Master thesis

Sharing stories about politics: Social realities of Brexit on Twitter

Using symbolic convergence theory for analysis of social media sites

Adrijan Markovikj

s2043645

Faculty of Behavioural, Management and Social sciences

Communication Studies

Examination committee / Supervisors

Dr. S. A. de Vries

MSc. R. Marinescu-Muster

May 2020

UNIVERSITY OF TWENTE

Table of contents

Abstract

This study proposes a descriptive model for social realities as a response and reflection to real world events shared through dramatized stories and developed as a result of online group communication. The model was developed by aligning theoretical concepts from symbolic convergence and narrative paradigm with social media site communication and used on the case Brexit on Twitter. Previous research using these theories was conducted with a focus on organizational settings, interpersonal communication or visual cues that lead to the development of topics and social realities. However, this study postulates that symbolic convergence theory has a place in analyzing todays social media communication and the process of storytelling in online group communication. The methods of analysis describing social realities on social media are based on text analysis and mining, topic modeling and network analysis. With these methods research was able to uncover patterns of communication in the stories shared about the political event Brexit. The most frequent patterns were negative dramatization of stories that address British politicians, mocking parliament and distrust of government. Storytelling on social media sites, and in this case Twitter, creates communities which share stances and viewpoints on social, political, or economic issues. This study finds that the narrative evolution of social realities on social media sites can be heavily influenced by negatively dramatized stories that shift viewpoints about real-world events and political issues of whole communities.

Keywords: Social reality; Dramatization; Brexit; Storytelling on social media

1. Introduction

Brexit is arguably the most significant referendum in the history of the European Union (Hobolt, 2016). The referendum vote passed on 24th of June 2016 ("Brexit: Your simple guide to the UK leaving the EU", 2019) in favor of Britain leaving the European Union (52% leave against 48% remain votes). While the British exit was ongoing for the past 4 years (2016 June – 2020 January), it shook the British nation dividing it into Leavers and Remainers (Hobolt, 2016; Goodwin & Heath, 2016), even well after the vote had passed. The topic of Brexit was prevalent throughout those four years because it brought the British people to the brink of social, economic and political crisis in their debates. These Brexit debates spread across social media sites, specifically on Twitter (Hall *et al.*, 2018; Bassilakis *et al.*, 2018). The focus here is on Twitter because this social media site has ballooned in use for political campaigning (Southern & Lee, 2018) and public relations and address (Collins *et al.*, 2019). Brexit as an unprecedented event of a European Union member state planning to leave, implying global political and economic ramifications (Hobolt, 2016).

The global impact of Brexit and the prevalence of the topic on Twitter in recent years has attracted research from multiple disciplines. For example, social and political sciences (Goodwin & Heath, 2016; Hobolt, 2016) with a focus on the social implications of Brexit and explaining the referendum vote. There are also computer and data science studies (Bassilakis *et al.*, 2018; Grčar *et al.*, 2017; Hall *et al.*, 2018; Llewellyn & Cram, 2016; Niklander, 2017) that provide insight into the opinion or stance of Twitter users on the issue of Brexit, all with similar classification methods (based on sentiment analysis). In addition to similar approaches towards data classification, those studies suggest that online group communication develops topics of discussion that reflect events in the real world. For instance, Grčar *et al.* (2017) predicted the stance of British voters as Leave by analyzing the diffusion of information on Twitter as a social network, while identifying online communities. In light of their findings, they point out that communication on social media sites (namely Twitter) has an impact on society and real-world events.

According to symbolic convergence theory (Bormann, 1985; Bormann *et al.*, 2001), in group communication, real-world events are accounted for as stories that become topics of

discussion for a group. This storytelling leads to the development of social reality, as well as the development and sustaining of group consciousness among group members and communities (Bormann, 1985; Griffin, 2011, pp.247-257). Bormann *et al.* (2001) further elaborate that symbolic convergence theory takes into consideration three types of realities, social, symbolic and materialistic. According to Bormann *et al.* (2001), social reality is our everyday life composed of relationships, social status, hierarchies and agreements among many other characteristics. While, the symbolic reality is one that we develop and sustain in order to "make sense of the world" which is composed of ideologies, worldviews and orientations. Finally, the materialistic reality consists of material facts, for instance objects. They emphasize that sharing stories in group communication leads the group members to also share their "sense of the world" or symbolic reality. This, in turn, develops a social reality that is shared between the group members. Castor (2005), in a study of organizational decision-making of faculty members, sheds light on the development of social reality in group communication. She finds that in addition to the process that symbolic convergence theory accounts for, vocabulary and use of language also play an important role in this development. According to Castor (2005) reality is socially constructed through communication. By analogy, this study applies her statement to online group communication on Twitter and aims to uncover: "**How can we use symbolic convergence theory to describe social realities developed as a result of online group communication on social media sites?**"

To answer the posed research question this study develops a descriptive model based on the framework of symbolic convergence theory. Moreover, this study chooses Twitter as a case for online group communication on social media sites. The developed descriptive model for describing social realities on social media sites will be refined and operationalized for analysis of Twitter. Symbolic convergence is the theoretical backbone of this study and has been used to evaluate visual symbolic cues that lead to a group developing their social reality, such as political cartoons (Bormann, 1978), satirical images (memes) about athletes (Page *et al.,* 2016), and political campaigning material (Page & Duffy, 2018). Additionally, it has found use in organizational decision-making (Bormann *et al.,* 1994, 2001; Saffer, 2016, 2018) as well as group communication on social networks and social media sites (Duffy, 2003; Page, *et al.*, 2016). Applications of symbolic convergence on social media sites focus more on analysis of visual or symbolic cues (memes, political campaigning posters) as a part of the stories, while

applications for organizational decision-making focus more on group homogeneity and text of statements and stories. By using key concepts from symbolic convergence with the methodology and analysis presented in literature about Brexit on Twitter, this study derives a description of social realities developed as a result of online group communication. This is done by operationalizing the descriptive model on the case of Brexit on Twitter and data collection in the period between the 5th and 12th of September 2019. In the timeframe this study was developed, Brexit's due date was supposed to be on 31st of October in 2019. On the 28th of August 2019, with the Queen's approval, the British prime minister scheduled a prorogation of parliament to take place between the 9th and 12th of September the same year (Woodcock, 2019). This created a lot of turmoil in Britain and the topic of Brexit on Twitter became popular among users on the social media site in this period.

This study will provide a theoretical contribution to social networks and social media sites analyses by using symbolic convergence theory on nowadays online group communication and establishing a model with which we can describe the development of social realities on social media sites. The social realities that social media users develop through communication are a response to real world events. Understanding these responses as stories which were agreed upon and shared by user groups will complement analysis of social movement organizations, activist groups, organizational decision-making, content, and discourse analysis on social media sites. For instance, this could prove useful for analyzing the stance and viewpoint of a user group on a political issue, or the public perception of a celebrity based on the shared stories.

This study will be organized as follows, in the second chapter the theoretical framework is provided, containing the descriptive model and the connection between key concepts and Twitter communication. The methodology is presented in chapter 3, including an operational model and design for data analysis. In chapter 4, the results will include a description of the data collected, sentiment and text analysis, identification topics of discussion and description of the discovered social realities within the data set. chapter 5 gives a discussion of the results and findings, as well as the limitations and implications of this study. And chapter 6 presents the conclusion.

2. Theoretical framework

In this chapter the study aims to develop a descriptive model for the development of social realities by using the framework of symbolic convergence theory. First a look at the theory's roots and development is presented. In the next section of this chapter, the key concepts, structure (or basic paradigmatic assumptions) and method of analysis of the theory's framework are provided. This includes important and seminal applications of the framework. Furthermore, in section 2.2 the connection between Twitter communication and key concepts of the framework are shown. The descriptive model constitutes the final section of this chapter.

2.1 Symbolic convergence theory

Symbolic convergence theory's conception was inspired by a communicative phenomenon that was observed in 1970 by Richard Bales. What Bales (1970) discovered was dramatization, a phenomenon that occurs in group communication when a story of a group member triggers a reaction which leads the whole group to start a discussion revolving around the theme and topic of the *dramatization*. He also provided examples of such messages besides stories, for instance jokes, analogies and metaphors can also trigger a group and cause the dramatizing chain reaction. These findings were picked up by Bormann with which he developed a method of analysis (1972) and eventually a general theory of communication, symbolic convergence theory (1982, 1984, 1985).

This theory is both objective and interpretive (Griffin, 2011, p.13-23), it aims to uncover a communicative truth while using an interpretive method of analysis that stems from rhetoric criticism (Bales, 1970; Burke, 1968,1985; Bormann, 1972). The "communicative truth" that objectifies this theory is the development of social realities and group consciousness it seeks to uncover (Bormann, 1985; Griffin, 2011, pp.247-257). Symbolic convergence received a lot of criticism in the first two decades of development (in the 70's and 80's of the last century) due to the choice of words used to coin key concepts, adopting theoretical concepts from other disciplines and the uncertainty and ambiguity caused by the interpretive method of analysis it used (Bormann *et al.*, 1994, 2001, 2003; Olufowote, 2006; Swanson, 1977). Early in this theory's development, in a reflective view, Swanson (1977) pointed out that the method of analysis (and eventually the theory itself) can present a credible tool to analyze social reality. Total symbolic convergence theory has 18 defined theoretical concepts (Bormann *et al.*, 2001;

Refer to Appendix A). This study will use key concepts that can be applied to communication on Twitter and fit in the scope of the study, focusing on the structure of the theory, the development of social realities, and what constitutes these realities.

Social reality is developed as a result of group communication (Bormann, 1982) and represented by the stories people share in a certain context. Symbolic convergence theory (Bormann, 1985) posits that people in group communication develop social realities and sustain "group consciousness". This theory has a three-part structure (Bormann, 1985, pp. 129-130):

1) "Discovery and arrangement of recurring communicative forms and patterns that indicate the evolution and presence of a shared group consciousness."

2) "Describing dynamic tendencies within communication that explain why group consciousnesses arise, continue, decline, or disappear and the effect they (group consciousnesses) have in terms of meanings, motives and communication in the group context. A basic dynamic is represented by the sharing of group fantasies."

3) "Factors that explain why people share the fantasies they do, when they do."

According to Bormann (1985), group consciousness refers to the interpretation of the collective social realities developed through group communication. For example, his notion of group consciousness represents the common social realities of one or multiple groups in one society. In addition to accounting for communal or societal groups, group consciousness implies the mutual motives, emotions and collective actions of a group. The method of analysis he used to uncover what he called group consciousness and social reality is named fantasy theme analysis (Bormann, 1972). Originally, Bormann (1972) described the stories people share as fantasies. He specifically coined the term fantasy instead of story due to the early focus on "imaginative language" and the analogies, metaphors and jokes a group of people use to describe events. The method of fantasy theme analysis was based on Bales' findings (Bales, 1970; Bormann, 1972). This method was developed around the central idea of dramatism, which Bales used to interpret the emotion, action, motivation and meaning as the rhetoric approach to discourse analysis. Bales' "rhetoric criticism" centered around dramatism as a method of analysis of language used in communication which was developed by the philosopher and rhetoric Kenneth Burke (Burke, 1968; Burke, 1985). Burke mainly focused on how to interpret action and motivation through

emotion dependent on what people say in a given situation or enacted scene (drama as an act of communication). This provided the basis for Bormann and his colleagues to develop an interpretive framework for analysis of how people develop symbolic and moreover social realities.

Based on fantasy theme analysis, Bormann was able to discover themes of stories that were created by sharing stories within a group. He described the social realities developed in this process by categorizing the recurrent themes that emerged by sharing dramatized stories (Bormann, 1972). The categorization scheme was: fantasy themes, rhetoric visions and types of themes and visions. Fantasy types are recurrent themes in the group communication process that emerge from sharing fantasies, which are recounted in detail over a period of time (Bormann, 1972). A "rhetorical vision" contains the shared themes he assumed can be of several generalized types of "fantasies". Most importantly, Bormann (1972) posed critical questions that can serve as a guideline for discovering these themes, which suggest researching should consider certain dramatic elements. Most notable from the list of critical questions (Bormann, 1972, pp. 401-402) suggest investigating typical scenarios, inherent meanings and emotional evocations in the dramas enacted through the stories. As factors that explain why stories are shared, when they are shared, Bormann posited dramatizing messages (1972). He suggests that dramatized stories are fundamental to the development of social realities. Dramatizing messages in Bormann's symbolic convergence theory are aligned with dramatism (Bales, 1970; Burke, 1968; Burke, 1985; Bormann, 1972) and are defined as the emotion, meaning, action and motive that a story describes (Griffin, 2011). Dramatization of stories is the fundamental concept of symbolic convergence, igniting the development of social realities in group communication.

Furthermore, as noted in the introduction, Bormann postulated that from the standpoint of rhetoric criticism there are three types of realities, social, symbolic and materialistic (Bormann *et al.*, 2001). According to them, our everyday life composed of relationships, social status and hierarchies represents our social reality. While we develop and sustain the symbolic reality in order to "make sense of the world", which is composed of ideologies, worldviews and orientations. Finally, the materialistic reality consists of material facts, for instance our physical surroundings, objects, etc. Even though most of the early criticism that symbolic convergence theory received was from the discipline of social constructionism in communication (Bormann *et*

al., 1994, 2001, 2003), social constructionists provide key concepts that bridge the gap between social and symbolic realities. Castor's (2005) statement that people socially construct reality through communication is aligned with how social reality is posited in connected and symbolic convergence theory literature. Moreover, according to Kincaid (1979) a social reality is a symbolic representation of reality. It is symbolic because people use language (as a set of symbols) to share their stories in the communication process. This indicates that use of language and vocabulary will be crucial for text analysis of tweets while using the symbolic convergence framework. In essence, the stories people share in group communication can serve as a "lens for interpretation" (Fisher, 1984; Stache, 2017) for real world events from both actual (i.e. news, argumentation, debate) and fictional (i.e. fantasies) text.

Fantasies that people share in communication are posited as stories when symbolic convergence theory was aligned (Bormann, 1985) with narrative paradigm theory (Fisher 1984). This alignment postulated communal narratives, stories that prevail in a group or community over time and serve as the "means to understand that group's social reality" (Dickerson, 2008, pp. 768-769). The narrative paradigm assumes that people are social beings that communicate by sharing stories. Both symbolic convergence theory and narrative paradigm draw from rhetoric approaches to analyse communication and both have a critical connection with dramatism (Fisher, 1984). Fisher suggests that narration implies use of words that have sequence and meaning for those who interpret or create the stories. In other words, he describes the recurring forms or patterns of communication as a part of the stories, and this concept overlaps with what was postulated by Bormann in his studies of fantasy themes, types, rhetoric visions and group consciousness (Bormann, 1972, 1982, 1984, 1985; Borman *et al.,* 1994, 2001). To clarify, Fisher (1984, p. 7) added that "each of these concepts translates into dramatic stories constituting the fabric of social reality for those who compose them". As characteristics of stories, both theories include plot, scene and characters. Fisher (1984) and Bormann (1985) indicate that the plot represents the action in the language of the story, with a location where that action took place as the scene, and the characters of the story that were part of the described action. The stories that a group shares can consist of recurring forms and patterns of communication, characterized by the use of language, plot, scene and characters. By discovering and arranging forms or patterns, development of social realities can be described, while also accounting for the dynamic tendencies postulated in the 3-part structure of symbolic convergence.

As defined by symbolic convergence theory, the basic "dynamic tendencies" in communication are the sharing of the stories within the group (Bormann, 1985). For example, in the chosen case of Twitter, stories are shared by retweeting a message, which means that the dynamic tendencies in this study are retweets of users' stories. Other dynamic tendencies in communication can be understanding, agreeing and acting on information in the stories (as collective action of a group) (Kincaid, 1979; Bormann, 1985). According to Kincaid (2002), dramatization of events plays an important role in communication, and emotion is a crucial element of dramatization. Duffy (2003) also noted that a social reality can be credibly described when the shared stories address events in a dramatic form. By dramatizing stories about events that had happened or are happening, people come to hold a common image and share meanings (Duffy, 2003). The latter two suggestions (Kincaid, 2002; Duffy, 2003) add to Bormann's (1972) critical dramatic elements, indicating that investigating dramatizing messages should primarily focus on the emotions of the stories. Dramatic elements have also been the focal point of research regarding communication on social media sites with the symbolic convergence framework.

More recent developments (in the 21st century) have used the symbolic convergence theory framework and adaptations of the fantasy theme analysis to discover instances of social realities (as themes, visions and their types) and shared meanings developed through online communication (Duffy, 2003; Page *et al.*, 2016; Page & Duffy, 2018) or organizational decision making (Saffer, 2016; 2018). By researching communication of hate groups online, Dufy (2003) was able to discover the recurrent themes of shared stories, done by comparing of vocabulary sets of bulletin board members that shared the same motivation for hate speech. She concluded with the first attempt to postulate online social realities. According to Duffy's findings "it can be argued that (virtual) rhetorical visions as seen in the study establish the basis for virtual rhetoric communities." (Duffy, 2003, p.309). This notion of online social realities (or virtual rhetoric visions/communities) propelled further research. Similar outcomes, as online group communication that leads to the development of social realities, were result of a study that researched memes (satirical images with text as symbolic expressions). Page et al. (2016) discovered two prevalent social realities in their review of online communication through memes about a famous football player. Furthermore, a similar approach was used to assess the credibility of political figures as depicted in their visual campaigning material posted online

(Page and Duffy, 2018). These three studies, respectively, put their focus on the rhetoric and interpretative approach to discovering social realities and researching dramatic elements. Alternative approaches and uses of the framework are presented in the studies of Saffer (2016, 2018). He used the symbolic convergence theory to uncover shared meaning in organizational settings (Saffer, 2016) and as an outcome of engaging in organizational decision-making processes (Saffer, 2018). In his studies, Saffer adapts the fantasy theme analysis method to an "evaluative theme analysis" which allowed him to discover shared meanings. Shared meaning in his study was operationalized as the similar, cohesive statements that participants effectively shared and agreed on. Implying that in the process of group communication and converging towards a "common image" people share meanings, actions, motives and emotions as a part of their social realities (Bormann, 1972; 1984; Saffer, 2016).

To describe social realities, this study focuses on the key concepts that constitute social reality in online group communication. According to Bormann (1985) and Fisher (1984) the basic concept is the stories people share when communicating, characterized by the plot, scene and characters and containing recurring forms or patterns of communication. Furthermore, the use of language and vocabulary of the stories are important to categorize as forms or patterns of communication. Based on the suggestions Duffy (2003) and Kincaid (2002) dramatized stories in this study are defined as the emotion shown in the content of stories. While the basic dynamic tendency in group communication is sharing of stories (Bormann, 1985; Bormann *et al.*, 2001) which leads to the development of topics (as fantasy themes) and themes of discussion (as rhetorical visions). Shared meaning is defined as the common image people come to hold (Duffy, 2003; Olufowote, 2017; Saffer, 2016) that overlaps between a story and a topic or theme that story developed. Finally, social reality is represented symbolically using language (Kincaid, 1979; Castor, 2005) and consists of the dramatized stories, topics and themes of discussion. In the next section of the chapter, the connection between these theoretical concepts and communication on Twitter is presented.

2.2 Key concepts and connection to Twitter

The key concepts of symbolic convergence theory that are applicable to the case of this study are in the left column of table 1. While the overlap with communication on Twitter is

shown in the right column. These key concepts and their characteristics will be used in a descriptive model in order to describe the development of social reality on Twitter.

Table 1. Symbolic convergence and communication on Twitter

Theoretical concepts	*Twitter communication overlap*
1. Stories, that contain forms or patterns of communication in their content. Characterized by use of language, plot, scene and characters (Bormann, 1972, 1985; Bormann *et al.*, 2001; Fisher, 1984; Stache, 2018, pp. 576-578)	1. Tweets 1.1. Content characterized by the: 1.1.1. Use of language 1.1.2. Use of hashtags 1.1.3. Number of retweets, favorites & replies
2. Dramatized stories as the emotion, action, motive and meaning of a story. The most important characteristic of dramatizing stories for this study are the emotional "evocations". (Bormann, 1972, 1985; Bormann *et al.*, 2001; Duffy, 2003; Griffin, 2011; Kincaid, 2002; Dickerson, 2008, pp.768-769)	2. Tweets are dramatized with an emotional load 2.1. Emotions in tweets are measured by: 2.1.1. Sentiment of the tweet 2.1.2. Intensity of the sentiment
3. Sharing of dramatized stories, defined as the dynamic tendency in the storytelling process. (Bales, 1970; Bormann, 1972, 1985; Bormann *et al.*, 2001; Fisher, 1984; Stache, 2018, pp. 576-578)	3. Dynamic tendencies as: 3.1. Retweeting dramatized tweets 3.2. Favoriting dramatized tweets 3.3. Replying to dramatized tweets
4. Rhetorical vision, as a prevalent theme of a group or community that can contain more than one topic and holds shared meanings (Bormann, 1972, 1985; Bormann *et al.*, 2001; Olufowote, 2006, 2017)	4. Themes of discussion on Twitter 4.1. "#Brexit"
5. (Fantasy) Themes of dramatized stories, defined as the dramatized stories that are repeatedly shared within a group. (Bormann, 1972, 1985; Bormann *et al.*, 2001; Olufowote, 2006, 2017)	5. Topics of discussion within the theme "#Brexit" 5.1. Developed by repeatedly retweeting a dramatized tweet (by multiple users)
6. Shared meaning, defined as the meaning a group shares on a particular theme (as a complementary concept to [fantasy] themes). Bormann, 1972, 1985; Bormann *et al.*, 2001; Griffin, 2011; Olufowote, 2006, 2017)	6. The "common image" users come to hold when they take part in the development of topics (i.e. Users that retweeted a dramatized tweet, which eventually developed into a topic of discussion, share the "nascent" meaning of the topic)

This study considers that the process of communication is storytelling. The tweets that users share are the stories which content is characterized by the use of language, use of hashtags and tweet attributes (favorites, retweets, replies). In addition to this, the plot, scene and characters of a story add to the context of the tweet. Hash-tagged words on Twitter commonly

represent the keywords of a tweet and can become the hashtags for topics and themes of discussion. Further, dramatized tweets show emotions in their text which will be determined by the sentiment and intensity of the stories. The dynamic tendency of sharing stories on Twitter is retweeting. However, this social media site also offers users to favorite and reply to stories which can also represent dynamic tendencies. Rhetorical vision in the chosen case is the theme of discussion "#Brexit". While, the topics of discussion within "#Brexit" are the fantasy themes in symbolic convergence theory. Lastly, shared meaning is the common image a group shares has about a topic within this theme, which can be discovered by analyzing the topics of discussion. Although this study has chosen "#Brexit" on Twitter as the case to apply symbolic convergence theory and develop a descriptive model, this theory's concepts can also be aligned with most, if not all, social media sites used today (i.e. Facebook, Reddit, YouTube, etc.). What all social media sites have in common is the connectivity between users in a networked environment that allows them (the users) to post and respond to content, essentially sharing their stories with each other. In addition, all social media sites allow their users to react to content, in the form of likes, favorites, shares or reposts (or retweets in the chosen case) and replies. The theoretical descriptive model presented in the next section is inclusive of all the theoretical concepts that can be aligned with social media site communication and is operationalized for the case of "#Brexit" on Twitter in chapter 3.

2.3 Descriptive model

According to Fisher (1984, p.7), in order to describe social reality, we need to understand "what constitutes the fabric" of social realities. While, Bormann *et al.* (2001) indicate that researching social reality of a group or social realities within a theme should not exclude the symbolic reality. Elaborating further, they suggest that groups of people interact symbolically and a group's symbolic reality accounts for their social reality. By taking the same approach to answer the posed research question, the developed descriptive model (figure 1) implements the theoretical concepts that allow for the description of social realities by analyzing tweets. Stories will represent the tweets, characterized by language and hashtags used, what they describe (plot, scene and characters). From the aspect of narrative paradigm, stories serve as a lens for interpretation of events. The sharing of stories is the dynamic tendency in storytelling (Fisher, 1984; Bormann 1985). On Twitter sharing is represented by retweeting. Theoretically, on this social media site, favoriting and replying to tweets can be considered as dynamic tendencies. In

addition, dramatization of tweets leads to dynamic tendencies (or retweeting). On the other hand, for different social media sites, sharing (Facebook) and reposting (Reddit) represent one concept of dynamic tendencies in group communication. Other dynamic tendencies, for instance on Facebook, are represented by liking or reacting and replying to a story. While on Reddit, besides replies to stories, the social media site allows users to upvote (or like) and downvote (or dislike) a story. With a focus on the emotions and emotional load of the story, dramatization implies the sentiment and intensity of stories. However, on social media sites today stories can contain emojis or emoticons, as well as memes, videos and gifs to portray user's emotions complementary to the text of the story. In other words, stories are dramatized by sharing emotions in a variety of ways, not only through what users say in their text but also through visual cues. When dramatized stories are shared repeatedly within a group, they develop topics of discussion. These topics are a part of the theme of discussion (in this case "#Brexit") and have the common characteristics of the stories that developed them, in terms of plot, scene and characters. Furthermore, the meaning of a dramatized story that developed into a topic is shared within the group that was part of that discussion.

This shared meaning should overlap between the story that developed into a topic and the topic itself. Olufowote (2017) implies that discovering shared meaning can be done by analyzing the content of the recurrent topics or themes and comparing it to the stories that are repeatedly shared. On Twitter, topics can be discovered by the retweet count of stories. Additionally, the number of favorites and replies to a story could be credible measurement for discovering topics on Twitter. However, retweets are the most important measurement for discovering topics in the chosen case, because they directly represent the sharing of stories (Grčar *et al.*, 2017). On other social media sites, for instance YouTube, themes of discussion are already categorized by the social media site itself (as YouTube topic, i.e. basketball), however topics within the themes can be discovered by the view counts of a (video) story, the count of replies, likes or dislikes, and shares. Very similarly, on Facebook topic discovery can be done by analyzing the count of likes or reactions to stories, as well as the count of shares and replies.

Although shared meaning is a characteristic part of social reality, this study will partially grasp the meanings shared in "#Brexit". According to Saffer (2016) discovering shared meaning requires a homogenous group engaged in a network of relationships. In addition, for adequate

representation of a group's shared meaning he strongly suggests using the contemporary evaluative theme analysis. This is further addressed in section 3.1 and discussed in chapter 5. The model is operationalized in the next chapter and adjusted to fit with the scope and case of this study contrasting the (theoretical) descriptive model in Figure 1.

Figure 1. Descriptive model of social reality on social media sites

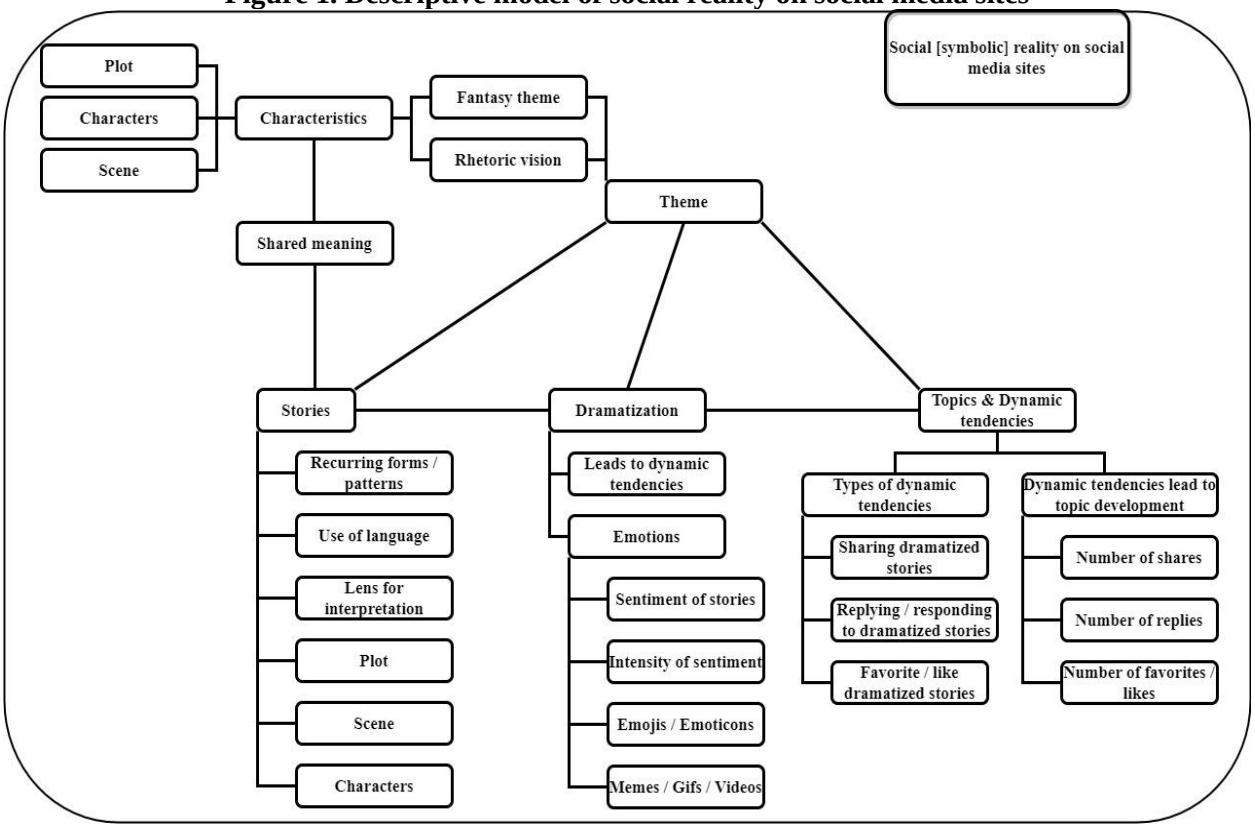


Figure 1. Note: This model is proposed by this study.

For example, with this model a social reality as a result of online group communication can be described through the theme created in the storytelling process that contains multiple topics. It is characterized by the number of times a story, that developed into a topic or theme, has been shared by a group, the meaning a topic or theme implies, as well as the emotions of the story. For instance, in the chosen case of "#Brexit", a story with negative sentiment, retweeted thousands of times by a group which talks about the economic situation of Britain would characterize one social reality within the theme. While a story with positive sentiment, also retweeted thousands of times, about leaving the European Union would characterize another social reality of (possibly) a different group of users.

3. Method

In this chapter the methodology, operational model, and the design for data analysis for this study are presented. Firstly, the methodology of this study is based on research on Twitter (Grčar *et al.*, 2017; Hall *et al.*, 2018) (specifically on the theme Brexit) in addition to methods and applications of symbolic convergence theory (Bormann, 1985; Borman *et al.*, 2001; Duffy, 2003) and overlapping procedures (Castor, 2005; Fisher, 1984; Page *et al.*, 2016; Page & Duffy, 2018). Further, description of social reality on Twitter will require designing an operational model that fits communication on the social media platform. In other words, the concepts from symbolic convergence theory need to coincide with the flow of communication on Twitter. Lastly in this chapter, data analysis is designed based on text analysis and text mining methods (Silge & Robinson, 2018), sentiment analysis (Hurlimand *et al.*, 2016; Jockers, 2017; Nielsen, 2011; Niklander, 2017), topic discovery and network analysis (Grčar *et al.*, 2017; Hall *et al.*, 2018; Quraishi *et al.*, 2018).

This study proceeds with a mixed research method. For research on Twitter data Grčar *et al.* (2017) suggest that quantitative methods should be used: measurements such as retweet counts, sentiment scores and hashtag counts (Grčar *et al.*,2017; Llewellyn & Cram, 2016) and gaining general insight about emotions from sentiment analysis on the whole collected dataset (Hall *et al.*, 2018). In addition to quantitative methods for analyzing Twitter data, Hall *et al.* (2018, p. 25) state that for a high level overview of online group communication, research needs to "drill down" into the content to discover what is shared and discussed. This implies qualitative methods and description of the content from the data. Niklander (2017) indicates that it is important to add context to the sentiment and qualitatively analyze the discourse on Twitter. For the qualitative part of the analysis, this study will use the interpretive framework of symbolic convergence and the overlapping concepts from the narrative paradigm presented in chapter 2. Tweets will be treated as stories, with a plot, scene and characters. The use of language will be analyzed by arranging patterns from the text of the tweets, including hashtags. Selected stories that show relatively high sentiment intensity (both positive and negative) and stories that developed into topics of discussion will be put through the sieve of the theoretical framework. With this interpretive approach, this study will be able to show the emotional load of stories, the development of stories into topics, and to describe the uncovered social realities within

"#Brexit". The following two sections show how the theoretical concepts fit with Twitter communication and data analysis methods.

3.1 Operational model

Contrasting the descriptive model in figure 1 (chapter 2), operationalization of the model required a fit between the concepts from symbolic convergence theory and Twitter communication presented in figure 2. The recurring forms or patterns of communication fit as a characteristic of tweets, as the use of language and hashtags. By categorizing these forms and patterns as a part of the shared stories this study will determine vocabulary sets (words and hashtags in the tweets) and discover the topics that developed in the theme. Although hashtags on Twitter start off as keywords, when they become popular, they develop into topics of discussion (Xiong *et al.*, 2019) as well as themes of discussion, as in the case of this study. The concept of dramatization is determined and measured by the kind of sentiment and intensity of sentiment, respectively. From the aspect of symbolic convergence there is no exact measurement of dramatization,. Moreover, conclusions drawn by using this theory are dependent on researcher insight rather than robustness or application of the theory (Dickerson, 2008, pp. 768-769). Dramatization so far has been interpreted based on the emotion, meaning, action, and motivation portrayed in a story (Bormann *et al.*, 2001; Duffy, 2003; Griffin, 2011; Saffer, 2016; Page *et al.*, 2018). However, by using sentiment analysis and measured intensity of sentiment, this study will both quantitatively and qualitatively analyze the emotions portrayed in stories. According to Duffy's (2003) and Kincaid's (2002) suggestions, emotions are the crucial dramatic element. This study limits dramatization to emotions in stories on Twitter as measurable by sentiment analysis. On this social media site, retweeting is effectively sharing of a tweet which fits with the concept of dynamic tendency in communication as sharing a story. Grčar et al. (2017) find that retweets are more than just simply sharing, but are also agreeing with a tweet. This proves retweets are very valuable when researching groups on Twitter, because of users' tendencies to share and agree with a story they develop a topic of discussion through retweets. A story itself, is born at the hands of the user as a tweet. In accordance with symbolic convergence, a user dramatizes their story, showing emotions when describing an event that is characterized by what they write, what hashtags they use, also how they portray the scene, plot and characters. This dramatized story is tweeted out and picked up by a part of the Twitter community. Twitter offers

Figure 2. Operational model for social realities developed on Twitter

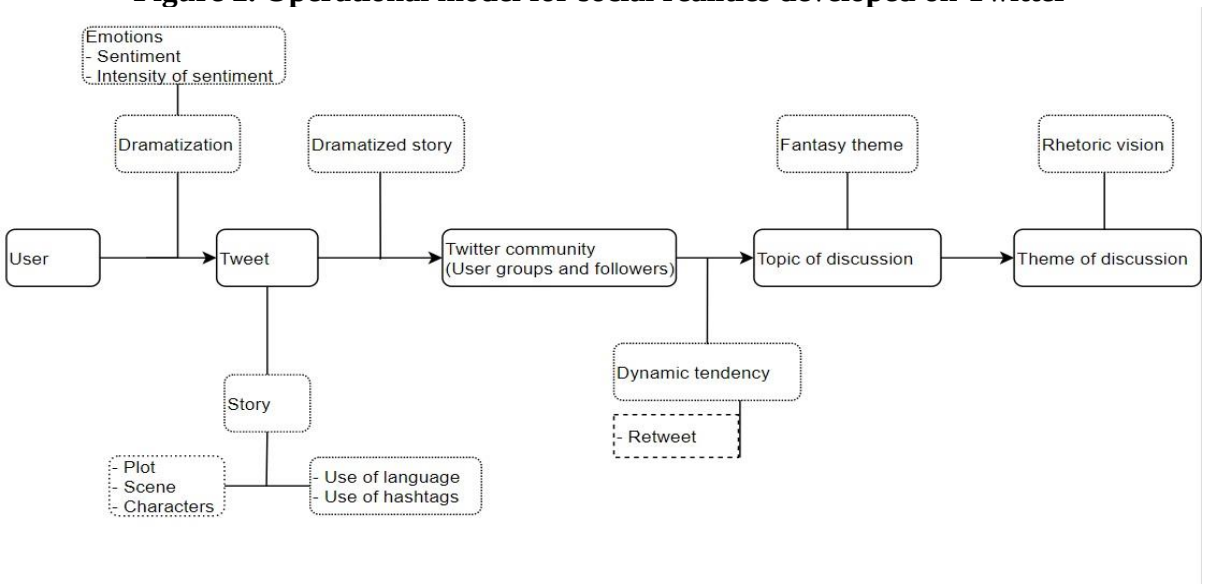


Figure 2. Note: Arrows represent the linear flow of stories / tweets on the social media site Twitter. Boxes with dotted outline represent the theoretical concepts.

its users the possibility of following each other, following trending topics and trending hashtags, thus allowing users to build a network of relationships with other users. So, a Twitter community is composed of groups of users following a certain topic, theme, or another user. For instance, the 10,000 people that follow "#Brexit" and all retweet the same story form a community that shares a topic of discussion and agrees with the same tweet. If this tweet happens to be with negative sentiment intensity that referred to the referendum with negative emotions and was shared or agreed upon in a community, it will typify one social reality on the theme Brexit.

Shared meaning was excluded from the operational model (in figure 2 above) due to the limitatyions on uncovering shared meanings. In other words, grasping shared meaning as a wholistic concept presented in symbolic convergence requires fantasy or evaluative theme analysis (Bormann, 1972, 1985; Saffer, 2016), which is outside of this study's scope. Both methods of theme analysis require supervised classification of data, such as manual annotation or labeling of data, intercoder agreement and cross validation. Further explication why this study chose not to use either fantasy or theme analysis is predetermination of the theme of discussion (Brexit). By taking an alternative route to the method of analysis in symbolic convergence theory, this study is researching social reality top to bottom, starting from the theme "#Brexit".

This will be done by using unsupervised data classification on the text and hashtags of the tweets and using insights from studies that have previously researched Brexit on Twitter.

3.2 Data analysis

For the description of social realities developed through Twitter communication this study will use methods and procedures presented in research on Twitter data and text mining. Data analysis is based on the text mining procedures presented in Silge and Robinson (2019) and the standard proposed by Hall *et al.* (2018) (refer also to appendices B & C). For analysis of social media in or during political events, Hall *et al.* (2018) propose a standard procedure that includes a ,general overview of the discussions over time, sentiment analysis, text analysis, discovery of topics, network of relationships between Twitter users and visualization. This study follows the procedure accordingly and adds the interpretive methods from symbolic convergence theory. This will include:

1) Text mining and analysis (Silge & Robinson, 2019), for discovery of patterns of communication and data manipulation (further addressed in 3.2.2).

2) Categorizing the most popular hashtags in the data set to discover topics of discussion (Bassilakis *et al.*, 2018; Grčar *et al.*, 2017; Llewellyn & Cram, 2016).

3) Sentiment analysis and classification of tweets (Grčar *et al.*, 2017; Hall *et al.*, 2018; Niklander, 2017; Silge & Robinson, 2019).

4) Discovering stories that a group of people shared and developed into a topic by retweets that used one of the popular hashtags. Visualization of a network of retweets (Grčar *et al.*, 2017; Hall *et al.*, 2018) from one of the most prevalent topics in the data set, determined by the popular hashtags.

5) Experimental Latent Dirichlet Allocation (non-parametric LDA) topic modeling (Graham & Ackland, 2015), with randomized Gibbs sampling.

Text mining and text analysis will be used to measure the relative term frequency of words and hashtags from the data set of tweets that will be collected. Additionally, words will be weighted depending on how often they appear in each tweet and across all tweets. This will provide an inverse document term document frequency score, the words with the highest score

will be the terms that characterize the stories on the theme of discussion. Text mining methods also will allow categorization of the most popular hashtags to discover the topics developed in the data set. Silge and Robinson (2019) in the text mining procedures also provide steps for the cleaning and manipulation of data sets, as well as preparing the data for sentiment analysis.

For dramatization score and classification of text this study will use sentiment analysis. The AFINN sentiment lexicon was chosen for this task (Nielsen, 2011; Jockers, 2017) due to the intensity of sentiment it provides. In this lexicon words carry a sentiment intensity between -5 (extremely negative) and 5 (extremely positive). The method that will be used in conjunction with this lexicon is bag of words. This is an unsupervised, and arguably raw, method for sentiment analysis and classification. With bag of words, each word of each tweet is scored separately and an aggregate score for the sentiment of each tweet is provided. An unsupervised method of classification was chosen due to the insight provided from previous studies about Brexit on Twitter (Bassilakis *et al.*, 2018*;* Grčar *et al.*, 2017; Hall *et al.*, 2018; Llewellyn & Cram, 2016). All these studies show that discussions on this theme on Twitter are mostly with negative sentiment. By using unsupervised methods, this study will be able to show if this theme of discussion remained negative over time. Niklander (2017) argues that these methods present a challenge because of the inability to determine false positives or false negatives. However, with this study's mixed method, selected tweets will be analyzed with the framework of symbolic convergence and it will be possible to interpret the emotions in the stories. With this sentiment analysis, tweets will be classified as one of 3 types, negative, positive and undecided.

Besides classification according to sentiment scores, this study also discovers topics of discussion through the popular hashtags in the data set. The most popular hashtags will present different topics in which this study expects groups of users to develop different social realities. Through retweet counts the stories, that users shared and agreed with the most, can be graphed in a network of relationships. Grčar *et al.* (2017) suggest that retweet communities provide an overview of the discussion by linking users who agree on certain topics. The network will show how topics developed and which stories became topics of discussion. This will be combined with the sentiment classification and communities will be linked together based on the sentiment of the tweets they shared. For this task, this study will select one of the most popular hashtags and visualize a network of retweets that used the same hashtag. With visualization of a network of

retweets that all contain the same hashtag as the topic of discussion an accurate description of social realities developed through online communication will be achieved. The communities in this network will be discovered by graphing the data in Gephi (Bastian, Heymann & Jacomy, 2009) and using an algorithm based on the Louvian method (Blondel *et al.*, 2008). This algorithm computes a score which represents the network's connection strength in terms of density and sparsity as modularity, from -1(sparse) to 1(dense). A high modularity score means that users in the network are densely connected within their communities, but loosely connected to other users in other communities. In addition to density, this algorithm will detect the number of communities in the network from the selected Twitter data set as modular classes, providing a modular class number for each community. Stories that were shared in the largest communities from this set and that were effectively agreed upon by hundreds or thousands of users will be selected for further analysis as typical for developed social realities.

For a general overview of the topics of discussion that characterize developed social realities, this study will use a Latent Dirichlet Allocation topic model. This experimental method for discovering topics generates a random number of documents dependent on the size of the data set. It is commonly used to uncover *latent* variables from text (among many other uses for non-parametric generative models; see also: Hong & Davison, 2010; Wood, 2014). In this study's case, latent variables in the text will be the words that characterize a document. With this method, tweets are the documents that contain terms (or words and hashtags) and the collected data set is a corpus of documents. Essentially, the data set will be manipulated into a *document-term matrix*. Words will be weighted according to their frequency of use, respective ranking equivalent to their frequency, how many times they appear in each document (tweet) and across documents. This model will output either the maximum number of topics or a set number of topics. The model will be limited to 4 topics in order to gain insight concerning 4 different social realities from a sample of the data set. Without placing a limit on the output of the model, it will generate as many topics as it can compile from the text of the tweets. This limitation will allow a description of 4 topics of discussion that are typical for the social realities developed on "#Brexit". The detailed procedure for a true randomized, generative LDA topic model with Gibbs sampling for Twitter data is available in Graham and Ackland (2015). The precision of this model for use on text and corpuses is heavily debated; however, with recent advances (in the past decade) it has been optimized to work very efficiently with short text, such as tweets.

Newman *et al.* (2011) argued that LDA topic modeling requires regularization parameters for short text; however, with further optimization Buntime and Mishra (2014) prove the consistency of Gibbs sampling with this method. Seeing how tweets are limited to 240-word characters and span up to 300 total character spaces, Twitter users need to come across clearly for their story to be seen, retweeted and make their opinion known. That means the topic model will provide a general overview of the 4 biggest topics based on the short text from collected Twitter data.

The presented procedures for data analysis will be done with R (R Core Team, 2018) and RStudio (RStudio Team, 2016), that provide the tools for statistical computing methods and the means for data collection. Additionally, for graphing the networks of retweets and the discovery of communities, this study will use Gephi (Bastian *et al.*, 2009). The latter provides the algorithm for discovering communities and network visualizations. By using presented methods and procedures in this chapter, this study will provides a description of developed social realities as a result of group communication on Twitter.

3.2.1 Data collection

While the British referendum, Brexit, was unravelling towards the end of 2019, the Queen and prime minister of England scheduled a prorogation of the longest parliamentary session in British history to take place between 9th and 12th of September 2019. This unforeseen prorogation ignited a protest on the streets of London, a showing of total lack of professionalism by British politicians in televised parliamentary sessions, talks of a “no deal Brexit” which implied a rushed exit from the European Union, adding chaos to the global ramifications in all of the social, economic and political domains which the exit from the Union already implied. The British people, politicians and media took their debates to and voiced their opinions on Twitter. This study will collect Twitter data from the 5th to the 12th of September. Data collection in this period will capture responses of Twitter users to the mentioned real-world events. For collecting the data an open Twitter API will be used with an interface in R. Retweets will be included in the collection procedure and the data will be collected on the hashtag Brexit (as the theme of discussion) in English. However, the open Twitter API poses restrictions and limitations for data collection. This free API allows collection of only a random portion of publicly available tweets in the last 24 hours (from the moment of data collection). With this collection method, it will be

impossible to capture all topics of discussion as a whole in "#Brexit", and some popular stories might sneak in as a retweet without capturing the original story.

3.2.2 Data manipulation and cleaning

Text in tweets contains special characters, URLs, emoticons or emojis, numbers and punctuation marks. In order to analyze the text from the tweets, all the above were cleaned, as were stop words and personal pronouns in English (words such as: to, the, was, I, etc.). Although stop words and personal pronouns add to the context of a whole tweet, they are the most used words and do not carry meaning by themselves. Stop words will be of no use for the analysis with the bag of words method. Manipulation of data sets implies conversion between formats and data structures, for instance converting a data set of tweets into a document-term matrix which will be used for topic modeling, switching between text format and a table layout, graphing networks of tweets, visualization of data, etc. Cleaning the text of the tweets and data manipulation will be done by using the "tidyverse" (Wickham, 2017) and "tidytext" (Silge and Robinson, 2019) packages in R. In addition, data manipulation and cleaning change the shape of the collected data, particularly the text of stories in the chosen case. For instance, the average character length of collected stories was 170 characters before cleaning and 151 characters after cleaning. Also, all emojis and links to memes, gifs, or videos are removed by cleaning the text of the tweets. This heavily impacts the content of the stories; however, this study focuses on text analysis and encoded emojis or links to visual media are not directly beneficial to the analysis.

4. Results

In this chapter description of the data is followed by the results of the data analysis, and the description of developed social realities in the theme "#Brexit" on Twitter. This includes an overview of the collected data and the sentiment scores of tweets over time. This is followed by the patterns of communication in the stories presented as vocabulary sets that contain words and hashtags. In addition to the overview of the stories, vocabulary sets will be categorized according to sentiment analysis and unsupervised classification. Finally, this chapter presents discovered social realities by visualizing one of the most popular topics in the data set. From the analysis this study expects that most, if not all, patterns of communication found in previous research about Brexit on Twitter (Grčar *et al.*, 2017; Hall *et al.*, 2018) will recur in the collected set. Considering the prorogation of parliament, the brewing protests, and the bitter debate during the

data collection period, this study expects that dramatized stories about these real-world events will create topics of discussion and develop social realities.

4.1 Data description

During one week of data collection (05.09.2019-12.09.2019) 279,199 tweets by 103,828 users were captured in the English language with "#Brexit", as seen from table 1. From which 53,912 were unique tweets by 26,449 users. While 225,287 were retweets by 77, 379 users. From the total collected tweets, on average a tweet received 904 retweets (presented in table 2); however, this includes tweets that were not retweeted. By adjusting the data set and removing the not retweeted stories, the summary for retweet counts shows that the average retweet count per tweet is 1065. Despite the indication that in the collected data set a lot of stories were heavily shared and agreed upon, creating many communities with different social realities, the difference between the minimum and maximum retweets is more than 10 standard deviations apart. This means that in the collected data set, there are outliers that have a much larger retweet count than the usual or average retweet count. This can be due to the means of collection and having tweets *sneak in* (Grčar *et al.*, 2017) without capturing the whole discussion revolving around those stories or even without capturing the original story. For the distribution of tweets per day refer to appendix D. The size of the data set is relatively small compared to previous research on this case (Grčar *et al.*, 2017; Hall *et al.*, 2018); however, this study shows that this sample is large enough to describe the developed social realities. A description is accomplished through analysis of patterns of communication and dramatization of stories in the data set.

Table 1. Collected data

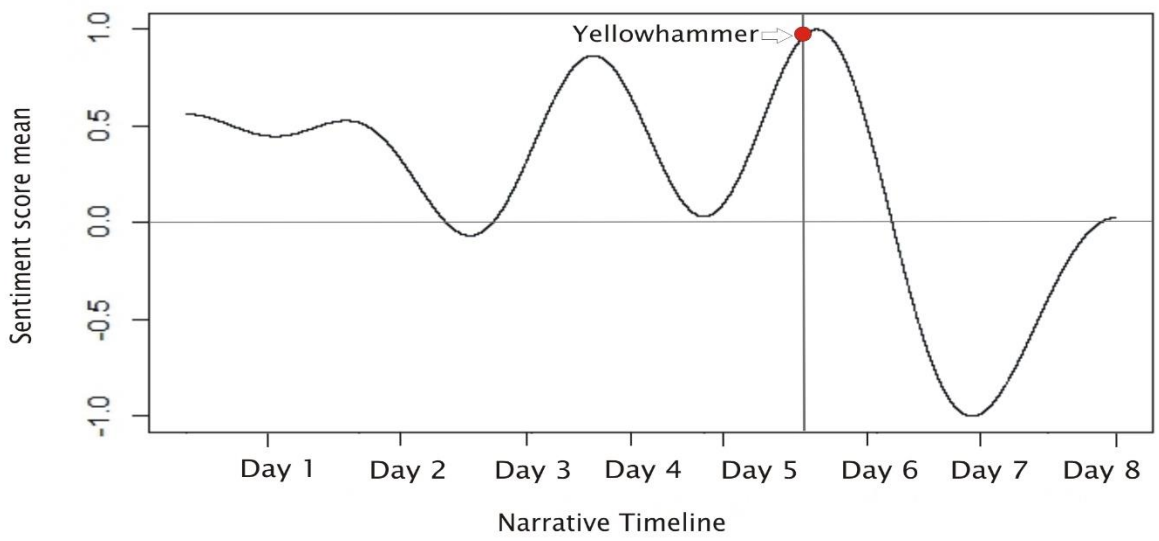
	Total collected	*Unique tweets*	*Retweets*
Tweets	279,199	53,912	225,287
Users	103,828	26,449	77,379

4.1.1 Sentiment analysis and overview of dramatization

Dramatization of stories was focused on the emotions portrayed in the tweets measured by sentiment analysis and the intensity of sentiment. For this task the AFINN sentiment lexicon was used; it provides a collection of 2,477 words that carry a weight based on their intensity of sentiment. The scores range between -5 and 5. This lexicon was used in conjunction with a bag of words method, providing a raw intensity score for each word in each tweet and returning a sum of the raw score. Tweets were classified in three sets, negative, undecided and positive, depending on their score. From the sample of unique tweets (in figure 4) it is noticeable that the

stories in the collected data set started with a positive sentiment intensity and over time became more negative. The *y* axis in the figure represents the mean of sentiment scores over the narrative time relative to the period of collection on the *x* axis. The sentiment scores in figure 4 were scaled to -1 (completely negative) and 1 (completely positive).

Figure 4. Sentiment of #Brexit stories over time



This graph shows how the discussions on "#Brexit" developed over time. For the first 5 days of data collection the overall sentiment of stories was somewhat positive or close to undecided. However, before day 6 of data collection there was a narrative turning point that is represented as a downwards dip, of the line in figure 4, to completely negative sentiment. During that period (around 9th of September 2019) the Yellowhammer operational report was discussed in parliament and later leaked to the British public on the 11th of September. This report dealt with the possibility of Britain leaving the European Union with no trade deals, which implied no financial, medical, or military support from European Union member states. Twitter users responded to those events by sharing negatively dramatized stories which in turn shows overwhelming negative sentiment between the 4^{5h} and 7th day of data collection. As the discussion about Brexit on Twitter evolved over time, the figure above shows that in the sample of unique tweets users responded with negatively dramatized stories to real-world events. The raw average of the sentiment score from this sample is -0.4, while for all the captured tweets the average is -0.5 (refer to table 2-1 in appendix E). The raw mean (without scaling to -1,1) of the

overall collected data compared to the raw mean of the sample of unique collected tweets indicates that in this case negatively dramatized stories were shared more times than stories that were classified as positive or undecided.

Table 3. Classified tweets

Sentiment	*Classified total collected tweets*	*Classified unique tweets*
Positive	98,602	16,783
Undecided	59,767	15,113
Negative	120,830	22,016

From the sample of unique tweets 16,783 were classified as positive, 15,113 as undecided and 22,016 as negative, as shown in the right column of table 3 above. While from the total collected data set including the retweets, 120,830 were classified as negative, 98,602 as positive and 59,767 as undecided. In this data set, from the unsupervised classification method based on sentiment analysis, it is apparent that negatively dramatized stories dominated the Brexit discussion on Twitter.

4.2 Patterns of communication

Stories are a basic form of communication according to narrative paradigm (Fisher, 1984; Stache, 2017) and symbolic convergence (Bormann, 1985; Bormann *et al.*, 2001). In the chosen case of this study stories are represented by the tweets. The patterns of communication from the stories were uncovered by analysis of their text, namely what words were used, how often they were used, as well as the used hashtags. In addition to frequency of terms, this study also uses *term frequency inverse document frequency (tf-idf)* as a measurement for the weighting of words that characterize the data set. This will be done by following the procedure and methodology presented by Silge and Robertson (2019). Inverse document frequency shows how often a word appeared in each tweet and across all the collected tweets, based on the relative term frequency and the inverse scores, “tf-idf” was calculated. Although this method has been proven to work for large corpuses of text data or documents (e.g., a collection of books, a corpus of news articles, etc.), Newman *et al.* (2011) show that term frequency and inverse document frequency are valuable measurements when analyzing short text such as tweets. This whole part of the procedure resulted in categorizing vocabulary sets from the whole collected data set, as well as the vocabulary of each classified group by sentiment analysis (positive, negative, and undecided). The categorized vocabularies include patterns of communication discovered from the use of language, use of hashtags and dramatization. Previous research on the theme Brexit on Twitter finds that users of this social media site have specific “Brexit vocabularies”. This

includes patterns such as: “revoke article 50” or “revokea50”, “leave” or “leavers”, “remain” or “remainers”, “reesmogg”, “nigelfarag[e]” (Grčar *et al.*, 2017), among other Brexit-specific terminology. It is expected that these patterns will recur in the collected data set and will characterize the vocabulary of users on this theme.

In tables 4, and 5 the vocabulary of the whole collected data set is presented by arranging the top 15 frequent words (as terms) and top 10 popular hashtags. The left half of table 4 shows the frequent words from the whole collected dataset of tweets, while the right half shows the frequent words from the sample of unique tweets. From the side to side comparison, it is noticeable that there is not much difference in which words were the most frequent between the sample of unique tweets and the whole data set. As is expected from this discussion, the top 3 most frequent words are “brexit”, “eu”, and “uk”, because of the theme of discussion “#Brexit”. Every user shared a story that referred to the referendum and the United Kingdom leaving the European union. For further analysis of frequent terms these three words were excluded, since they will always be the 3 most popular terms in this data set and are “common sense” given the theme of the discussion. From the rest of the frequent terms, the relative term frequency indicates that stories revolved around the British prime minister, Boris Johnson, the parliamentary sessions that were scheduled to end in the period of collection, and the debate of deal or no deal Brexit. In addition to these words, Brexit-specific vocabulary known from previous research on this topic appears in this data set as well. The words deal, nodeal, remain, and leave are such examples. Due to the scheduled prorogation of parliament and the upcoming deadline of Brexit (31st of October) at that time, users responded to these real-world events on Twitter.

Table 4. Frequent terms

All collected Tweets			*Collected unique tweets*		
Word	Number of times used	Relative term frequency	Word	Number of times used	Relative term frequency
brexit	318,683	0.082	brexit	59,004	0.095
eu	46,615	0.012	eu	7,170	0.012
uk	40,080	0.010	uk	6,424	0.010
people	39,462	0.010	people	5,150	0.008
parliament	28,218	0.007	deal	4,699	0.007
deal	26,645	0.007	boris	4,215	0.007
boris	26,162	0.007	leave	3,877	0.006
leave	25,376	0.007	parliament	3,651	0.006
voted	22,366	0.006	vote	3,300	0.005
johnson	21,780	0.006	remain	2,713	0.004
remain	21,342	0.005	johnson	2,535	0.004
referendum	20,980	0.005	nodeal	2,484	0.004
party	18,011	0.005	election	2,460	0.004
vote	17,971	0.005	borisjohnson	2,388	0.004
election	17,587	0.005	yellowhammer	2,318	0.004

**Note: The whole collected data set of tweets contained 50,013 unique words from a total of 3,890,362 words; The sample of collected unique tweets contained 45,653 unique words from a total of 622,685 words.*

Notably from the comparison of the top 15 frequent terms, in the sample of unique tweets the word yellowhammer appears 2,318 times. However, this word is not part of the top 15 in the whole data set with retweets. Yellowhammer refers to an internal report of the British government that was not fully known to the public until September 2019. This report dealt with the option of Britain exiting the European union without any deal whatsoever. It implied no trade deals, no military or security options, and no import of medical supplies from Europe. The yellowhammer report caused a divide in the response to British parliamentary proceedings when it surfaced and was made known to the public. The publicly televised parliamentary sessions during this period looked more like a dog and pony show, with politicians filibustering (talking for the sake of talking) for hours, laying down in the middle of sessions, refusing to come to a resolution whether Brexit was going to end with a deal or not. This study expects that the stories that addressed the yellowhammer report will show negative emotions and created a negative topic of discussion within the theme. Table 5 shows the most popular hashtags relative to the collected data set. As expected, "brexit" is the most used hashtag across the whole data set, because the data was collected with that hashtag. In the sample of unique tweets, it is apparent that the developed topics of discussion dealt with the European Union, the British prime minister, and the afore mentioned yellowhammer report. In addition to these three prevailing hashtags, in the set of all collected tweets, "nodeal" and the variation "nodealbrexit" dominated the stories. Although, this is not true for the hashtags in the unique tweets. In the right half of table 5, "brexitshambles", "brexitchaos" and "bbcqt" (referring to BBC news) are amongst the most popular hashtags in the set. This indicates that from the collected data set, although thousands of unique stories were created with these 3 hashtags, they were not shared as much as stories with the hashtags "eu", "yellowhammer", "borisjohnson" and "nodeal" or "nodealbrexit". In addition to this apparent difference in popular hashtags by comparison, in this collected data set, more users shared and effectively agreed with the hashtag "remain" than "leave", although in real life the British public was decisive in its vote to leave the European Union. Prorogation was also a popular hashtag in the data set as a whole, being used 3,579 counting the retweets, as the end of one of the longest parliamentary sessions in British history was getting closer.

Table 5. Popular hashtags

All collected tweets		*Collected unique tweets*	
Hashtag	Number of times used*	Hashtag	Number of times used*
brexit	153,170	brexit	54,876

eu	11,512	eu	2,536
yellowhammer	9,471	borisjohnson	2,492
borisjohnson	5,894	yellowhammer	1,942
nodeal	4,462	uk	1,754
nodealbrexit	4,087	brexitshambles	1,196
remain	3,902	remain	1,179
parliament	3,845	brexitchaos	1,121
prorogation	3,579	bbcqt	1,033
britishindependence	3,043	nodeal	956

**The number of times each hashtag was used is relative to the collected data set.*

Patterns of communication in the collected data set are already apparent from the general overview of sentiment analysis, term frequency and popular hashtags. Stories on the theme "#Brexit" were mostly negative, revolved around the British prime minister, the European union, the yellowhammer report and a no deal Brexit. The most frequent words show that users responded to the parliamentary events and the social divide between *leavers* and *remainers* continued in this discussion. The main character in the stories was Boris Johnson, the main scene was the British parliament and the plot is exemplified by the debate of leaving or remaining in the European Union, a no deal Brexit, prorogation of parliament and the yellowhammer report.

As one selected topic, this study will further analyze the discussion created with the hashtag "yellowhammer" (in section 4.3.1). Firstly, due to this hashtag's rise to popularity in such a short time. From the 7th to the 12th of September there were more than 2,000 unique tweets and close to 9,000 retweets with this hashtag (refer to appendix F). In addition to the relatively quick development of this topic, the yellowhammer report caused an outrage on the part of the British public as well as being a narrative turning point in the Twitter discussion. It is also expected that it will contain different social realities by different communities that took part in the discussion "#yellowhammer". The other topics discovered through the popular hashtags are also expected to contain different social realities and different dominant dramatization of stories (in terms of emotions as positive, negative and undecided). With an in-depth analysis of one dominant topic, this study will be able to describe of social realities. Although this is posing a limitation on the scope of the study, describing each social reality within each topic of discussion will not add value to the description will be. By bracketing one topic of discussion within the Brexit theme, this study shows typical social realities within the yellowhammer topic. For further analysis and categorization of patterns in the 3 classified groups and the dominant topic of "yellowhammer" discovered through hashtags, this study uses the whole collected data set inclusive of retweets. As noted above, based on the insights of Grčar *et al.* (2017) retweets are considered as sharing and agreeing with a story while representing the dynamic tendency in

communication on Twitter, which is crucial for this study. The patterns found in the unique tweets will not be completely excluded from this study and are presented in appendix G. In the following section dramatized patterns are shown and analyzed, which are categorized dependent on the unsupervised classification method, with sentiment analysis, as negative, positive and undecided.

4.2.1 Dramatized patterns

Dramatization in storytelling leads to the dynamic tendencies of sharing and agreeing with a story. In this study dramatization is characterized by the emotions shown in the stories as measured and classified by sentiment analysis. As seen from the general overview of dramatization, negative emotions dominated the storytelling on "#Brexit". In this section a side by side comparison of frequent terms, popular hashtags and words that characterize the theme are shown below. These dramatized patterns are categorized as classified by the sentiment analysis method.

Table 6. Frequent terms categorized according to classification with sentiment analysis

Negative			*Undecided*			*Positive*		
Word	Number of times used	Relative term frequency	Word	Number of times used	Relative term frequency	Word	Number of times used	Relative term frequency
people	18,422	0.0116	parliament	9,377	0.0165	referendum	14,376	0.0115
deal	13,263	0.0083	people	6,890	0.0121	people	14,150	0.0113
parliament	12,366	0.0078	price	6,160	0.0108	voted	13,208	0.0106
boris	12,035	0.0076	boris	5,893	0.0104	leave	11,083	0.0089
leave	11,731	0.0074	nodeal	4,691	0.0082	deal	10,574	0.0085
johnson	11,095	0.0070	yellowhammer	4,242	0.0075	remain	8,247	0.0066
remain	11,025	0.0069	johnson	3,696	0.0065	boris	8,234	0.0066
british	9,397	0.0059	election	3,602	0.0063	vote	7,707	0.0062
election	8,549	0.0054	wetherspoons	3,409	0.0060	party	7,143	0.0057
news	8,402	0.0053	labour	3,237	0.0057	johnson	6,989	0.0056
labour	7,922	0.0050	party	3,118	0.0055	government	6,835	0.0055
party	7,750	0.0049	vote	2,963	0.0052	parliament	6,475	0.0052
voted	7,488	0.0047	deal	2,808	0.0049	labour	5,987	0.0048
vote	7,301	0.0046	democracy	2,753	0.0048	media	5,859	0.0047
nodeal	7,243	0.0046	stopthecoup	2,668	0.0047	support	5,723	0.0045

**Note: The negative stories contained 28,929 unique words from a total of 1,588,437 words; The undecided stories contained 22,258 unique words from a total of 569,073 words; The positive stories contained unique 26,199 words from a total of 1,249,132 words.*

As seen in the table above, all three types of classified stories, as negative, undecided, and positive, had Boris Johnson as the main character and parliament as the main scene. However, yellowhammer is among the most frequent words in stories that were classified as undecided, despite the expectation of an overall negative dramatization on this topic. All stories also frequently contained the labour party, which is the British political party that abruptly switched sides from being *leavers* to becoming *remainers.* Also common for all three types of stories is the term people's vote, which refers to the people's choice to leave the European Union

as decided in the 2016 referendum vote. It is also apparent that all types of stories referred to the crucial political question of a deal or no deal Brexit.

To depict which words are characteristic for the dramatized stories, term frequency and inverse document frequency were combined to score each word from each type of dramatized story. This score was derived by categorizing the stories according to the classification and creating a corpus of three documents. Each document contained the terms from one of the three types.

Table 7. Characteristic words from each type of story

Negative			Undecided			Positive		
Word	Number of times used	Relative term frequency inverse document frequency	Word	Number of times used	Relative term frequency inverse document frequency	Word	Number of times used	Relative term frequency inverse document frequency
majorgeneral	867	0.0006	wins	100	0.0002	corbynschaos	1,470	0.0013
granddaughter	865	0.0006	tolls	1,306	0.0025	fouryearold	1,195	0.0011
referrals	761	0.0005	sophistry	924	0.0018	printed	3,141	0.0010
toplady	725	0.0005	datenight	2,141	0.0015	wins	1,719	0.0006
burner	1,947	0.0005	branson	520	0.0010	unfitness	625	0.0005

**The calculated weight of words according to term frequency inverse document frequency is relative to the collected data set.*

As shown in table 7, the words that characterize the negatively dramatized stories refer to the major general of the army, and the top lady or the Queen. The other 3 characteristic words in this type of stories are too general to interpret. While, the undecided stories are characterized by branson, referring to Richard Branson, a British philanthropist. In this category, tolls refer to the act of bells tolling to signal the end of Brexit or parliament. Although the rest of the 3 displayed words in the undecided stories are also too general to interpret, sophistry stands out. By definition it refers to a plausible yet misleading or fallacious argument and is considered to have a heavily negative connotation. Despite the negative intensity this particular word is characteristic for the undecided stories, which raises a flag for the precision of the classification method. Lastly, stories that were dramatized with positive emotions are characterized by the phrase corbynschaos, referring to the chaos that Jeremy Corbyn caused by supporting the *leave* side and afterwards siding with the *remainers.*

Table 8. Popular hashtags categorized according to classification with sentiment analysis

Negative		*Undecided*		*Positive*	
Hashtag	Number of times used	Hashtag	Number of times used	Hashtag	Number of times used
brexit	62,243	brexit	40,086	brexit	50,841
eu	5,832	parliament	2,728	yellowhammer	4,089
yellowhammer	3,408	nodealbrexit	2,707	eu	3,660
borisjohnson	2,105	nodeal	2,686	euarmy	2,090

remain	1,643	reesmogg	2,229	remain	1,911
uk	1,426	datenight	2,141	britishindependence	1,850
peoplesvote	1,344	borisjohnson	2,090	nprosjohnson	1,699
nodeal	1,302	eu	2,020	remainers	1,171
prorogation	1,275	yellowhammer	1,974	bbcqt	1,111
johnson	1,135	prorogation	1,613	leave	1,076

**The number of times each hashtag was used is relative to the collected data set. Negative stories contained 7,298 unique hashtags from a total of 144,467 hashtags; Undecided stories contained 6,689 unique hashtags from a total of 107,587 hashtags; Positive stories contained 6,982 unique hashtags from a total of 119,154 hashtags.*

Table 8 shows the top 10 popular hashtags in all three types of stories. As noted, hashtags with frequent use and sharing develop into topics, and from this data set it is apparent that the topic of "#yellowhammer" prevailed in the negatively and positively dramatized stories. However, despite the indication that yellowhammer as a frequent term was found in the undecided stories, from the categorization of hashtags yellowhammer is not as prevalent a topic in the undecided stories compared to the other types of stories. What is interesting from the use of hashtags is that yellowhammer tops the charts for both the negative and positive stories. This could be due to misclassification of false negatives or positives with the unsupervised method, or it implies that stories shared on the topic developed vastly different social realities in terms of dramatization.

4.3 Describing social realities

After discovering the patterns of communication, this study proceeds with the description of developed social realities in this section. For the general overview of social realities that were developed, LDA topic modeling was used. In particular, the model used was generative and used randomized Gibbs sampling. The method employed creates documents dependent on the size of the input data set, which was the sample of 53,912 unique tweets, and generated 36,000 documents (for the code used and the control list refer to appendix H). Thus 70% of the collected unique tweets are treated as as documents that contain terms. The output from this model are the 4 biggest topics and their top 10 keywords from the collected unique tweets, presented in table 9.

Table 9. Topic model top 10 keywords per topic

Brexit and economy	Brexit and parliament	News about Brexit	Brexit politicians
stocks	update	thread	britishindependence
kag	hours	hours	poverty
gbpusd	brilliant	welcome	chukaumunna
banking	perfect	shambles	opportunities
varadkar	brexitnow	quits	thoughts
amid	unbelievable	suspension	summary
drama	guardian	ringo	blamethetories
warns	jacobreesmogg	sums	chukacheck
reject	sterling	hilarious	findchuka

The largest discovered topic of discussion from the unique tweets was “Brexit and economy”. As seen from the output, the top keywords of shared dramatized stories on this topic refer to the economic state of Britain, the stock market, fluctuation of the British currency compared to American dollars, and banking. A story that was shared and agreed upon thousands of times on this topic is presented below.

→ “The man managing Norway’s $1 trillion wealth fund has vowed to invest in the UK no matter what type of #Brexit occurs as he believes the economy will power ahead. The fund owns ~1.5% of all global listed stocks so knows a thing or two about markets... https://t.co/HdDXE2BQq6”

→ *Retweet count: 1,965; Dramatization: Positive; Intensity: 3.*

This story about the British economy had 1,965 retweets at the end of data collection, out of which 1,095 retweets were captured with the collection methods. It was classified as positively dramatized with an intensity score of 3. The main character in this story is Norway’s wealth fund manager who believes in the market of United Kingdom and supports any kind of Brexit. He vowed that he would invest in their economy and as proof the user that shared this story linked a news article (https://t.co/HdDXE2BQq6). Although the overall story is indeed positive and suggests a bright future for the economic state of Britain, this user added an ironic expression at the end of the story aimed at those who would question it. The social reality developed through this story is one of economic security regardless of a deal or no deal Brexit that was shared between a community of 1,965 users.

Brexit and parliament is the second largest topic of discussion from the output of the topic model. From the top keywords on this topic it is apparent that the main character in the stories was Jacob Rees-Mogg, the elected president of parliament at the time of data collection. It is not a surprise that this politician in particular was the main character in the discussion, because he was caught on camera resting and lying down on a bench in parliament during the referendum resolution debate. Also notable is that users sharing stories on this topic used the phrase “brexitnow” referring to the end of the referendum resolution. A characteristic story on this topic was:

→ ”My evening at the Houses of Parliament. #ReesMogg #NoDeal #DateNight #Parliament #Brexit #NoDealBrexit https://t.co/WgxXrw66YZ”

→ *Retweet count: 51,855; Dramatization: Undecided; Intensity: 0.*

Although from the text of this story dramatization is undecided with 0 intensity, this story was a joke. It contained a satirical video montage (https://t.co/WgxXrw66YZ) of Jacob Rees-Mogg being caressed by a twitter user while he is lying down in the middle of a parliamentary session. It is up for debate whether satire or jokes can be classified as positive, negative, or undecided. This story was shared 51,855 times and was the most shared story in the whole collected data set; however, only 2,143 retweets were captured with the collection method. It developed a satirical social reality shared between a community of 51,855 users that mocked the British president of parliament.

As the third largest topic in the sample from the collected data set, News about Brexit, developed social realities that included stories about the shambles in parliament and the suspension (or prorogation) of parliament. As well as stating that the sessions were hilarious. A negatively dramatized social reality was developed by sharing the following story:

→ “Lord James has been threatened with the police, told to retract comments he made in the House of Lords on #EU Defence Union plans, is facing demands to resign and told to remain silent on the issue from now on. #EUArmy #Brexit Pls RT, let people know. https://t.co/d354wscZI2”

→ *Retweet count:4,353; Dramatization: Negative; Intensity: -4.*

This story talks about “threats” made to a British politician which were a consequence of his comments on the plans of the Defense Union when and if Britain exits the European Union. Lord James addressed the parliament and talked about the dangers of leaving and how Britain will transfer its forces to Brussels (UK column reporters, 2019). He was dubbed as the hero of Brexit by the *UK column*, and this story was shared 4,353 times out of which this study captured 2,471 retweets. It also contained a link to the news article by the *UK Column* (https://t.co/d354wscZI2) and developed a social reality shared between a community of 4,353 users. In this social reality the users agreed that James is the hero of Brexit and spread the word about the threats made to the politician.

Lastly, the smallest of the 4 topics from the collected unique tweets was Brexit and politicians. Besides the top keyword being "britishindependence", stories shared on this topic were referring to Chuka Umunna, a British politician that has been the target of ridicule by Twitter users. This is due to his constant switching of political parties in his career and remaining unheard on the Brexit issue. Although the topic model suggests that he was a part of the keywords that characterize this topic, this study only captured 98 tweets and retweets that contained his name. This is due to the occurrence of these keywords only in the documents that were allocated to the *politicians* topic and they were scored as important keywords. On this topic of discussion, the following story developed another satirical social reality:

→ "I have no idea which genius made it, but it was worth the creation of Monty Python for this alone. It's perfect. #BorisJohnson #JeremyCorbyn #Brexit #BrexitShambles #SundayThoughts #SundayMotivation #SundayMorning https://t.co/XlsFL64qr2"

→ *Retweets: 18,338; Dramatization: Positive; Intensity: 2*

There was a story about the proceedings in parliament that directly addressed the Sunday session on the 7th September 2019 during the referendum resolution. It refers to Boris Johnson and Jeremy Corbyn as the main characters and contains a video montage (https://t.co/XlsFL64qr2) of the parliamentary session spliced with an iconic movie scene. Only 76 retweets of this story were captured by data collection, however it is notable due to the total number of times it was shared. According to the unsupervised method of classification with sentiment analysis, this story was positively dramatized. This is another indication of the issue of unsupervised classification and how intricate uses of language and context (i.e. sarcasm, satire, metaphor, etc.) present a problem for the chosen method. It is arguable that due to the laughter caused by jokes such stories are perceived as positively dramatized, yet this story makes a mockery of the parliamentary proceedings. It developed a social reality shared between a community of 18, 338 users that mocked the referendum debates in parliament.

From the 4 topics it is apparent that 4 different social realities were developed, attracting thousands of users to share and agree with the messages and creating communities on the social media site. What all 4 developed social realities have in common, besides being developed through dramatized stories as a result of group communication, is legitimization. Two of the notable stories had news articles adding to the relevance and legitimization of the messages they

shared. While the other two had satirical videos that mocked the British parliament. This indicates, relative to this collected data set, that stories that were dramatized with satire attracted a bigger chunk of the Twitter users and developed social realities shared in large communities.

4.3.1 Social realities in "#Yellowhammer"

The Yellowhammer Report is an operational document made by the British government that concerns exiting the European union without making a deal about trade and import of supplies, among other things (Wood, 2019). It was "leaked" to the public on 11th September 2019. As shown in figure 4, this report presented a narrative turning point in the discussions on Twitter. Stories shared with this hashtag were mostly dramatized with negative emotions and caused the whole "#Brexit" theme of discussion to shift from somewhat positive to completely negative dramatization. This study collected a total of 10,975 tweets by 8,912 users (shown in table 10) on the topic "#yellowhammer" and the variation "#operationyellowhammer" which caused this dramatic shift in the dramatization of stories. The collected data set shows that this topic developed within 4 days, starting on the 9th of September 2019 (refer to appendix F). On this topic there were 8,848 retweets by 7,293 users and 2,128 unique tweets by 1,619 users from the collected data (presented in table 11). This specific hashtag developed into a topic of discussion quickly, and overall the stories shared on the topic were negative. From the total pool of 10,976 "#[operation]yellowhammer" tweets collected, 7,852 were classified as negative, while 2,095 were undecided and only about 10 % (or 1,029) were positive. Figure 5 presents the progressive decline from overall positivity in the dramatized stories to overwhelming negativity by the end of the narrative timeline At the end of data collection, the discussion on "#Yellowhammer" had an

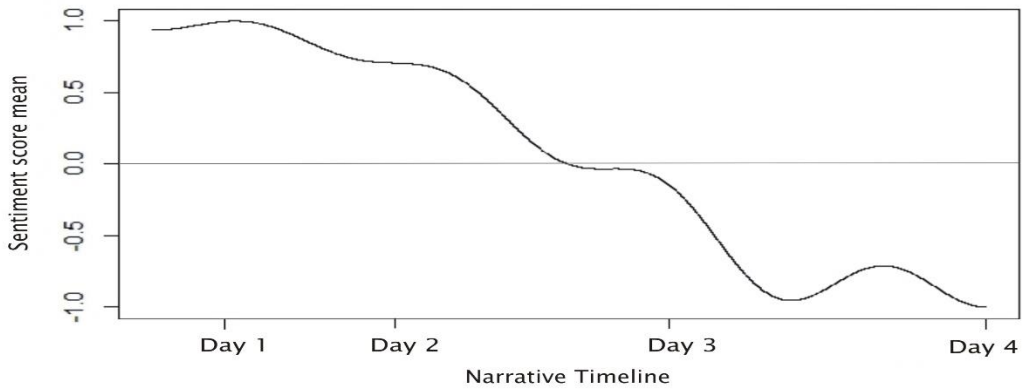
Table 10. Tweets collected with #Yellowhammer

	All collected	*Retweets*	*Unique tweets*
Number of tweets	10,976	8,848	2,128
Number of users	8,912	7,293	1,619

Table 11. Classification of #Yellowhammer tweets

	All collected	*Retweets*	*Unique tweets*
Positive	1,029	541	488
Undecided	2,095	1,517	578
Negative	7,852	6,790	1,062

Figure 5. Sentiment of #Yellowhammer stories over time



average sentiment score of -1 on a scale from -1 to 1. Within two days this topic was dominated by negatively dramatized stories. The report itself has negative implications for the British public, such as no trade deals and no European support whatsoever. The dramatized stories on Twitter as a response to events can be a direct reflection of the emotions that the British public had in the real world. Most stories that were shared with this hashtag in the collected set were negatively dramatized. This is also apparent by comparing the retweet counts of different dramatized types of stories, presented in table 12. Stories that were classified as negative had significantly more retweets than both positive and undecided

Table 12. Descriptive summary of retweet counts on #Yellowhammer

	Minimum	*1st Quartile*	*Median*	*Mean*	*3rd Quartile*	*Maximum*	*Standard deviation*
All collected	0	7	721	1,122	1,534	5,167	1,432
Positive	0	0	2	25	19	313	47
Undecided	0	1	19	436	1,087	1,434	548
Negative	0	69	1,146	1,448	1,609	5,167	1,548

stories. On average, from the collected data set, negatively dramatized stories that used "#yellowhammer" or "#operationyellowhammer" were shared and agreed with 1,448 times, with one story having 5,167 shares as the maximum. This confirms that the prevalent stories which developed social realities were a negative reflection on real-world events.

Table 13. Characteristic words in #Yellowhammer

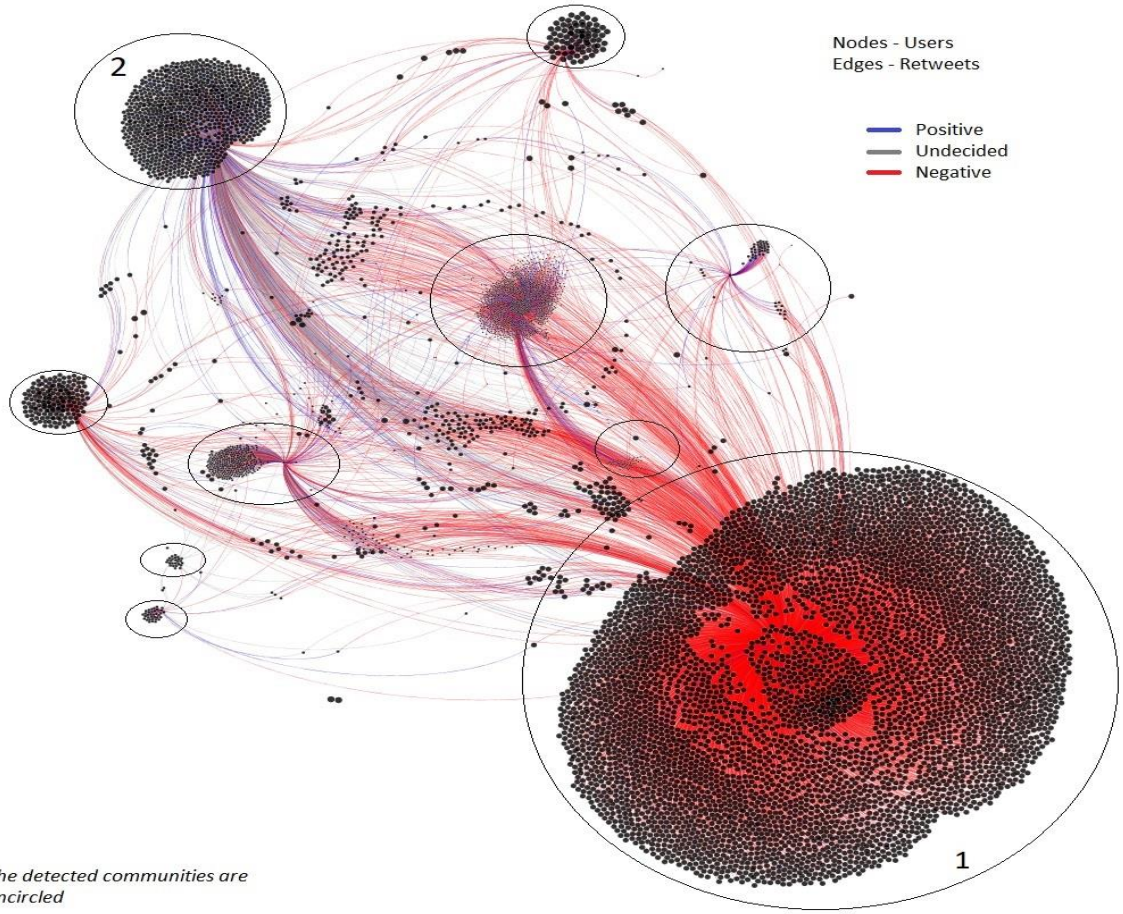
Word	Number of times used	Relative term frequency inverse document frequency
printed	3,137	0.0358
unredacted	3,145	0.0179
compile	1,071	0.0122
idleness	720	0.0082
complied	708	0.0081

On this topic, dramatized stories that users created and shared were characterized by the words presented in table 13. They included requests to see the compiled, unredacted and printed yellowhammer report, as well as the idleness and compliance of the British politicians in revealing the report. The user groups that shared stories with the selected hashtag used 5,793 unique words from a total of 121,518 words. Considering that in this collected data set there is a dramatized story that was shared over 5, 000 times, most of the frequent words and popular hashtags in the vocabulary from the "#Yellowhammer" topic come from that one story. The frequent words and popular hashtags on this topic are presented in appendix I.

For a better description of developed social realities with this hashtag, the selected data set was graphed in a forced network of retweets. This was done by manipulating the collected data set, extracting stories that contained "#yellowhammer" and the variation "#operationyellowhammer", and visualizing the forced network in Gephi (Bastian *et al.*, 2009).

Users are represented as nodes and the connections between them as edges. If users shared a story they are connected by the edges (as vertices) to the user that posted the story.

Figure 6. #Yellowhammer



With visualizing a forced retweet network this study also discovered 10 different communities detected through algorithmic methods, that developed respectively different social realities, presented in figure 6. In the visualization of a forced retweet network two communities stand out, which are labeled as 1 and 2. Community 1 is the largest community as computationally decided. This community developed a negatively dramatized social reality. The second largest community, labeled with 2, developed a mixed social reality which leaned towards undecidedness. The communities were discovered by using the Louvian method as a standard community detection algorithm (Grčar *et al.*, 2017) which is available in Gephi (refer to appendix J). There are 10 modularity classes representing each community, with an overall 0.5 modularity score in the network. This indicates that the network had strong and dense connections within the communities and somewhat strong connection between communities. The

graphed network had 3 weakly connected users and 8,930 strongly connected users. The communities' modularity and size were maximized by using the standard algorithm and by grouping the communities dependent on which story they retweeted and how the story was dramatized. The chosen method for visualization forces the nodes to form communities around their mutual connections (the edges as retweets) and the detected communities are solely computationally decided. The yellowhammer discussion was visualized with this method because it shows the different communities and developed social realities in proportional size to the number of shares of stories. Overall, the communities presented in figure 6 mostly shared negatively dramatized stories, however we can also notice mixed emotions from community 2 which shared all three types of dramatized stories.

Community 1 had 6,286 connected users which developed their negatively dramatized social reality through the following stories:

→ *1)* "I want to read the unredacted #Yellowhammer report and I believe it needs to be printed across all media with the clear guidance that it is NOT a worst case scenario. The government think we're too stupid to understand it. Please RT if you're in agreement, thanks. #Brexit"

→ *Retweets: 5,167; Dramatization: Negative; Intensity: -1.*

→ *2)* "The doctor who helped compile #operationyellowhammer and took on @Jacob_Rees_Mogg tells us why he thinks Brits will die if there's a no deal #Brexit https://t.co/ZTi67LN964"

→ *Retweets: 1,174; Dramatization: Negative; Intensity: -4.*

Story *1)* was the most shared story that used this hashtag in the collected dataset. From the total of 5,167 retweets, this study captured 3,136 in the data collection period. It was originally misclassified with the unsupervised methods as positive with an intensity of 1. Due to the content of this story and use of language, the score was reversed. This story vilifies the government, as one of the two main characters, by stating that they believe the people are "…too stupid to understand…". This short statement in the second sentence of the story shows negative dramatization. The user that created the story explicitly implored other users to share it if they agree with the opinion that the yellowhammer report needs to be fully revealed to the public.

A doctor who confronted Jacob Rees-Mogg about the no deal Brexit is the main character in the second most shared story (*2*) in this community. It is accompanied by a video which presents the doctors opinion backed up by "facts" (https://t.co/ZTi67LN964). This story was shared 1,174 times, out of which 1,073 retweets were captured in the data collection. Users agreed with the opinion that people will die if there is no deal Brexit. The user who created this story directly referred to Jacob Rees-Mogg by mentioning him in the text ("@Jacob_Rees_Mogg"). This community developed a negatively dramatized social reality that promulgated the belief that the government is villainous and undermines its own population, as well as sharing the opinion that no deal Brexit will result in deaths.

The second largest connected component in the graph is community 2 which contained 1,469 users. The two most prevalent stories that developed the social reality of this community are undecided:

→ *1)* "Dominic Grieve skewers Boris Johnson for his "manly idleness" <U+0001F602> #Brexit #Prorogation #Yellowhammer https://t.co/j99vEcqJCK"

→ *Retweets: 1,434; Dramatization: Undecided; Intensity: 0.*

→2) "It sounds as if the government won't be releasing the #Yellowhammer #NoDeal docs. Andrea Leadsom - Putting out #Yellowhammer docs would just "concern people" #r4today #bbcbreakfast #brexit https://t.co/NfPsziLyaQ"

→ *Retweets: 1,053; Dramatization: Undecided; Intensity: 0*

Both undecidedly dramatized stories were created by the same user, and where shared by 1,434 users and 1,053 users respectively. With the means of collection, this study captures a total of 786 retweets combined from both stories. The first story is a portrayal of the critique by politician Dominic Grieve towards Boris Johnson for his idleness. It is accompanied by a clipped video (https://t.co/j99vEcqJCK) by the BBC. While the second story is also essentially a comment on a BBC news report video (https://t.co/NfPsziLyaQ), stating that people will just be concerned by the revelation of the worst case scenario report, yellowhammer.

Although both biggest communities in the captured discussion about yellowhammer on twitter shared stories about the implications and ramification of a no deal Brexit, they developed two very different social realities. Community 2 mostly shared stories that were a statement of

fact backed up by news reports, developing a social reality that was a response to the parliamentary proceedings on the 9th of September. While community 2, shared stories that vilified the British government and their prime minister, developing a social reality that was a reflection to the outrage of the British public at the decision to not reveal this report. Notably, the most shared story in this discussion had no extra content (such as a video or a link to a news report), but instead explicitly asked users to share if they agreed with the statement that the government undermines the British public.

The stories shared on the theme Brexit on Twitter were mostly dramatized with negative emotions. In the narrative timeline of this theme, the topic yellowhammer played a big role as a turning point that caused negatively dramatized stories to dominate the discussion. The discovered social realities were developed by stories that had British politicians as the main characters, had parliament as the scene and the political debate of leaving from the European Union with or without a deal as the plot. The top shared stories by Twitter users vilified the politicians and government, or mocked them, most often referring to the prime minister Boris Johnson and Jacob Rees-Mogg. One frequently used phrase and hashtag to describe the parliamentary proceedings in dramatized stories was Brexit shambles, addressing the chaos caused by the prorogation. While stories that discussed the operational report Yellowhammer mostly expressed negative emotions towards British politicians and government, with one social reality developed around the statement that the British government is undermining its population. From the analysis and results of describing the developed social realities on the theme, this study finds that the most popular stories that became topics of discussion disagreed with the British government's actions in the period between the 5th and 12th September 2019.

5. Discussion

This study used symbolic convergence theory and concepts from narrative paradigm to propose a model for description of social realities developed as a result of online group communication on social media sites. Although this theory has been used in the past for analysis of group communication, research was mostly conducted with a focus on in-person communication or in an organizational setting. Furthermore, use of symbolic convergence theory for social media sites has had visual cues as a point of focus. This study shows how this theory with concepts from narrative paradigm can be used to analyze text and emotions of stories shared

on social media sites. With alignment between the theoretical concepts and online group communication on social media sites, the place of symbolic convergence theory for use in nowadays communication is solidified. This contributes to further use of symbolic convergence for such analysis, and even broader use for analyses of social movement organizations, activist networks, political, social and economic issues by providing a theoretical framework for describing social realities as the topics user groups developed, on a theme of discussion, through sharing stories which they dramatized with emotions. In addition to sharing dramatized stories, users form communities around dramatized stories that they agree with.

As an answer to the research question, this study posits that social realities developed as a result of online group communication on social media are the shared emotions and stories within communities as a response to real world issues, which can create topics of discussion not only within but also between different communities that influence users viewpoint of current events. Description of social realities developed on social media sites also implies the use of methods and procedures from data science, to categorize and analyze patterns of communication, the emotions of the stories and to visualize the communities that developed these social realities. Based on the results from this analysis, one finding stands out, besides negatively dramatized stories being shared the most. The discussion that revolved around the yellowhammer report shows how a couple of negatively dramatized stories can influence the whole theme of discussion ("#Brexit"). Essentially, these stories represent a narrative turning point that swayed the whole theme towards negatively dramatized stories over a period of two days. It is also interesting how quickly (within 4 days) the yellowhammer discussion developed social realities and attracted multiple communities to share and agree with stories on this topic. This could be due to the outrage a no deal Brexit provoked in the British public and shows how quick users were in "voicing" their opinion and sharing their stories on Twitter. A smaller finding are the satirical visual cues that accompanied some of the most popular dramatized stories, mocking the government and politicians mercilessly. However, it is already known that visual cues play a role in the development of social realities from previous studies using symbolic convergence theory.

Furthermore, even though the model proposed by this study (in figure 1) accounts for multiple social media sites, includes visual cues and a few types of dynamic tendencies, the operational model (figure 2) for the chosen case of "#Brexit" on Twitter was limited. This can be

improved by including qualitative and quantitative analysis of visual cues (such as mems, gifs, videos and emojis or emoticons) as complimentary to sentiment analysis for a more refined depiction of dramatization. In particular, emojis or emoticons are in heavy use on todays social media sites and directly represent emotions and feelings. Future studies that use the model should include emoji analysis with sentiment analysis for a more accurate dramatization score in terms of emotions. On top of the mentioned visual cues, an operational model for describing social realities developed as a result of online group communication can be refined by including all of the dynamic tendencies that a chosen social media site offers its users. Dynamic tendencies such as mentioning users, responding (as replies) and reacting (as dis/likes, up/downvotes, etc.) to dramatized stories.

5.1 Limitations

In terms of data collection methods this study was limited by using a free, open API for Twitter. The disadvantages of using the open API are access to only a portion of publicly available tweets, as well as the ability to acquire all 90 data points that Twitter provides for each posted tweet. In addition to the API limitation, this study only collected tweets that used "#Brexit". Studies by Grčar *et al.* (2017), Hall *et al.* (2018), and Bassilakis *et al.* (2018) suggest that for data collection over a longer period of time, aggregating specific key words, hashtags and topics proves to be a more successful method for capturing the whole discussion. However, this collection method was outside the scope of this study, considering the short collection period and the limited open API. Another limitation in this study, due to the scope, is the collection of data only in English. The Brexit referendum had a direct impact on the European Union as a whole, and indirectly impacted the rest of the world in terms of politics and economy.

As mentioned, the unsupervised classification method based on sentiment analysis by using a bag of words approach is limited and not as accurate as supervised classification. The main difference being that supervised classification implies human annotation, inner-coder agreement and validation, which were outside of the scope for this study. Even though the unsupervised classifying method used is nowhere near a gold standard for classification of social media data, it still provides a credible general overview of the sentiment from the data set. From the aspect of dramatization, this general overview of sentiment and classification in three groups proved to be very valuable. Yet, the bag of words approach to sentiment analysis and text

analysis has issues when dealing with wordplay, jokes, sarcasm, irony, etc. Arguably, these forms of communication can be easily detected with human supervision, but for unsupervised classification they require advanced methods that were outside of this study's scope. For instance, such methods are recent developments with word-character neural networks and pre-trained models for sarcasm detection which are very hardware expensive.

Lastly, in the descriptive model in chapter 2, shared meaning was included, contrasted by the choice to exclude shared meaning from the operational model and analysis. It is debatable whether the selected tweets as dramatized stories that created topics of discussion and developed social realities can be inferred as shared meanings of a community. However, the difficulty of discovering shared meaning arises with the operational definition by Saffer (2016, 2018), who states that discovering shared meanings in a network of relationships requires supervised annotation, evaluating statements, non-parametric tests for homogeneity, cross-validation and *k-clique* analysis. K-clique analysis is social network analysis of groups and subgroups and their mutual relationships determined by how strongly they agree or disagree with given statements. This, again, was outside of this study's scope.

5.2 Implications and future direction

For future research, this study suggests that the presented methods of analysis and procedures can be used for discourse analysis, activism, social movement organizations, [political] debates on social media, and other popular topics., e.g. the upcoming United States presidential elections in 2020. By collecting data over a longer period and gathering the timelines of online stories by candidates, future research can describe and understand the created social realities, analyze the shared dramatized stories by discovering patterns of communication and graphing a network of retweets. Such an approach can enable future studies to describe and understand online group consciousness, in which case a method for analyzing collective action will be necessary. This study encourages further development of the descriptive model and believes that such a model can be implemented even outside of the political context. For instance, as Duffy's and others studies suggest, famous sports personae (Page *et al.*, 2016) and online hate groups (Duffy, 2003) can also be analyzed with this framework to uncover the social realities they develop. The benefit from such analyses can be an understanding why social media users share the stories that they do, and when they do it. For example, symbolic cues such as

memes and emojis can be considered as factors of dramatization. By applying the presented concepts and the proposed descriptive model to analyzing the stories shared by political candidates in the 2020 US election, research will be able to discover exactly which stories and what kind of dramatization developed a social reality that attracted a large social media community. This can be valuable to analyze the public's view of a candidate or to rank the candidates according to their dramatization scores (combining emotions and caused dynamic tendencies). As for dynamic tendencies in online group communication, future research can also consider favoring or liking, mentioning users and replying to stories in addition to sharing or retweeting. Lastly, analyzing a network of relationships with the approach this study used is not limited to a forced retweet network. Taking into account different dynamic tendencies, as mentioned above, can add value to the description of developed social realities. Networks of replies or mentions, for instance, can show which dramatized stories attracted the most responses by social media users or which users frequently shared dramatized stories that created communities.

6. Conclusion

This study proposes a descriptive model for social realities developed on social media sites by using the theoretical framework from symbolic convergence theory and narrative paradigm, implying how research can use these rhetorical and interpretative concepts to analyze todays online communication. Analysis with the proposed model allows online group communication to be viewed as a process of storytelling, in which users share their emotions through their stories and create communities which agree with a stance or viewpoint on real world events. Application of this model on the theme of Brexit on Twitter required operationalization by aligning theoretical concepts with group communication on social media sites and using methodology for social network and media sites analysis. The operational model used in the case did not account for visual cues such as emojis. However, this is valuable for future research, to refine the model with inclusion of emoji analysis adding to the accuracy of emotional overview of stories. Further, the approach and procedure for analysis allow discovery of trends in storytelling, or in other words patterns of communication while sharing dramatized stories and creating communities. This in turn benefits the understanding of social media users' stance and viewpoint of real-world events and social, political, or economic issues. The apparent patterns in this study show that negatively dramatized stories cause a drastic change in the

narrative evolution of a theme of discussion. Moreover, negative dramatization was commonplace by Twitter users sharing stories about Brexit politics compared to the classified positive and undecided dramatized stories. The typical discovered social realities developed by the biggest communities mocked the British government, the prime minister, and president of parliament as a response to the shambles caused by the prorogation and the referendum debate in parliament. Also, the largest community that shared stories about the no deal Brexit report, yellowhammer, developed a negatively dramatized social reality that portrayed the government and parliament as villains, and emphasized their inability to trust their public. The social realities of "#Brexit" on Twitter had politicians as the main characters, parliament as the scene and prorogation, deal, or no deal as the plot at the center of the stories that developed into topics. In particular, the quickly developed social realities on the topic yellowhammer show how negative stories can easily sway the emotional load of the discussion.

While this research was conducted, the British prime minister and president of parliament were re-elected, and they managed to exit the European Union with a deal. However, the analysis and results of this study show the negative emotions and mocking politicians or government dominated stories shared on the Brexit topic. This contrast between what the results show and what happened in the real world can be due to the divided society in Britain, with a large portion of it's voter pool on the margins of society or not using Twitter at all. On the other hand, the discussion about yellowhammer on Twitter shows how quickly stories are born, developed into topics and social realities which dramatically change the narrative landscape. However, taking into account the contrast between what happened in reality and the social realities of Brexit on Twitter, topics, social realities and moreover stories about politics can dissipate as quickly as they are developed on social media sites. From the theoretical standpoint, the social media stories did influence how users perceived British politics and their government, although after a period of 2 months (November 2019) people in the real world supported the same political figures that they vilified and mocked on Twitter. Ironically, the politicians that were a target for mockery and slander due to the chaos created in the period between 5th and 12th September 2019, were the same politicians who managed to save Britain's exit by striking a deal with the European Union. All stories are not created equal but can become equally powerful when they are dramatized and shared by communities through online group communication on social media sites.

References

Bales, R. F. (1970). *Personality and interpersonal behavior*. Holt, Rinehart and Winston. New York.

Bassilakis, A., Diermeier, M., & Goecke, H. (2018). Converging Media versus Diverging Politics-the Brexit Twitter on Debate. In *CESifo Forum* (Vol. 19, No. 4, pp. 29-35). München: ifo Institut–Leibniz-Institut für Wirtschaftsforschung an der Universität München.

Bastian M., Heymann S., Jacomy M. (2009). *Gephi: an open source software for exploring and manipulating networks.* International AAAI Conference on Weblogs and Social Media.

BBC News reporters (2019). Brexit: your simple guide to the UK leaving the EU. BBC News; retrieved from: https://www.bbc.com/news/uk-46318565

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, *2008*(10), P10008.

Bormann, E. G. (1972). Fantasy and rhetorical vision: The rhetorical criticism of social reality. *Quarterly journal of speech*, *58*(4), 396-407.

Bormann, E. G. (1985). Symbolic convergence theory: A communication formulation. *Journal of communication*.

Bormann, E. G., Craan, J. F., & Shields, D. C. (1994). In defense of symbolic convergence theory: A look at the theory and its criticisms after two decades. *Communication Theory*, *4*(4), 259-294.

Bormann, E. G., Cragan, J. F., & Shields, D. C. (2001). Three decades of developing, grounding, and using symbolic convergence theory (SCT). *Annals of the International Communication Association*, *25*(1), 271-313.

Bormann, E. G., Cragan, J. F., & Shields, D. C. (2003). Defending symbolic convergence theory from an imaginary Gunn. *Quarterly Journal of Speech*, *89*(4), 366-372.

Bormann, E. G., Koester, J., & Bennett, J. (1978). Political cartoons and salient rhetorical fantasies: An empirical analysis of the'76 presidential campaign. *Communications Monographs*, *45*(4), 317-329.

Buntine, W. L., & Mishra, S. (2014, August). Experiments with non-parametric topic models. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 881-890).

Burke, K. (1968). Dramatism. *International encyclopedia of the social sciences*, *7*(445-452).

Burke, K. (1985). Dramatism and logology. *Communication Quarterly*, *33*(2), 89-93.

Castor, T. R. (2005). Constructing social reality in organizational decision making: Account vocabularies in a diversity discussion. *Management Communication Quarterly*, *18*(4), 479-508.

Collins, S. D., DeWitt, J. R., & LeFebvre, R. K. (2019). Hashtag diplomacy: twitter as a tool for engaging in public diplomacy and promoting US foreign policy. *Place Branding and Public Diplomacy*, *15*(2), 78–96. https://doi.org/10.1057/s41254-019-00119-5

Dickerson, A. (2008). Symbolic convergence theory. In L. L. Kaid & C. Holtz-Bacha (Eds.), *Encyclopedia of political communication* (Vol. 1, pp. 768-769). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412953993.n654

Duffy, M. E. (2003). Web of hate: A fantasy theme analysis of the rhetorical vision of hate groups online. *Journal of Communication Inquiry*, *27*(3), 291-312.

Fisher, W. R. (1984). Narration as a human communication paradigm: The case of public moral argument. *Communications Monographs*, *51*(1), 1-22.

Goodwin, M. J., & Heath, O. (2016). The 2016 referendum, Brexit and the left behind: An aggregate‐level analysis of the result. *The Political Quarterly*, *87*(3), 323-332.

Graham, T., & Ackland, R. (2015). Topic modeling of tweets in R: A tutorial and methodology

Grčar, M., Cherepnalkoski, D., Mozetič, I., & Novak, P. K. (2017). Stance and influence of Twitter users regarding the Brexit referendum. *Computational social networks*, *4*(1), 6.

Griffin, E. A. (2011). *A first look at communication theory*. McGraw-Hill Education; *(8)*

Hall, W., Tinati, R., & Jennings, W. (2018). From brexit to trump: social media's role in democracy. *Computer*, *51*(1), 18-27.

Hobolt, S. B. (2016). The Brexit vote: a divided nation, a divided continent. *Journal of European Public Policy*, *23*(9), 1259-1277.

Hong, L., & Davison, B. D. (2010, July). Empirical study of topic modeling in twitter. In *Proceedings of the first workshop on social media analytics* (pp. 80-88).

Hürlimann, M., Davis, B., Cortis, K., Freitas, A., Handschuh, S., & Fernández, S. (2016, September). A twitter sentiment gold standard for the brexit referendum. In *Proceedings of the 12th international conference on semantic systems* (pp. 193-196).

Jockers, M. (2017). Introduction to the Syuzhet package. https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html,

Kincaid, D. L. (1979). The convergence model of communication.

Kincaid, D. L. (2002). Drama, emotion, and cultural convergence. *Communication Theory*, *12*(2), 136-152.

Llewellyn, C., & Cram, L. (2016, March). Brexit? analyzing opinion on the UK-EU referendum within Twitter. In *Tenth International AAAI Conference on Web and Social Media*.

Newman, D., Bonilla, E. V., & Buntine, W. (2011). Improving topic coherence with regularized topic models. In *Advances in neural information processing systems* (pp. 496-504).

Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*.

Niklander S. (2017) Emotional Computing and Discourse Analysis: A Case Study About Brexit in Twitter. In: Stephanidis C. (eds) HCI International 2017 – Posters' Extended Abstracts. HCI 2017. *Communications in Computer and Information Science*, vol 713. Springer, Cham

Olufowote, J. O. (2006). Rousing and redirecting a sleeping giant: Symbolic convergence theory and complexities in the communicative constitution of collective action. *Management Communication Quarterly*, *19*(3), 451-492.

Olumide Olufowote, J. (2017). Symbolic Convergence Theory. *The International Encyclopedia of Organizational Communication*, 1-8.

Page, J. T. & Duffy M. (2018) What Does Credibility Look like? Tweets and Walls in U.S. Presidential Candidates' Visual Storytelling. *Journal of Political Marketing*, 17:1, 3-31, DOI: 10.1080/15377857.2016.1171819

Page, J. T., Duffy, M., Frisby, C., & Perreault, G. (2016). Richard Sherman speaks and almost breaks the internet: Race, media, and football1. *Howard Journal of Communications*, *27*(3), 270-289.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/

Saffer, A. J. (2016). A message-focused measurement of the communication dimension of social capital: Revealing shared meaning in a network of relationships. *Journal of Public Relations Research*, *28*(3-4), 170-192.

Saffer, A. J. (2018). The Outcomes of Engagement in Activism Networks: A Co‐creational Approach. *The handbook of communication engagement*, 285-299.

Silge, J., & Robinson, D. (2019). Text mining with R: a tidy approach. O'reily,

Southern, R., & Lee, B. J. (2019). Politics as usual? Assessing the extent and content of candidate-level online campaigning at the 2015 UK general election. *Journal of Elections, Public Opinion and Parties*, *29*(2), 179–198. https://doi.org/10.1080/17457289.2018.1515209

Stache, L. (2017). Fisher narrative paradigm. In M. Allen (Ed.), *The sage encyclopedia of communication research methods* (Vol. 2, pp. 576-578). Thousand Oaks, CA: SAGE Publications, Inc doi: 10.4135/9781483381411.n205

Swanson, D. L. (1977). A reflective view of the epistemology of critical inquiry.

UK Column reporters, (2019). 'Hero of Brexit' Lord James of Blackheath threatened over EU Defence Union. UK column; retrieved from: https://www.ukcolumn.org/article/hero-brexit-lord-james-blackheath-threatened- over-eu-defence-union

Wheeler, B., Seddon, P., & Morris R. (2019). Brexit: All you need to know about the UK leaving the EU. BBC News. retrieved from : https://www.bbc.com/news/uk-politics-32810887

Wickham, H. (2017). Tidyverse, https://CRAN.R-project.org/package=tidyverse

Wood, I. (2014, December). Recent advances and applications of probabilistic topic models. In *AIP Conference Proceedings* (Vol. 1636, No. 1, pp. 124-130). American Institute of Physics.

Wood, J. (2019). Operation Yellowhammer No Deal Brexit planning documents say there will NOT be food shortages - but prices will go up, two-day queues at the channel will last three months and financial services will be hit. Mailonline, Daily mail UK, retrieved from : https://www.dailymail.co.uk/news/article- 7453695/Operation-Yellowhammer-No-Deal-documents-released-say.html

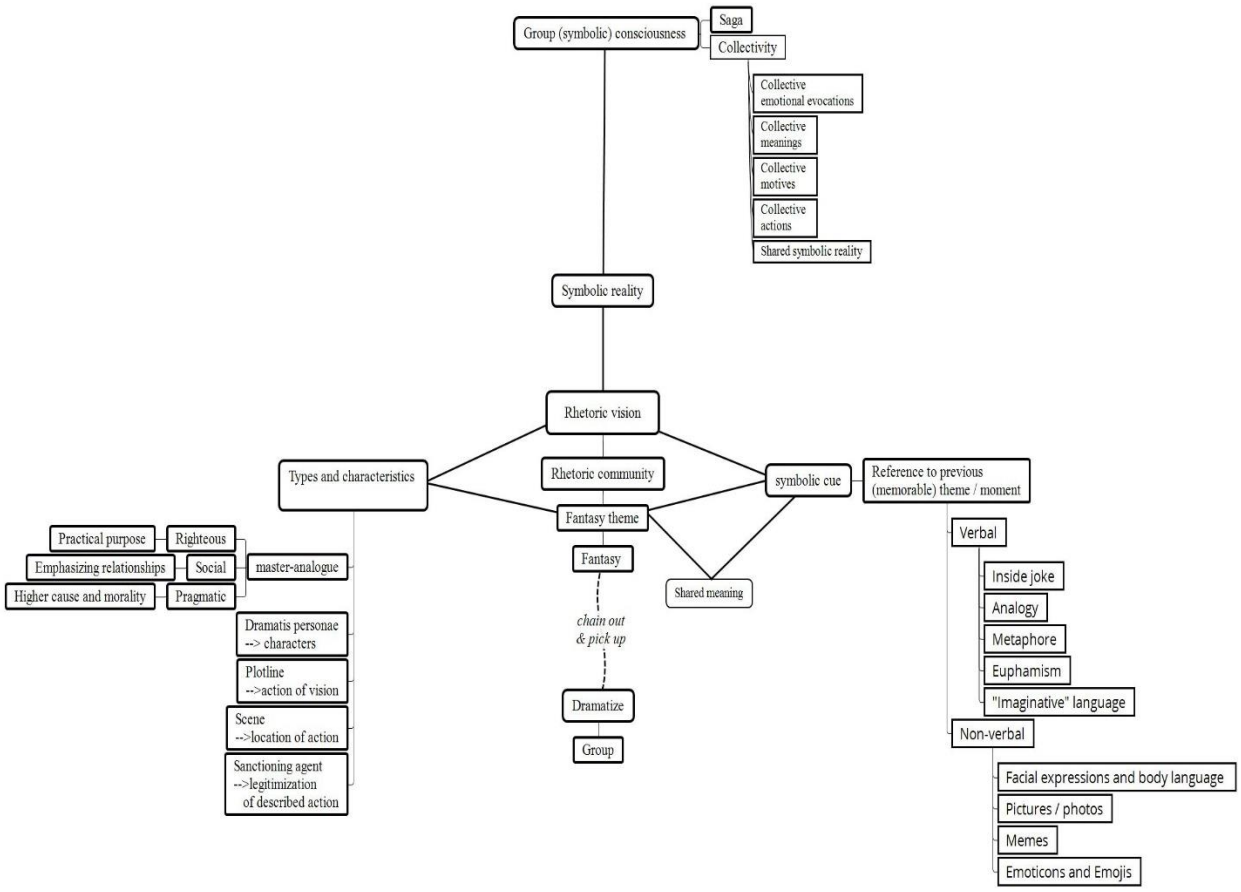
Xiong, Y., Cho, M., & Boatwright, B. (2019). Hashtag activism and message frames among social movement organizations: Semantic network analysis and thematic analysis of Twitter during the# MeToo movement. *Public Relations Review*, *45*(1), 10-23.

Woodcock, A. (2019) "Boris Johnson's suspension of parliament ruled unlawful by court", *Independent UK*, retrieved from: https://www.independent.co.uk/news/uk/politics/parliament-prorogue-suspend-unlawful-boris-johnson-latest-scottish-supreme-court-brexit-no-deal-a9100396.html

Appendices

Appendix A

Figure 1-1. Map of Symbolic convergence theory



In figure 1-1 all of the concepts from symbolic convergence theory are presented with added examples under *symbolic cues*. This study used the concepts that overlap with Twitter communication and fit in the scope of the study. For detailed characteristics and definitions of all concepts refer to Bormann *et al.* (2001) (see also: Dickerson, 2008, pp.768-769; Griffin, 2011; Olufowote, 2006, 2017)

References

Bormann, E. G., Cragan, J. F., & Shields, D. C. (2001). Three decades of developing, grounding, and using symbolic convergence theory (SCT). *Annals of the International Communication Association*, *25*(1), 271-313.

Dickerson, A. (2008). Symbolic convergence theory. In L. L. Kaid & C. Holtz-Bacha (Eds.), *Encyclopedia of political communication* (Vol. 1, pp. 768-769). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412953993.n654

Griffin, E. A. (2011). *A first look at communication theory*. McGraw-Hill Education; *(8)*

Olufowote, J. O. (2006). Rousing and redirecting a sleeping giant: Symbolic convergence theory and complexities in the communicative constitution of collective action. *Management Communication Quarterly*, *19*(3), 451-492.

Olumide Olufowote, J. (2017). Symbolic Convergence Theory. *The International Encyclopedia of Organizational Communication*, 1-8.

Appendix B

Figure 2-1. *Tidytext* procedures for text mining and analysis

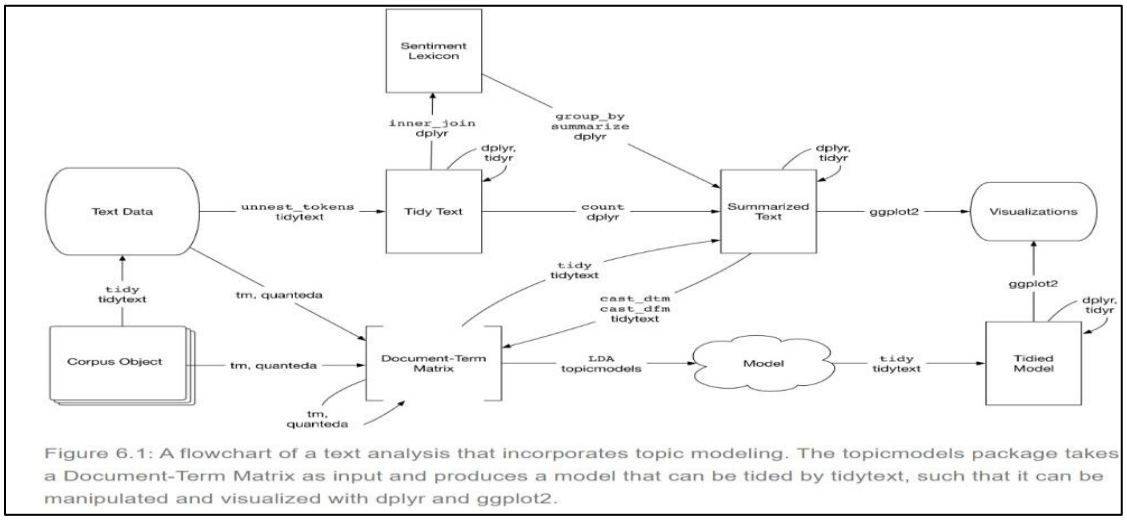


Figure 6.1: A flowchart of a text analysis that incorporates topic modeling. The topicmodels package takes a Document-Term Matrix as input and produces a model that can be tided by tidytext, such that it can be manipulated and visualized with dplyr and ggplot2.

Figure 2-1 has been taken from Silge and Robinson (2019) in accordance with Creative Commons license.

This figure represents the full procedure for text analysis, including manipulation and cleaning of text data, sentiment analysis and topic modeling. For a full and in-depth overview of the whole procedure "Text mining with R" is freely available online (https://www.tidytextmining.com/).

Reference

Silge, J., & Robinson, D. (2019). Text mining with R: a tidy approach. *O'reily*. https://www.tidytextmining.com/

License

Appendix C

Figure 2-2. Data analysis procedure for political events in social media sites

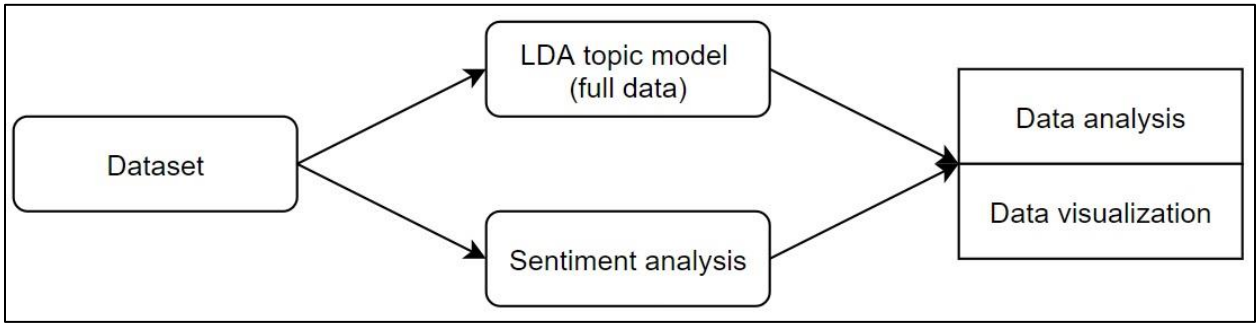


Figure 2-2 represents the standardized procedure for analysis of social media sites data in political events. Adapted from Hall et al. (2018)

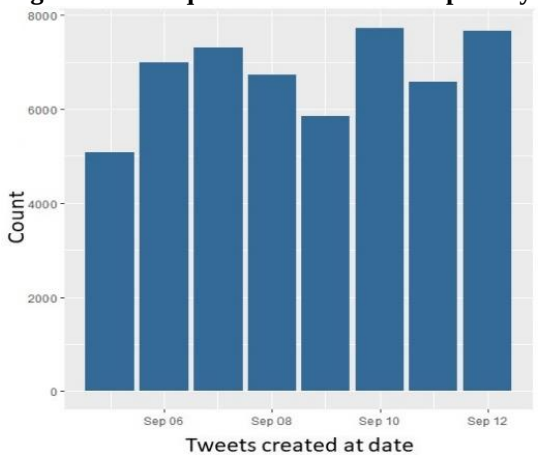
The figure above shows an adaptation of the standardized data pipeline and analysis for social media sites in political events proposed by Hall *et al.* (2018). The procedure includes LDA topic modeling, sentiment analysis, further data analysis (such as discovering patterns) and data visualization. Visualizations are done by creating a network of users and tweets from the collected dataset. In the chosen case of this study, networks of users will be visualized based on the sentiment of retweets.

Reference

Hall, W., Tinati, R., & Jennings, W. (2018). From brexit to trump: social media's role in democracy. *Computer*, *51*(1), 18-27.

Appendix D

Figure 3-1. Unique tweets distribution per day



Appendix E

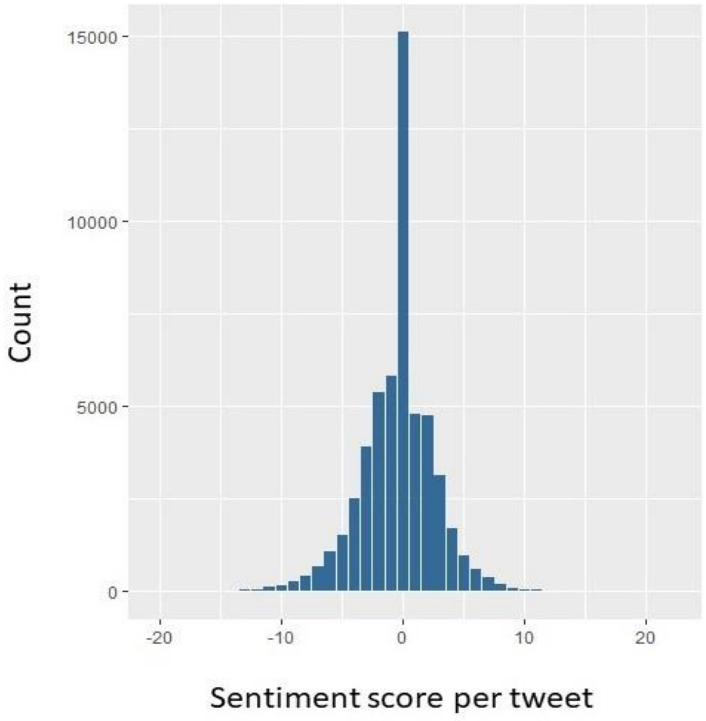
Table 2-1. Descriptive summary of retweet counts from the collected data

	Total data set of tweets	Adjusted set with no zero retweets
Minimum	0	1
1st Quartile	4	26
Median	119	221
Mean	904	1065
3rd Quartile	653	837
Maximum	51,855	51,855
Standard deviation	4,493	4,858
Relative total number of retweets	252,480,375	252,480,375

Table 2-2. Descriptive summary of sentiment scores from the collected data

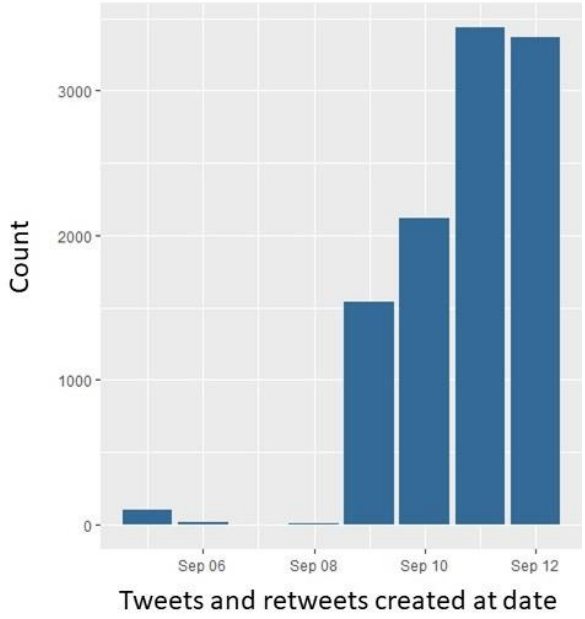
	Sentiment scores of all collected tweets	Sentiment scores of unique tweets
Minimum	-20	-20
1st Quartile	-2	-2
Median	0	0
Mean	-0.5	-0.42
3rd Quartile	1	1
Maximum	22	22
Standard deviation	3.2	3

Figure 4-1. Sentiment intensity distribution



Appendix F

Figure 5-1. Distribution of tweets on "#[operation]yellowhammer]



Appendix G

Vocabulary from stories classified as Negative excluding retweets:

Word	Number of times used	Relative term frequency
deal	2,590	0.0104
people	2,470	0.0099
leave	2,436	0.0098
boris	1,976	0.0079
parliament	1,868	0.0075
vote	1,533	0.0061
remain	1,388	0.0056
johnson	1,283	0.0051
nodeal	1,210	0.0048
voted	1,186	0.0048
yellowhammer	1,151	0.0046
government	1,141	0.0046
stop	1,120	0.0045
borisjohnson	1,119	0.0045
election	1,119	0.0045

Word	Number of times used	Relative term frequency inverse document frequency
wtf	68	0.0003
jailed	45	0.0002
violent	45	0.0002
symptom	99	0.0002
esa	97	0.0002

Hashtag	Number of times used
brexit	22,441
borisjohnson	1,175
eu	1,065
yellowhammer	967
uk	732
remain	625
leave	522
brexitshambles	465
bbcqt	463
parliament	461

Vocabulary from stories classified as Positive excluding retweets:

Word	Number of times used	Relative term frequency
people	1,804	0.0105
deal	1,374	0.0080
boris	1,205	0.0070
vote	1,104	0.0064
leave	1,039	0.0060
parliament	924	0.0054
remain	870	0.0051
election	782	0.0046
labour	723	0.0042
referendum	705	0.0041
party	694	0.0040
country	663	0.0039
voted	659	0.0038
democracy	643	0.0037
nodeal	632	0.0037

Word	Number of times used	Relative term frequency inverse document frequency
wins	100	0.0002
thegreatawakeningworldwide	23	0.0002
wonderful	59	0.0001
asianexit	20	0.0001
trumpit	20	0.0001

Hashtag	Number of times used
brexit	17,121
eu	827
borisjohnson	641
uk	469
yellowhammer	442
remain	343
brexitshambles	341
brexitchaos	337
bbcqt	310
nodeal	287

Vocabulary from stories classified as undecided excluding retweets:

Word	Number of times used	Relative term frequency
boris	1,034	0.0092
people	876	0.0078
parliament	859	0.0077
deal	735	0.0066
vote	663	0.0059
yellowhammer	657	0.0059
borisjohnson	650	0.0058
johnson	643	0.0057
nodeal	642	0.0057
labour	591	0.0053
election	559	0.0050
government	477	0.0043
remain	455	0.0041
voted	451	0.0040
mps	423	0.0038

Word	Number of times used	Relative term frequency inverse document frequency
decreases	29	0.0003
trumpbaby	26	0.0003
oldcorn	22	0.0002
jennings	11	0.0001
essay	29	0.0001

Hashtag	Number of times used
brexit	15,314
borisjohnson	676
eu	644
uk	552
yellowhammer	533
u	456
brexitshambles	390
brexitchaos	342
labour	294
bbcqt	260

Appendix H

```
                              ### LDA topic model ###
### Libraries used: topicmodels, slam, Rmpfr, tm. ###

tweetsText<- iconv(txtwork$text, to="utf-8")

tweetCorpus<-VCorpus(VectorSource(tweetsText))

dtmTM<-DocumentTermMatrix(tweetCorpus, control = list(stemming = FALSE, tolower =
TRUE, removeNumbers = FALSE, removePunctuation = FALSE))

term_tfidf<- tapply(dtmTM$v/row_sums(dtmTM)[dtmTM$i],dtmTM$j, mean) *
log2(nDocs(dtmTM)/col_sums(dtmTM>0))

median_tfidf<-summary(term_tfidf) [3]

dtmTM<-dtmTM[, term_tfidf>=median_tfidf]

forRemoval<-which(row_sums(dtmTM)==0,)

dtmTM<-dtmTM[row_sums(dtmTM) > 0,]

harmonicMean<-function(logLikely, precision = 2000L) {
+     llmed<-median(logLikely)
+     as.double(llmed - log(mean(exp(-mpfr(logLikely, prec = precision) + llmed))))
+     }

burnin = 1000

iter = 1000

keep = 50

sequ <- seq(2, 1000, 4)

Fitted_many<- lapply(sequ, function(k) LDA(dtmTM, k = 4, method = "Gibbs", control =
list(burnin = burnin, iter = iter, keep = keep)))
```

Appendix I

"#[operation]yellowhammer" frequent words and popular hashtags:

Word	Number of times used	Relative term frequency
yellowhammer	9,884	0.0813
worst	4,048	0.0333
government	3,919	0.0323
read	3,589	0.0295
media	3,433	0.0283
scenario	3,432	0.0282
understand	3,202	0.0264
stupid	3,169	0.0261
unredacted	3,145	0.0259
agreement	3,141	0.0258
printed	3,137	0.0258
guidance	3,135	0.0258
deal	1,806	0.0149
operationyellowhammer	1,368	0.0113
die	1,128	0.0093

Hashtag	Number of times used
yellowhammer	9,655
brexit	4,693
operationyellowhammer	1,366
prorogation	845
blackswan	180

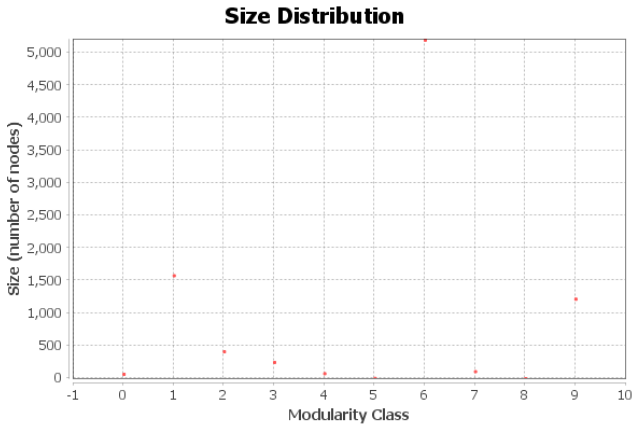
nodealbrexit	179
borisjohnson	159
nodeal	159
stopthecoup	88
brexitshambles	80

Appendix J

1. Modularity Report

Parameters: -Randomize: On; -Use edge weights: On; Modularity: 0.545

Results: -Modularity with resolution: 0.545; -Number of Communities: 10



Algorithm:

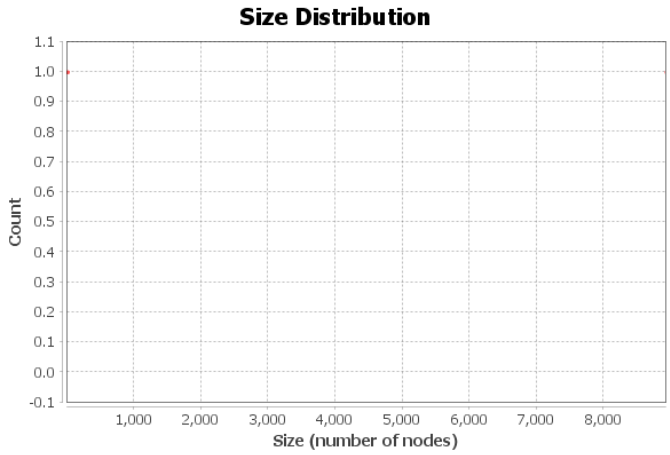
Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, Fast unfolding of communities in large networks, in Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P1000 Resolution:

R. Lambiotte, J.-C. Delvenne, M. Barahona Laplacian Dynamics and Multiscale Modular Structure in Networks 2009

2. Connected Components Report

Parameters: -Network Interpretation: directed

Results: -Number of Weakly Connected Components: 3; -Number of Strongly Connected Components: 8930



Algorithm:

Robert Tarjan, Depth-First Search and Linear Graph Algorithms, in SIAM Journal on Computing 1 (2): 146–160 (1972)

Graphs taken from Gephi (Bastian et al., 2009).

Bastian M., Heymann S., Jacomy M. (2009). *Gephi: an open source software for exploring and manipulating networks.* International AAAI Conference on Weblogs and Social Media.