leidende standaard. In Europa wordt dat geaccepteerd, vaak constructieve kritiek. ”Goede basis, moet wat verbeteren, gebeurt elke

Kunnen/willen we toe naar een wereld zonder palmolie?
Wereldwijd kunnen we niet zonder.

Ik zie dat de publieke opinie heel sterk geuit wordt op Twitter, en ik ben me ervan bewust dat Twitter enkel een medium is maar je moet het ergens vandaan halen. Wie/wat denk jij dat deze sterke mening drijft?
Begonnen door de NGOs (WWF, greenpeace). Toen was het ook echt nodig. WWF is heel erg constructief. Vooral kleine, fanatieke apen-liefhebbers, heel erg emotioneel op de orang-utan. Tastbaar, aaibaar. Puur gericht op de orang utan, het is goed of fout teriwjl de werkelijk niet zo zwart is. Twitter moet kort en krachtig.

Heeft deze sterke mening een effect op het beleid?
In het verleden heeft het wel degelijk gewerkt. Bedrijven zijn omgegaan door die campagnes ooit. Toen werkte het wel maar nu niet meer, averchts. Het is nooit genoeg.

Wie is er verantwoordelijk voor het beleid, NL, Europa, ..?
Werkt het op overheidsbeleid? Politieke belangen van andere landen spelen een belangrijke rol. Handelsbelangen, etc.

Beleid - politiek zijn twee verschillende dingen. Het beleid is wat rustiger. Zij proberen met de landen zelf tot oplossingen te komen. In de praktijk: debat, dan kamervragen, dan wordt er gereageerd en dan moet ze kunnen aantonen wat ze daaraan doet. Beleid zet zich dan wel in op verduurzaming. Niet iets van de laatste jaren.

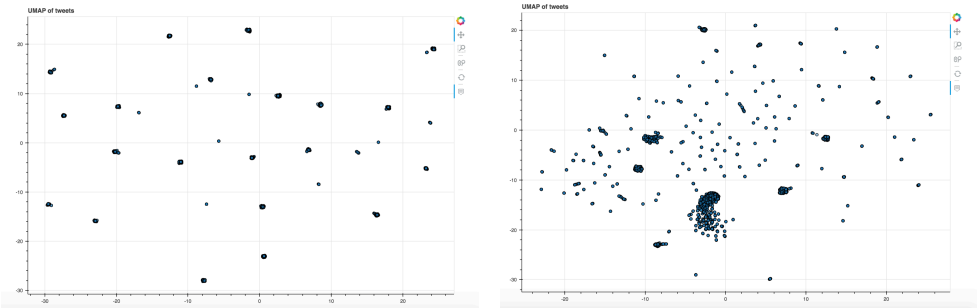
Er zijn veel tegenstrijdige belangen. In lijn met ’follow the money’, wie zijn de grote verdieners aan palm olie? En de grote verliezers?
Grootverdieners, Malaysie en Indonesie. Grote bedrijven maar ook de landen (boeren zijn heel erg afhankelijk).Vraag naar duurzaamheid komt uit het westen. De trend wordt ingezet en het idee is ook dat ook China dat gaat doen.

F

Spam

Peak 4: September 6, 2018

- Emergence: September 6
- Cause: Unknown, there is no clear hashtag or retweet causing this peak.
  - See figure F.2 for the comparison of two topic models. Figure F.2a shows how there are not clear groups of tweets like in figure F.2b.
  - A factor that might cause the peak is the user mention RSPOtweets. This is mentioned in 38% of the tweets.
  - When looking a bit deeper into who and what is tweeted with that, see figure 6.15. It does not seem like in section 6.1 these accounts are automated but it is also a bit strange. Additionally, the official limit is 100 tweets per hour (Twitter, 2018a), broken down in smaller sections. With 9 tweets/minute (as in figure 6.15a), the user is well above that.
  - From the reply graph in figure figure F.1 we can see that Nestle and PepsiCo play a central role.
  - We can see for a specific user called Freja_Petersen that her tweets do not exist anymore. The url says ”Sorry, that page doesn’t exist!”



(a) Topic model to show how splintered the topics are on peak 4

(b) Topic model to show the comparison of peak 1

Figure F.2: Two topic models showing the difference between peak 1 and 4. Peak 1 has a clear couple of reasons to peak.

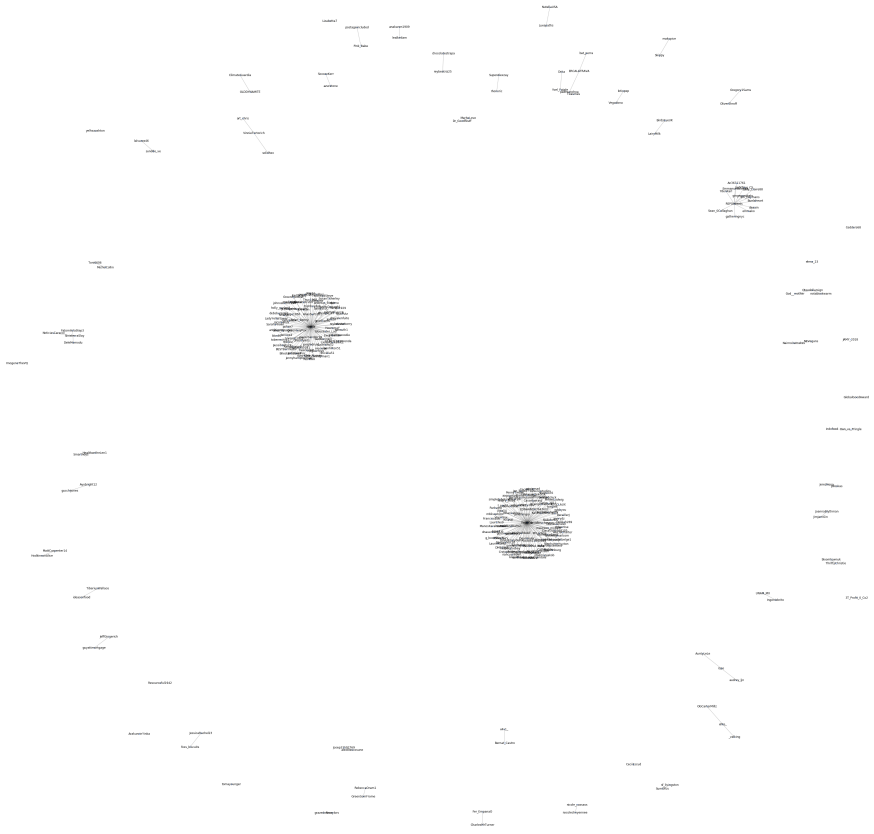


Figure F.1: Giant reply graph of peak 4 showing a central position from Nestle and PepsiCo