UiA University of Agder

Master

Fake News on Twitter related to the Refugee Crisis 2016: An exploratory case study

Tor Ole Lerbæk
Bjørn Erik Vik Olsen

SUPERVISORS
Tim A. Majchrzak
Jaziar Radianti

University of Agder, 2020
Faculty of Social Science
Department of Information Systems

Preface

This master thesis marks the end of the master's degree in Information Systems at the University of Agder 2020. It has been challenging, as well as a rewarding journey to undergo. New privacy regulations led to a lengthy Data Protection Impact Assessment (DPIA), and in combination with COVID-19, this study saw its fair share of challenges along the way.

The topic of fake news is something that is interesting, frightening, and currently very relevant for both academia as well as those interested in the public debate to study. To be able to work on a topic like this has been a great opportunity, and hopefully, some of the preliminary work conducted here can aid future research on the subject.

Different people deserve thanks for their support throughout the study.
1) We would like to thank our supervisors, Professor Tim A. Majchrzak and Associate Professor Jaziar Radianti. Thank you for all your help, emails, chats, video meetings, and general enthusiasm for the topic. Your guidance was invaluable to us, both in terms of knowledge of the subject area, as well as getting us through the tough times. 2) Thanks to the team at the University of Duisburg-Essen for letting us use the dataset and enable us to analyze an extensive amount of novel data. 3) Thanks to Professor Øystein Sæbø for personal involvement to ensure that all of the students were seen and heard through the process of their thesis - something which undoubtedly raised the quality of the work conducted. 4) Lastly, we would like to thank Eva Payne from Norwegian Centre for Research Data, who took personal contact to both explain and aid us through the process of completing the DPIA process. No doubt, she stepped out of her way to ensure that the study could be conducted.

Kristiansand,
04.06.2020

Tor Ole Lerbæk

Tor Ole Lerbæk

Bjørn Erik Vik Olsen

Bjørn Erik Vik Olsen

Abstract

Fake news has, in recent years, gained traction in the public media and as a research topic. Events such as the U.S 2016 presidential election, Brexit, the COVID-19 pandemic, amongst others, have seen traces of large amounts of fake news in social media. Social media sites like Twitter have enabled individuals, politicians, and companies to share content and opinions with a large number of people across the globe. This opportunity for mass communication has also led to Twitter becoming a place for fake news sharing. Various narratives by various actors partake in the same public discussions, and knowing what is true and what is fake is increasingly difficult.

The purpose of this study was to examine and analyze a previously not studied dataset of 14.3 million tweets related to the 2016 refugee crisis and attempt to find traces of fake news. The research approach chosen was an exploratory case study with mixed data analysis. The analyzed focused on finding the characteristics of tweets, the most prominent topics, identifying fake news, some of the actors (webpages) spreading fake news, and classify the type of fake news. To identify what content was fake, an extensive amount of literature in combination with three fact-checking services were utilized.

The findings reveal that the American presidential election was a very prominent topic; nevertheless, topics such as Brexit, the refugee crisis, Syrian conflict as well as other geopolitical events were present through several of the contextual findings related to the tweets. The most liked and shared tweets often saw a relation to Trump's election campaign and were susceptible to fake news. The webpages (URLs) used within tweets also saw a tight connection to the election, with substantial amounts of fake news. The content classified as fake news spiked throughout the year, with an increase towards election day. The (emotional) focus of the fake news was negative, aiming to target individuals and fearmongering towards refugees. A majority of the fake news utilized fabricated information, with little to no base in reality.

The results from the study confirm the massive political spamming related to the election previously shown in literature and some of the properties of the most prominent tweets. The refugee crisis became a global discussion and tied with large amounts of fake news aiming to discredit individuals and political opponents.

The term *fake news* is also problematic, and dependable on actor its definition and intention changes. With the eight types of fake news classifications used in this study, we argue that if the term should be used, it should be used in combination with these types. By doing so, fake news would be the case when an actor misleads the receiver of the news by design - regardless of severity.

Keywords: Twitter, Social Media, Topic Modeling, Fake News, Crisis, Refugees, Crisis informatics, Information Diffusion, Fake News Detection

Contents

List of figures

List of tables

1 Introduction
Fake news, despite its validity and definition (Gelfert, 2018; Murphy, 2018; Tandoc Jr, Lim, & Ling, 2018), has quickly become a topic of both interests and worry, much helped by the introduction and growth of social media (Allcott & Gentzkow, 2017; Shu, Sliva, Wang, Tang, & Liu, 2017). One of the strengths of social media is the low-cost economical cost and human efforts to partake in the public debate. At the same time, it also helps facilitate the spread of fake news as users of social media uncritically consume and spreads fake news throughout the media(s) (Gelfert, 2018; Shu et al., 2017). There have been many attempts to classify the term *fake news*, and we will provide a further granularity of this term in sections 2 and 3 of this thesis. To get a starting point, we refer to a literature review conducted by Gelfert on the subject:

> *Despite being a new term, 'fake news' has evolved rapidly. This paper argues that it should be reserved for cases of deliberate presentation of (typically) false or misleading claims as news, where these are misleading by design. The phrase 'by design' here refers to systemic features of the design of the sources and channels by which fake news propagates and, thereby, manipulates the audience's cognitive processes (Gelfert, 2018, p. 84).*

Context
While social media covers a range of mediums and platforms, the focus on this thesis has been on the platform Twitter, with further reasoning of why this selection (choice, approach) in paragraph 1.2. Some of the material utilized as a reference covers additional platforms, yet the maturity of the literature on the subject is still young (which we will elaborate on under chapter 2). This means that a broad approach to the research is utilized in this thesis.

As for the topic at hand, fake news in social media, this is set to the case refugee crisis 2015-2016, with data stemming from 2016 throughout an extensive Twitter tweet (posts) collecting conducted by the University of Duisburg-Essen and University of Agder related to the Rise SMA project (RISE_SMA, 2020).

This thesis does not focus on further establish new definitions or confirm already existing terminologies or point out the societal challenges imposed by fake news. Instead, a more exploratory case study is performed to study the extends of fake news presented on Twitter related to the public debate regarding the refugee crisis and the characteristics of this type of information.

The reasoning for choosing this topic
The motivation for conducting this study is fourfold; Firstly, it is within the interest field of the group's members, especially the issues that arise when dealing with fake news in the context of immigration. Secondly, the literature suggests a range of different approaches and topics for future research. It contains several gaps on the subject, and suggest several different future research directions, something which is backed up by a literature review (Al-Rawi, Groshek, & Zhang, 2019; Pal & Chua, 2019; Zannettou, Sirivianos, Blackburn, & Kourtellis, 2019). Thirdly, an assumption by the group is that exploratory studies that look at the actual actions of people on Twitter yield more precise results than those that ask people directly. We after trailed and expanded on the research, "*Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior*" (Talwar, Dhir, Kaur, Zafar, & Alrasheedy, 2019), by surveying over 100 students at the University of Agder, something which yielded inconclusive results. Based on this, the group believed more accurate results are to be found in the actual actions of people and not what they say under a study. While it was, and still is, not the

focus of pointing out this misalignment of words and actions, the problem is not new (Coulombe, 2014; Hair, Ringle, & Sarstedt, 2011). Furthermore, the recent situation with the COVID-19 pandemic has changed a lot of the world in a short amount of time, including social media. Users are sharing a large amount of information on Twitter, including fake news, which shows the timeliness and relevance of the topic (Garrett, 2020).
Lastly, as the University of Agder has both the Centre For Integrated Emergency Management (CIEM, 2019) in the combination of people involved in the RISE_SMA project (RISE_SMA, 2020), this enables resources sharing in the form of people and knowledge which the group sees as a benefit for the study.

Research problem

The thesis is exploratory, due to the nature of the topic, and the dataset available. The need for an open-ended approach is therefore chosen, as we do not know what the dataset contains. Thus, the research questions reflect this in their form:

1) What were the most liked and shared tweets in a dataset related to the 2016 refugee crisis, and what were the characteristics of tweets?
2) Using open-source services and a mixture of manual and automatic fake news detection, is it possible to find the presence of fake news in the tweets, and if so, what type of fake news is present?

Structure and content of the paper

This master thesis is structured as follows. In **Section two,** we look at the background literature, to get a better understanding of the context and field. We present the theoretical foundation and the literature review. **Section three** gives an overview of the research method, research strategy, software, data collection, data, analysis, and ethical challenges. **Section four** presents the findings of the data analysis. **Section five** discusses the findings. **Section six** is the conclusion where we summarize the findings, look at the limitation within this study, and possible future research. **Section seven** is the references and at the end appendices.

2 Literature

To get an understanding of what the current literature on the topic states, we identified the core themes of the thesis. Figure 1 illustrated the identified overlapping themes, which sets the direction of literature collecting. Please note that a range of synonyms was utilized, something which means that the search string is quite more refined than what illustrated. While both Fake News and Social Media had iterative small-scale literature reviews conducted, the core search held the primary focus. At the same time, the interconnected theme crisis did not receive a literature review beforehand. The subparagraph of this chapter will cover this process in more detail, with reasonings of the choices and utilization within the thesis.

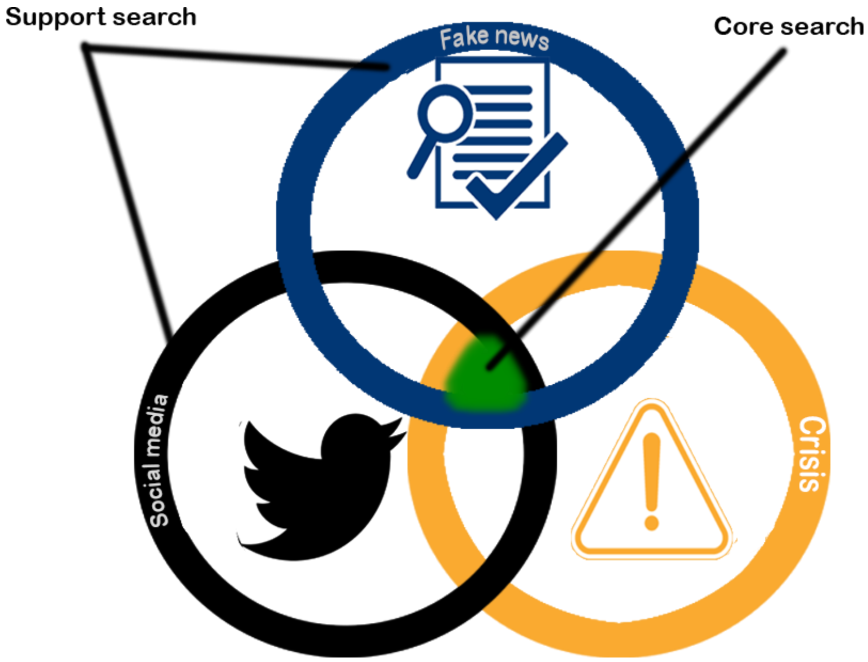


Figure 1 – Master thesis research topic, much inspired by Frank Danielsen (Danielsen, 2019)

2.1 Background: Social media

Social media, or social network and online social network which is often referring to the same (Heidemann, Klier, & Probst, 2010; Nied, Stewart, Spiro, & Starbird, 2017; Zakharchenko, Maksimtsova, Iurchenko, Shevchenko, & Fedushko, 2019). It has gained traction in recent years and is continuously in expansion. While the actual numbers of users are hard to measure, the forecasts for 2020 is 2.96 billion unique social media users worldwide (Statista, 2020). While one should proceed with caution when the source prevents the extraction of data sources (unless signing up for a subscription), Statista seems to be most reliable in this regard. It is has been used in IS literature before (Swar & Hameed, 2017), couplet with the fact that other reliable data sources on the topic seem lacking. Furthermore, back in 2014, Facebook reported having one billion users worldwide (Greenwood, Perrin, & Duggan, 2016). With the sheer amounts of different social media platforms, the numbers from Statista seems to be within reasonable estimates for this context.

Social media has seen a shift from being an online phenomenon in isolation (as a result of this often referred to as *Virtual Worlds* (Berente, Hansen, Pike, & Bateman, 2011; Urquhart & Vaast, 2012)). It has become a factor that affects our social behaviors in the real world (Urquhart &

Vaast, 2012). Likewise, one finds effects in the opposite direction as well, where the real world affects social media. Furthermore, social media communities have life cycles for the involved members similar to those of the real world (Füller, Hutter, Hautz, & Matzler, 2014). The lines between what is online and what is "real" seem to blur more as the time passes, coupled with the introduction of new platforms and an ever-increasing more *digital native* population (Prensky, 2001). It is crucial to keep in mind that while digital natives might primarily be the young that have grown up with this technology, digital natives are not exclusive to this group. Instead, dependable on other variables such as exposure to social media and experience, amongst others, are factors to take into consideration when addressing this group (Helsper & Eynon, 2010).

Social media is a collection of a range of different platforms and services. To point at one single platform or service would give a misleading picture of its diversity and reach. Rather, focusing on whether it facilitates social interactions or not. While tons of literature points out the different characteristics social media, to get a conceptual understanding we have chosen to use the foundation created by Cooke & Buckley (2008) to highlight the key aspects which have contributed to the growth of social media:

1) The growth of user-generated content which blurs the difference between professional and amateur created content (Cooke & Buckley, 2008). In this context, amateur refers to users that have neither education nor fulltime jobs related to the focus area. The primary example used here is how the media has altered its role in the direction of being a medium for facilitating the spread of news.
2) Media is moving towards the direction of what people want to read and focus on, rather than presenting ("forcing") news at people. The agenda of the public debate sets the lead of the media picture.
3) The structure of media has shifted from a monolithic model to a more fragmented/decentralized model. People can read, see, listen, and interact at a preferred level of depth that in advance is not decided by the conveyer of the news item.
4) The social interaction in terms of ranging, commenting, reviewing, and general response related to the content created by the media holds a direct impact on the success of the material. By success, it is in this context meant traffic generated, accessed by numbers, sharing, and actuality.

The result of this transformation highlighted above is that social media has become the preferred way of consuming (gathering, reading, sharing, and discussing) news (Zannettou, Sirivianos, Blackburn, et al., 2019).

2.2 Background: Fake News

Fake news on social media has gained much of its traction with the help of the American presidential campaign in 2016 (Allcott & Gentzkow, 2017; Budak, 2019; Fourney, Racz, Ranade, Mobius, & Horvitz, 2017), amongst others. Although the common understanding of the word *Fake News* is new, the spread of false information by the internet is not a new phenomenon (Murungi, Purao, & Yates, 2018). Sharing of incorrect or otherwise wrong information was still present long before the times of the internet. However, with the introduction of social media platforms, it has made it easier for people to consume and spread information on a much larger scale than previously possible.

Fake news spreads in various ways, and one very prominent example of this is those covered by known actors such as Twitter accounts linked to presidents (Bolsonaro, 2020; Trump, 2019). Another example is automated processes such as bots that solely operate to spread fake news

(Bastos & Mercea, 2019; Ferrara, Varol, Davis, Menczer, & Flammini, 2016; Nied et al., 2017). Although international actors seem most prominent here, recent studies show indications that local, (as in national users where the news item originates from), disseminates fake news on a larger scale than international users (Grøtan et al., 2019; Kogan, Palen, & Anderson, 2015).

In literature, the term *Fake news* has its origin from terms such as misinformation, disinformation, rumor, or hoaxes, amongst others (Ferrara et al., 2016). It is first in 2016 marks the real entrance of the term in the literature.
As mentioned in the introduction by quoting Gelfert (Gelfert, 2018), one could say fake news is false by design and intention. By now, one would think that people or automated processes always do an intentional spread of fake news with negative agendas. However, there are more shades to this, and arguable fake news should be viewed at as a collection of various terms which different aims and impacts, which will be elaborate into eight main categories below.

1) **Fabricated**. Are news items that have no roots stemming from the real world. An example of this can be the story about Hillary Clinton adopting an alien child (Heller, 2014; Zannettou, Sirivianos, Blackburn, et al., 2019).
2) **Propaganda.** News with a political focus and usually contains stories with little to none hold from the real world. The intention is often at staining the reputation of a politician, political party, or nation. Can affect democratic processes such as presidential elections or general elections such as Union election (Bastos & Mercea, 2019; Budak, 2019; Ferreira, 2018)
3) **Conspiracy Theories.** News that “explains” events and situations without containing any real evidence or evidence with such a lack of quality that it borderlines the two previous categories (Zannettou, Sirivianos, Blackburn, et al., 2019). The message of the new piece often plays at the trust of people by using phrases such as “what the news won't tell you,” which reduces the burden of proof needed to gain the trust of people.
4) **Hoaxes.** Type of news that contains elements that hold some truth in it, but at the same time contains false or misleading accusations. Sometimes referred to as half-truths (Zannettou, Sirivianos, Blackburn, et al., 2019). One example here might be news that contains the announcement of the death of a celebrity, which turns out to be bogus.
5) **Biased or one sided.** Type of news presented in such a way that it conforms to a biased opinion, with little to no room to give a balanced view of the topic at hand. Often found in echo chambers, with the subject often being individuals, political parties, events, or situations (Zannettou, Sirivianos, Blackburn, et al., 2019).
6) **Rumors.** While being a lesser form of false news, it has proven to be very useful in social media. Do not have to be all-good or all-bad, but can aid in creating uncertainty and uproar, which taints the public debate. One very prominent example here is the bombing of the Boston Marathon in 2013 (Kate Starbird, Maddock, Orand, Achterman, & Mason, 2014; Vosoughi, Roy, & Aral, 2018; Zannettou, Sirivianos, Blackburn, et al., 2019).
7) **Clickbait.** News containing misleading thumbnails and headings compared to the actual news piece. Its seen as the least harmful false news category. Alas, couplet with how many people read the news today (seeing thumbnail and reading header), this has proven to be an effective way of distributing false news in social media (Zannettou, Sirivianos, Blackburn, et al., 2019).
8) **Satire.** News containing humor and irony can be viewed as a double-edged sword. On the one hand, the news presented bears little to no problem as it intends to be a humoristic piece of information, on the other hand when spread in social media couplet with the same human behaviors as in clickbait, the original intention often is lost (Zannettou, Sirivianos, Blackburn, et al., 2019). A somewhat recent example is the case where the

Norwegian radio show called Radioresepsjonen (aired 31.10.19) made a satirical piece where the aim for the hosts was to act as embarrassing as possible by being a science-denying racist. It was warned before the section and after it, that the intention was to be as awkward as possible. Nevertheless, it was taken out of context and spun around in social media, which ultimately resulted in the piece removed from the coverage (Dahl. Ingvill Dybfest & Wergeland. Atle Jørstad, 2019; Kalajdzic. Pedja & Vigsnæs. Maria Knoph, 2019).

2.3 Background: Crises

When presented with the word *crisis*, associations to terminologies such as both *emergency* and *disaster* are not uncommon (Al-Dahash, Thayaparan, & Kulatunga, 2016; UNDRR, 2020). There is a need to define the use of the terminology we are looking at as crisis in this text gives the context of fake news in social media being studied. Following the studies by Al-Dahash et al., figure 1 was developed to illustrate how the three terminologies overlap. This text does not delve into the transition from crisis to disaster or emergency to disaster. Instead, it focuses on establishing which perspective the text utilizes.

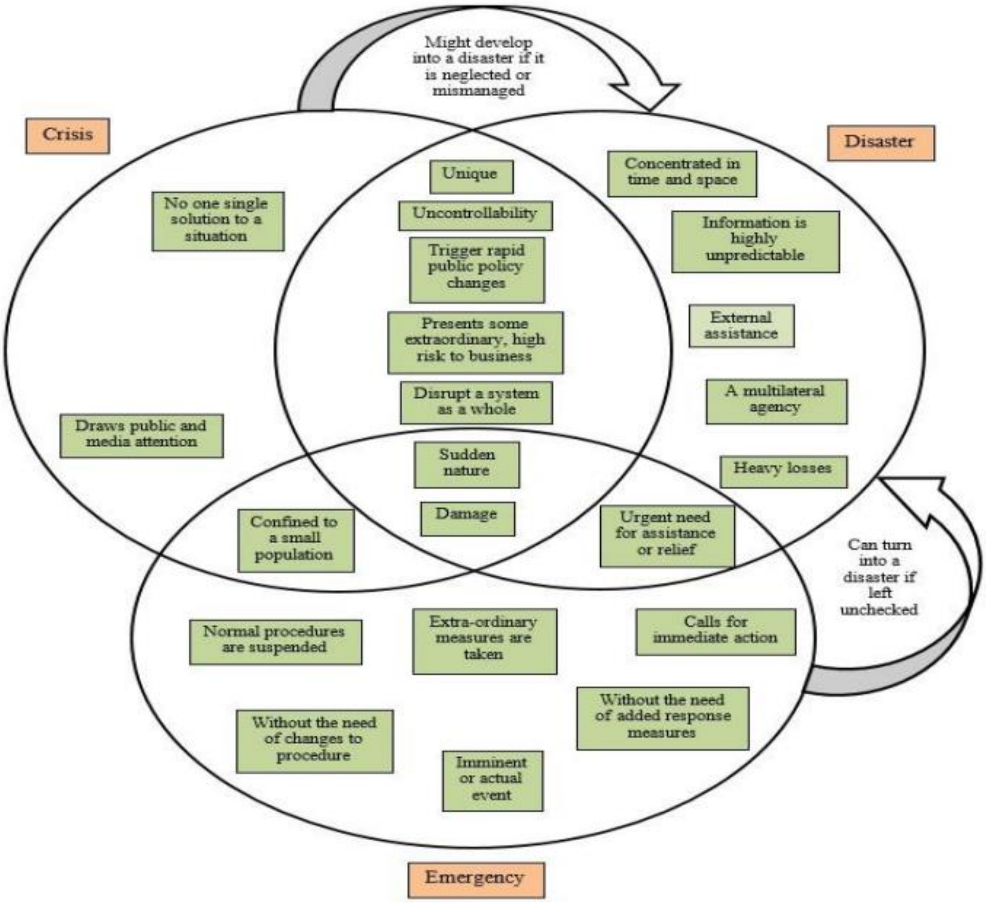


Figure 2 - Set Diagram to illustrate the interconnection of terminologies (Al-Dahash et al., 2016, p. 1197)

Furthermore, from same article Al-Dahash et al. creates figure 3, dwelling further into crisis in isolation:

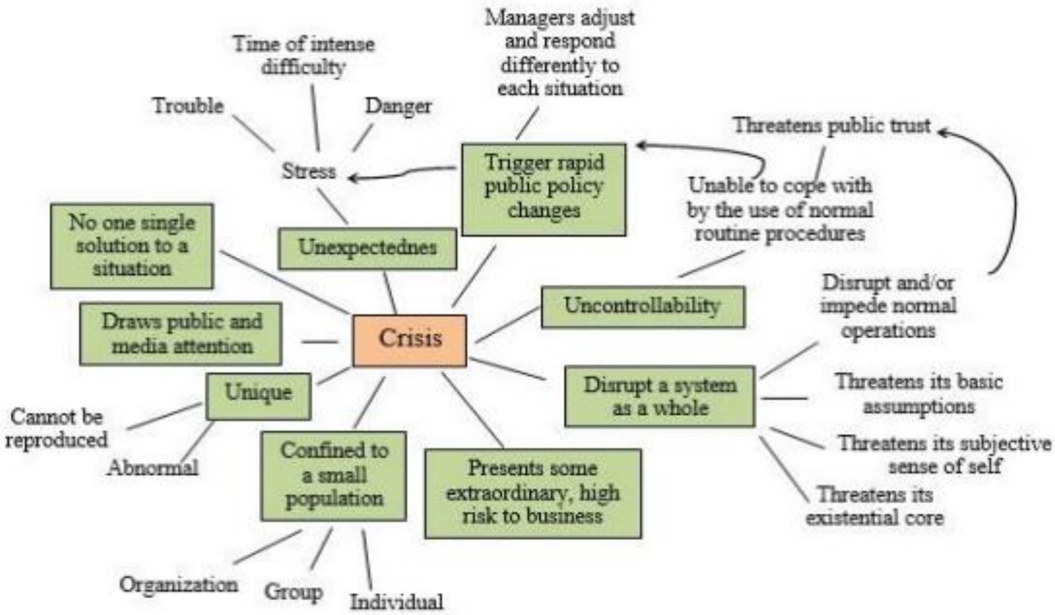


Figure 3 - Cognitive map of crisis (Al-Dahash et al., 2016, p. 1195)

Studying the figure above (figure 3), several of the elements visualized above is of relevance when addressing crisis in the context of the refugee crisis, as elaborated below:
1) *No one single solution to a situation* is prominent. The Refugee Crisis also referred to as the migrant crisis, saw its initial period of focus from 2015, – arguably ongoing, due to several geopolitical situations such as Syrian civil war (Holmes & Castañeda, 2016), violence in Afghanistan and Iraq, abuses in Eritrea, as well as poverty in Kosovo (BBC, 2016b). This situation led to a large(r than ordinary) number of people arriving the European Union from across the Mediterranean Sea or overland through Southeast Europe. BBC further highlighted this by creating the following graph based on data from Eurostat:

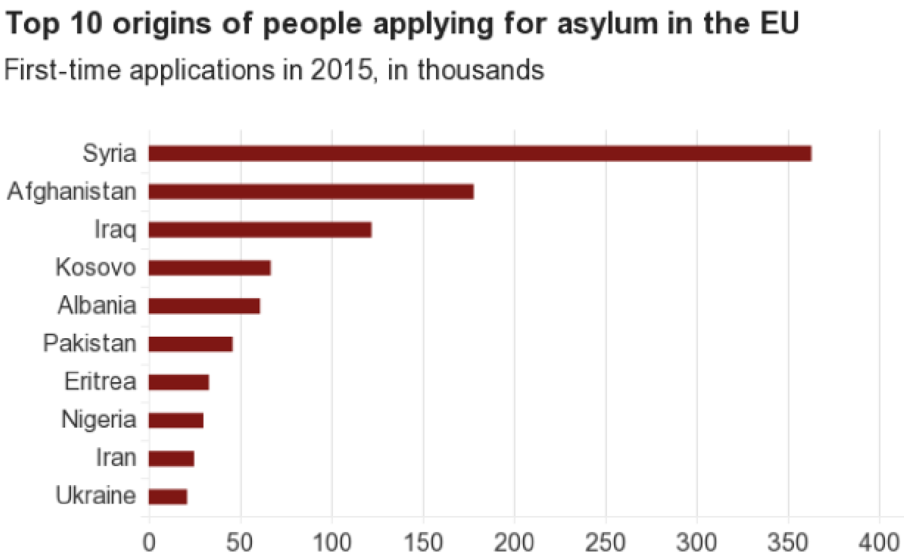


Figure 4 – Origins of people applying for asylum (BBC, 2016)

Going to the data source directly a similar chart can be viewed to highlight this increase in applications (here all countries within the EU combined to single graph):

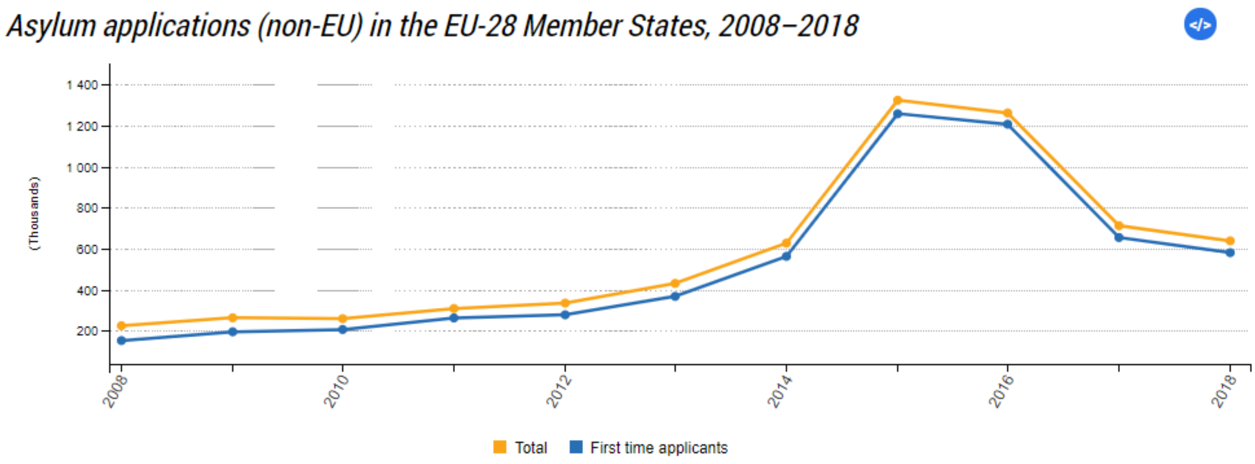


Figure 5– Asylum applications (non-EU), 2008-2018 (Eurostat, 2019)

2) It drew massively media attention, e.g., *Draws public and media attention* (Chouliaraki & Zaborowski, 2017; Georgiou & Zaborowski, 2017), which undoubtedly pave the way for classifying the situation as a crisis (Georgiou & Zaborowski, 2017). It affected both the media coverage (Berry, Garcia-Blanco, & Moore, 2016) as well as the public debate (Aigner, Durchardt, Kersting, Kattenbeck, & Elsweiler, 2017; Davidson & Farquhar, 2020), amongst others sources which will be covered more in-depth within chapter 3.

3) *Uncontrollability* was a particularly noticeable argumentation in the public debate, and political parties such as Fremskrittspartiet in Norway were often visible in the general discussion both on social media and traditional media to argue that the crisis was uncontrolled (amongst other claims), (Tjernshaugen, 2016). It is worth noting that while Norway is not officially part of the EU agreement, much of the legislation from the EU is shared with Norway as it falls within the European Economic Area (EEA) agreement.

4) *Disrupt a system as a whole / Presents some extraordinary, high risk to business*. The refugee crisis was by many, at least in Norway, viewed as a threat to the welfare system (Skjeggestad;, Sandvik;, & Omland, 2017; Tjernshaugen, 2016). This because the argumentation used was that refugees never would be able to achieve sustainable income for themselves and would be dependable on the welfare system throughout their remaining time of life. Worth noting here that the data material provided by Ny Offentlig Utredning (NOU) primarily focused on the long-term work for the government, but the public debate spun this around to fit the narrative. The cultural threat imposed by refugees was also a hefty debated topic in the general discussion (Lee & Nerghes, 2018).

5) *Unique*. Although immigration is not a new phenomenon in Europe (see image below), it was by the same reasons mentioned previously with the same type of argumentation spun as an extraordinary or unique, challenging situation for Europe as a whole.

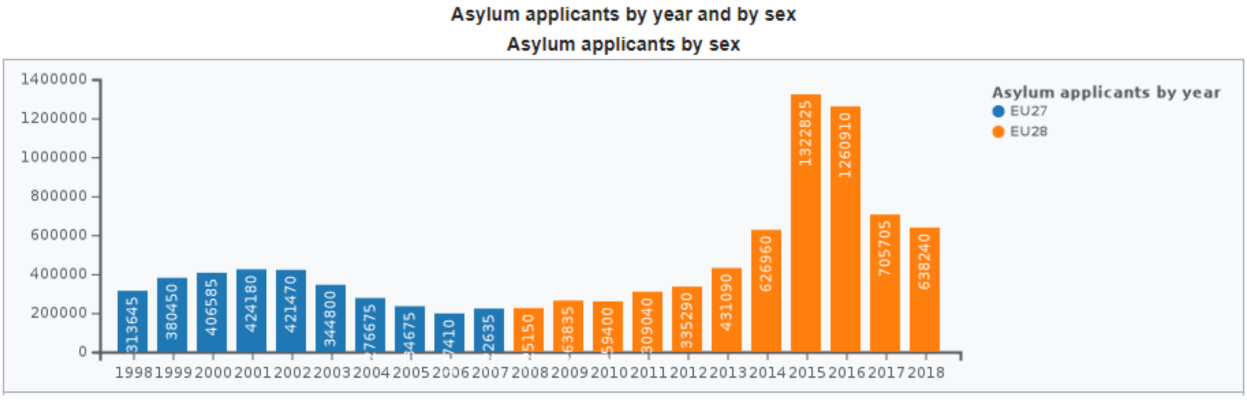


Figure 6 - Asylum applications. Data to the figure provided by Eurostat (Wikipedia, 2020)

6) *Trigger rapid public policy changes.* As the public perception of the crisis changed, so did policies. In September 2015, both Hungary and Croatia closed its borders to refugees (BBC, 2015a, 2015b). Next in line were Bulgaria in February 2016 (ITV, 2016). Couplet with the closing of borders, policies changed across Europe, and more hardline rhetoric appeared (Berry et al., 2016).

As for the remaining elements listed on the cognitive map (figure 3), these are linked either within the first four arguments or stand as their own as the public debate would talk about them in a more or lesser degree when the topic was/is addressed.

2.4 Bridging Social Media, Fake News, and Refugee crisis

As previously covered under 2.1, social media have enabled people to partake a more active role in conveying news, or said differently, decentralizing the news coverage (Cooke & Buckley, 2008). Celebrities, acquaintances, friends, or/and family members now act as a source for the news, which poses challenges in identifying what is real and what is false news. The level, both in terms of the sheer amount of news consumed and shared, and also trust and credibility, is now of such a character that human intuition alone is insufficient to identify what is real and what is false (Burbach, Halbach, Ziefle, & Valdez, 2019). Social media sometimes is referred to as a double-edged sword; to enable individuals to find news and groups of people that share similar views at the costs of critical voices and balanced opinions. It can contribute to the rise of echo chambers (M. Del Vicario, Gaito, Quattrociocchi, Zignani, & Zollo, 2018), and filter bubbles (Flaxman, Goel, & Rao, 2016; Spohr, 2017). It is worth noting that social media alone can be blamed for neither false news or echo chambers, as this is dependable of a range of factors such as sociopolitical actions (Del-Fresno-garcía & Manfredi-Sánchez, 2018). Still, the platform(s) have regardless enabled and facilitated that fake news has an arena to both publish and spread on a larger scale than traditional news (Pal & Chua, 2019). Furthermore, as digital natives become a broader part of the population (Helsper & Eynon, 2010; Prensky, 2001), increasing amounts of time are used on social media platforms (Broadbandsearch.net, 2020).

As for crisis, even when the context is set not to emergencies or disastrous short-time events, it is within reason to assume that the way people collect, share and discuss news regarding this with little different than traditional news in social media (Broadbandsearch.net, 2020; Greenwood et al., 2016; Swar & Hameed, 2017). As false news has become a widespread challenge in social media with especially Twitter being a platform of study (Al-Rawi et al., 2019; Bastos & Mercea, 2019; Buntain & Golbeck, 2017), it stands to reason to couple these together when tackling challenging geopolitical situations such as the refugee crisis. Lastly, studies have shown that the topic refugees in social media spur the public debate on these platforms, something that in turn increase the likelihood of the spread of fake news (Ayers, Hofstetter, Schnakenberg, & Kolody, 2009; Davidson & Farquhar, 2020; Li, Dombrowski, & Brady, 2018).

We argue that by using the literature above and the research analyzed in the central literature review below, it is valid to connect and couple these thematics together for the remaining of the text. As pointed out, more factors need to be taken into consideration when trying to explain these three thematics in detail, yet, this text focuses on addressing the usage and extend of false news in social media set in the context of crisis rather than focusing on all the interconnected aspects in depth.

2.5 Literature Review

In the following sections, we will elaborate on the process surrounding the literature review. The literature review chapter has been divided into subsections for increased readability.

2.5.1 Approach literature review

Figure 7 illustrates the overall process of collecting and selecting literature throughout the review. At the same time, the more fine-grained details are listed in their subparagraphs later in this chapter. Following known methodology (Kitchenham, 2004; Kitchenham et al., 2009; Webster & Watson, 2002), the aim of the review section in combination with the background theory chapter supplied above is to build a theoretical understanding and foundation on the field of study for the study group. Take notice of the strong-bordered “Preliminary literature study.” These are 18 articles identified in the previous year’s literature review by the group that has undergone the same iterations and filtering as those located on the left of it in the figure 7. It is also the reason why it is not further iterated on in the processes at the very right of the figure 7 compared to the 452 other articles.

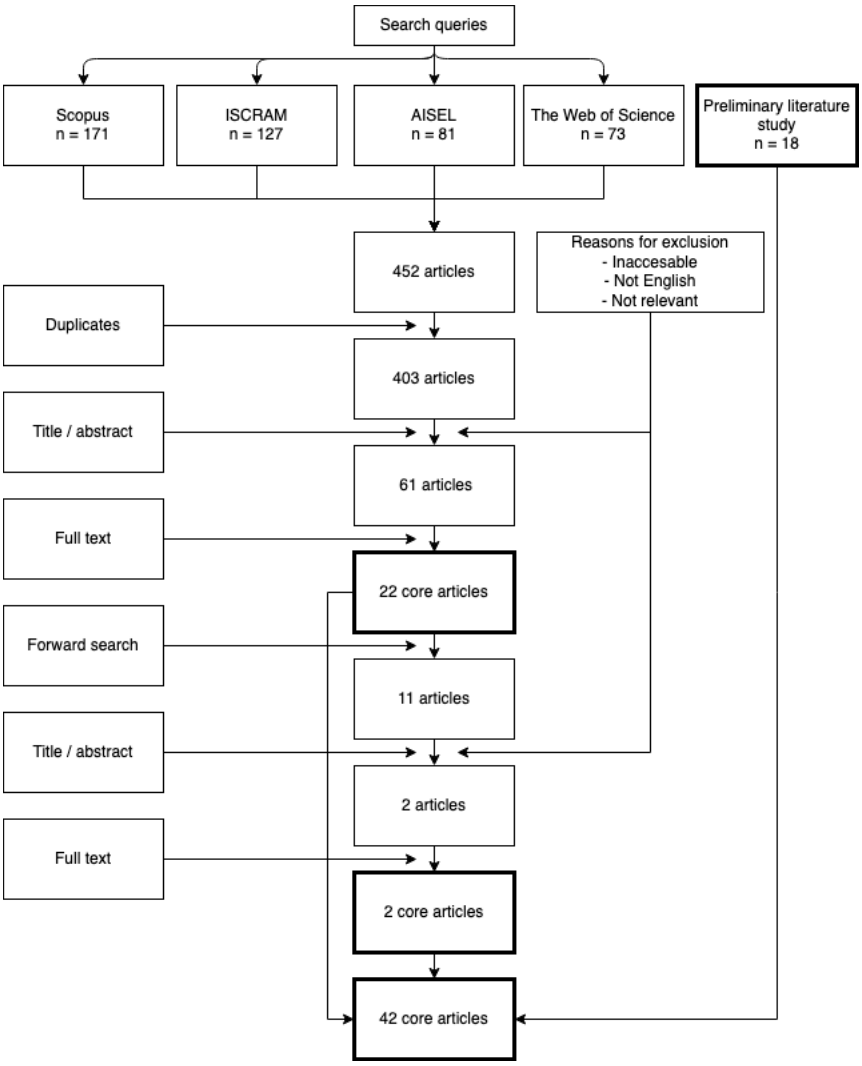


Figure 7- Literature review process diagram

2.5.2 Inclusion and exclusion criteria

Table 1 illustrates the inclusion and exclusion criteria were created. This table lists up the different criteria that the articles need to adhere to for either being included or excluded, which is of great help when trying to categorize a rather large number of articles effectively.

Criteria	
Inclusion (i)	
i1	Searchable at either Scopus, ISCRAM, Aisel, or Web of Science*
i2	Peer-Reviewed
i3	Name, abstract and text is coherence with the theme(s) **
i4	Published in journals or conferences
Exclusion (e)	
e1	Other languages than English
e2	Irrelevant thematic **
e3	Not complete (e.g., not reviewed, not fully published)
e4	Lack of sources in their argumentations
* Some of the articles had to be looked up at Google Scholar to get the full text	
** Overlapping process, not strictly limited to one category	

Table 1 - Inclusion and exclusion criteria

2.5.3 Inclusion criteria process

In the initial phases of the study, we used a reasonable amount of time at refining the search string, or preliminary searching. This extra effort to ensure that the reach of the search as well as the removal of some of the most prominent noise. As previously discussed, the topic of study is still young, which very much reflects the literature. When the final search string on all databases was complete (which is first criteria i1 for inclusion), the task shifted to analyzing whether it was peer-reviewed, name, abstract as well as full text to compare it against the thematics, through 4 iterations. Falls within i2 and i3 inclusion criteria. Below a small sample of this iterative process is attached (from the latest iteration of articles brought in from Scopus).

	Authors	Title	Year	Source title	Cited by	DOI	Link	Abstract	Document Type	EID
2	Oh O., Kwon K.H., Rao H.R.	An exploration of social media in extreme events: Rumor theory and twitter during the HAITI earthquake 2010	2010	ICIS 2010 Proceedings - Thirty First International Conference on Information Systems	83		https://www.researchgate.net/profile/Kyounghee_Kwon/publication/221599216_An_Exploration_of_Social_Media_in_Extreme_Events_Rumor_Theory_and_Twitter_during_the_Haiti_Earthquake_2010/links/0a85e533f3b71a6e7c000000.pdf	Due to its rapid speed of information spread, wide user bases, and extreme mobility, Twitter is drawing attention as a potential emergency reporting tool under extreme events. However, at the same time, Twitter is sometimes despised as a citizen based non-professional social medium for propagating misinformation, rumors, and, in extreme case, propaganda. This study explores the working dynamics of the rumor mill by analyzing Twitter data of the Haiti Earthquake in 2010. For this analysis, two key variables of anxiety and informational uncertainty are derived from rumor theory, and their interactive dynamics are measured by both quantitative and qualitative methods. Our research finds that information with credible sources contribute to suppress the level of anxiety in Twitter community, which leads to rumor control and high information quality.	Conference Paper	2-s2.0-84870963741
3	Castillo C.	Big crisis data: Social media in disasters and time-critical situations	2016	Big Crisis Data: Social Media in Disasters and Time-Critical Situations	73	10.1017/9781316476840	https://pdfs.semanticscholar.org/f631/dc866f369bc74f9cd925d20b4781fb0b3357.pdf?_ga=2.2980724.595849340.1578559089-898535808.1568793242	Social media is an invaluable source of time-critical information during a crisis. However, emergency response and humanitarian relief organizations that would like to use this information struggle with an avalanche of social media messages that exceeds human capacity to process. Emergency managers, decision makers, and affected communities can make sense of social media through a combination of machine computation and human compassion - expressed by thousands of digital volunteers who publish, process, and summarize potentially life-saving information. This book brings together computational methods from many disciplines: natural language processing, semantic technologies, data mining, machine learning, network analysis, human-computer interaction, and information visualization, focusing on methods that are commonly used for processing social media messages under time-critical constraints, and offering more than 500 references to in-depth information. © Carlos Castillo 2016. All rights reserved.	Book	2-s2.0-84978407174

Figure 8- Iterative filtering of literature

The next step was to check whether the articles found were published, which is criteria i4. While most of the articles were published, a few had to be manually removed later in the process.

2.5.4 Exclusion criteria process

Much of this process is similar to section 2.6.1, particularly the iterative cycle of inclusion criteria number two (i2 / e2). The first unique exclusion criteria are e1), other languages than English.
While the search filter removed most of the none-English articles, some concerning false news in the French and Ukrainian election slipped by. e3) relates to the state of the material. Some of the collected literature were preliminary, meant to shape further work. Thus, they were very short and focusing more on the idea / what to do, rather than doing the actual work. e4) tackles the lack of sources in the article. While the number of sources is not a direct indicator of the quality alone, having low amounts such as +- 10 couplet with little citations in the text gives less-than-ideal credibility to the text. Several articles were removed based on this criterion. Still, no article was removed due to e4 alone, as it usually was a combination of lacking quality/theme in text couplet with few sources, which reduced the perceived credibility of the article.

2.5.5 Metadata

As previously stated, the literature on the field of fake news has seen a significant increase over the last years. This trend seems to continue with an ever-increasing number of published articles and conference papers. It is worth noting here that the graphs and figures represent the Scopus search engine. While similar results can be made based on the other search engines used, these do not have (to best to the knowledge of the group) an automated visual representation at the same level as Scopus. These graphs and figures are based on the results of the main search string below. Furthermore, as a literature review was conducted autumn 2019 on the subject of the spread of fake news in social media, the results of these two searches have been merged in the final source list. We refer to Appendix D to see the first search string and visual models for the original search string in isolation.

TITLE-ABS-KEY ("Social Media" OR "Social Network") AND ("fake news" OR "false news" OR "false information" OR "false facts" OR "disinformation" OR "misinformation") AND ("refugee crisis" OR "refugee" OR "crisis") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (SUBJAREA , "COMP") OR LIMIT-TO (SUBJAREA , "SOCI") OR EXCLUDE (SUBJAREA , "ARTS") OR EXCLUDE (SUBJAREA , "PSYC") OR EXCLUDE (SUBJAREA , "BUSI") OR EXCLUDE (SUBJAREA , "ECON") OR EXCLUDE (SUBJAREA , "AGRI") OR EXCLUDE (SUBJAREA , "ENER") OR EXCLUDE (SUBJAREA , "ENVI") OR EXCLUDE (SUBJAREA , "MATH") OR EXCLUDE (SUBJAREA , "SOCI") OR EXCLUDE (SUBJAREA , "DECI") OR EXCLUDE (SUBJAREA , "MEDI") OR EXCLUDE (SUBJAREA , "EART") OR EXCLUDE (SUBJAREA , "MATE") OR EXCLUDE (SUBJAREA , "HEAL") OR EXCLUDE (SUBJAREA , "NURS") OR EXCLUDE (SUBJAREA , "PHYS")) AND (LIMIT-TO (LANGUAGE , "English"))

Figure 9 – Final search string on Scopus

The final search string contains several filters to prevent undesirable results. It is a result based on three iterations of search strings where different filters were applied alone and in isolation to establish the most effective search string for this thesis. There are four filters in-effect:
1) The string looks on the overlapping thematics of the thesis by utilizing several synonyms, 2) We only wish Scopus to present final/published articles and conference papers, 3) Exclusion of subject areas which holds little interest to the thesis and lastly, 4) Only English results.

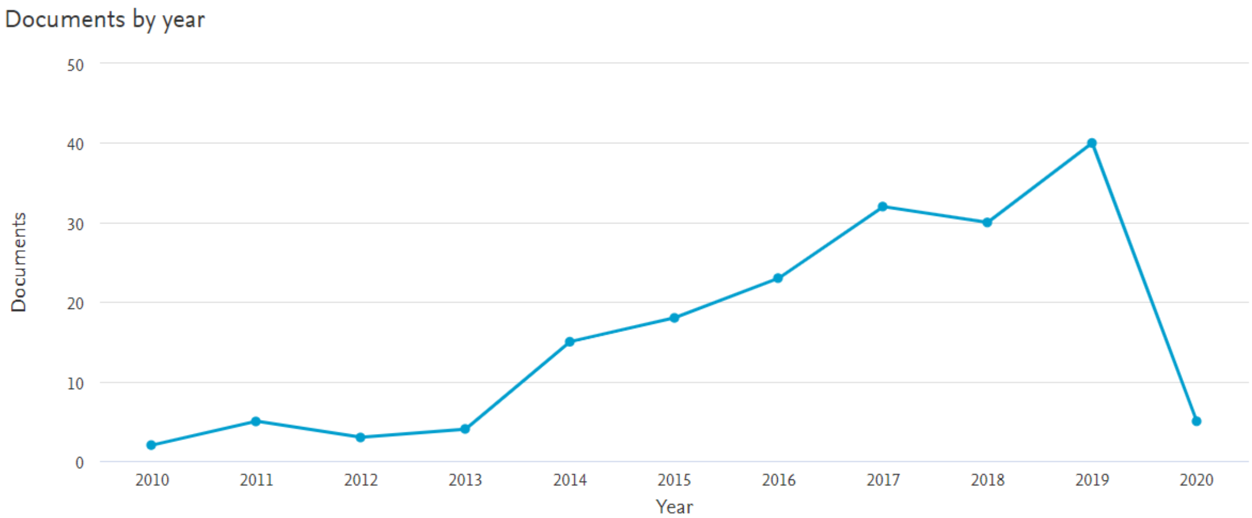


Figure 10 – Documents per year based on the main string, spring 2020 (Scopus, 2020a)

The documents by year graph highlight the rapid development of the field. It is ever-increasing (despite a small dip in 2018), and compared to the two documents by year with the search string above, it was in 2019 40 documents by year. There are spikes in the intervals 2013-2014, 2016-2017, and 2018-2019. Events such as Alberta floods, Lac-Megantic disaster, Boston Marathon Bombing, and Typhoon Haiyan might have contributed to the spike seen in 2013-2014, while the 2016-2017 interval could be related to the American presidential election.

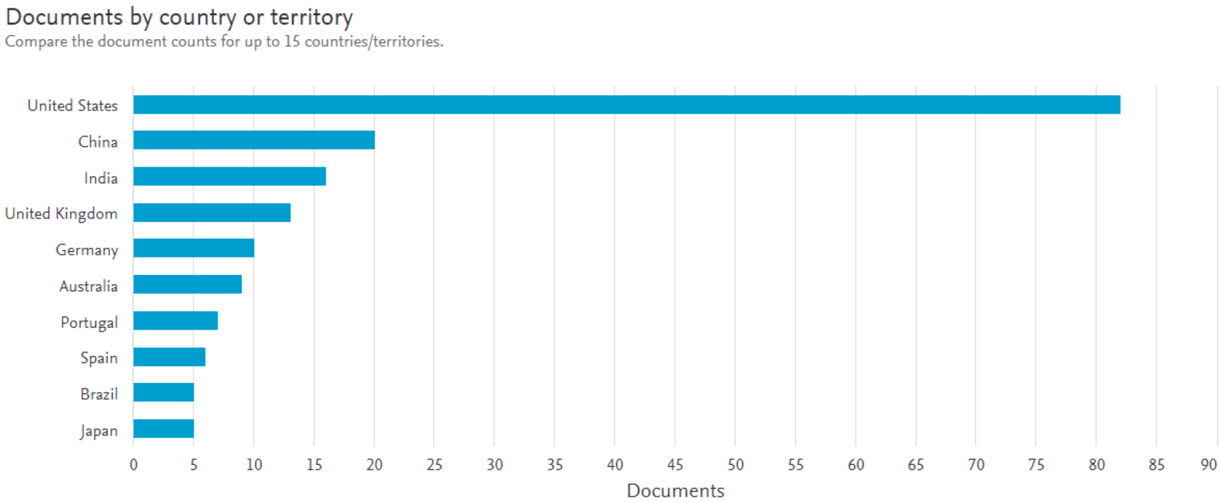


Figure 11 – Documents by country or territory based on the main string (Scopus, 2020a)

Reading the graph above, a clear overweight of literature stemming from America becomes visible. Approximately 46% of the research within the categories given is American. Reasons why could be many, but this trend continues in 2019-2020, meaning single explanations such as the American presidential election that spikes scientific interest in American falls short here.

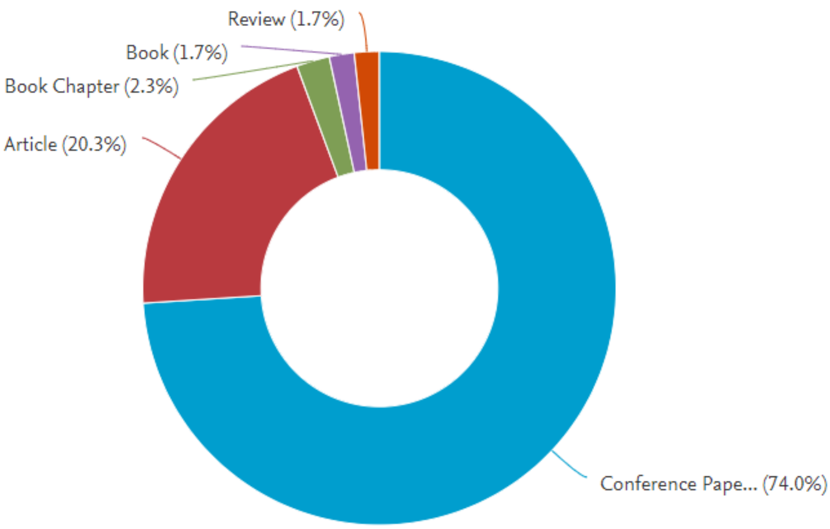


Figure 12 – Documents by type based on the main string (Scopus, 2020a)

The bar chart above highlights the maturity of the themes in questions (fake news in social media set in the context of a crisis); 74% of the literature stems from conference papers. If one were to exclusively use sources originate from Basket of Eight – which often is seen as a proof of quality (AIS, 2020), the literature is limited at the current time.

2.5.6 Concept Matrix

Following Webster and Watson guidelines in their article from 2002, the concept matrix should reflect the identified concepts in the collected literature (Webster & Watson, 2002). The first iteration of this concept matrix has its origin from the 2019 literature review conducted by the study group and has been updated to reflect the theme of the study. In its original form, it also covered concepts such as dissemination patterns (manual processes and automated processes) as well as selective presentation (echo chambers and filter bubbles). While these are of interest, they are outside the scope of the study. The concept matrix also highlights some of the challenges the current literature faces, with its mixtures of terminologies and concepts that sometimes overlap and sometimes not. The matrix has undergone several iterations, from a broad general overview to the more detailed overview illustrated on the next page.

The imagery attached in this text can be hard to read due to the size. Thus, the concepts are attached in the image below, followed by a short rundown of their intentions.

Methodology					Type		Fake news								Social Media Platforms				Fake news Detection				Disruption type		
Quantitative study	Qualitative study	Case study	Theoretical study	Literature review	Journal	Conferance paper	Fake news / fake stories	Conspiracy theories	Misinformation	False news / false information / Claims	Disinformation	Rumors / rumor cascade	Tweet / Retweet cascades	Hoax	Twitter	Facebook	LinkedIn	Reddit	Manual	Machine Learning	URL	Regular Expression	Crisis	Emergency	Disaster

Table 2 - Identified concepts

Methodology: Qualitative study, qualitative study, case study, theoretical study, and literature review. Acted more like a classifier for the group to aid the process while at the same time being of interest to understand the field better.
Type: Journal and conference paper.
Fake News: Fake News / fake stories, conspiracy theories, misinformation, false news / false information / false claims, disinformation, rumors/rumor cascade, tweet/retweet cascades, and hoax. As all articles in one way or another address false information, the phrasing and point of focus differ. To highlight this change, all identified concepts within the category were listed to see the grouping of concepts.
Social media platforms: Twitter, Facebook, LinkedIn, and Reddit. These to highlight the social media platform(s) studied.
Fake news detection: Different fake news detection methods. These have been identified in the literature, but there exist other detection methods. Manual, machine learning, URL, regular expression.
Disruption type: Crisis, emergency, and disaster. Linked to the theme crisis and used to show the different terminologies. See section 2.3 for more on why this is particularly important.

In appendix E, there are further details added to the concept matrix, such as the overlapping and occurrence of concepts.

			Methodology					Type		Fake news								Social Media Platforms				Fake news Detection				Disruption type		
Nr	Author(s)	Year	Quantitative study	Qualitative study	Case study	Theoretical study	Literature review	Journal	Conferance paper	Fake news / fake stories	Conspiracy theories	Misinformation	False news / false information / Claims	Disinformation	Rumors / rumor cascade	Tweet / Retweet cascades	Hoax	Twitter	Facebook	LinkedIn	Reddit	Manual	Machine Learning	URL	Regular Expression	Crisis	Emergency	Disaster
1	Del-Fresno-garcía M., Manfredi-Sánchez J.-L.	2018	X	X				X		X								X				X						
2	Ross B., Pilz L., Cabrera B., Brachten F., Neubaum G., Stieglitz S.	2019	X			X		X		X		X						X	X	X								
3	Burbach L., Halbach P., Ziefle M., Valdez A.C.	2019	X	X					X	X								X	X					X				
4	Budak C.	2019	X		X				X	X	X	X	X	X	X			X	X									
5	Al-Rawi A., Groshek J., Zhang L.	2019	X					X		X		X						X										
6	Bastos M.T., Mercea D.	2019		X	X			X		X		X			X	X		X										
7	Jang S.M., Geng T., Queenie Li J.-Y., Xia R., Huang C.-T., Kim H., Tang J.	2018	X					X		X			X					X				X						
8	Del Vicario M., Gaito S., Quattrociocchi W., Zignani M., Zollo F.	2018	X						X	X		X	X		X			X	X									
9	Campan A., Cuzzocrea A., Truta T.M.	2018					X		X	X								X	X									
10	Zimmer F., Scheibe K., Stock M., Stock W.G.	2019	X	X	X			X		X	X	X	X	X	X		X	X	X		X							
11	Pal A., Chua A.Y.K.	2019	X						X	X								X				X						
12	Babcock M., Cox R.A.V., Kumar S.	2019	X						X	X		X	X	X				X				X						
13	Zannettou S., Sirivianos M., Blackburn J., Kourtellis N.	2019					X	X		X	X	X	X	X	X		X	X	X		X							
14	Zannettou S., Caulfield T., De Cristofaro E., Kourtellis N., Leontiadis I., Sirivianos M., Stringhini G., Blackburn J.	2017	X						X	X	X	X	X	X	X		X	X	X		X			X				
15	Norah Abokhodair, Daisy Yoo, David W. McDonald	2015	X	X				X										X								X		
16	Soroush Vosoughi, Deb Roy, Sinan Aral	2018		X				X		X		X	X	X		X	X	X	X			X						X
17	Ross, B., Heisel, J., Jung, A.-K., Stieglitz, S.	2018	X					X		X		X	X	X			X		X									
18	Murungi, D., Purao, S., Yates, D.J.	2018		X	X				X	X		X	X					X	X			X						
19	Oh O., Kwon K.H., Rao H.R.	2010	X	X					X			X			X			X				X						X

Table 3 - Concept matrix of 2019 articles

2020 Articles																												
20	Kogan M., Palen L., Anderson K.M.	2015	X	X					X							X		X										X
21	Andrews C., Fichet E., Ding Y., Spiro E.S., Starbird K.	2016	X	X					X			X			X			X				X				X		
22	Nied A.C., Stewart L., Spiro E., Starbird K.	2017	X	X					X			X			X			X				X				X		
23	Li H., Dombrowski L., Brady E.	2018	X	X					X						X			X								X		
24	Aigner J., Durchardt A., Kersting H., Kattenbeck M., Elsweiler D.	2017	X						X				X					X				X				(X)		
25	Zannettou S., Sirivianos M., Caulfield T., Stringhini G., De Cristofaro E., Blackburn J.	2019	X		X				X				X	X				X										
26	Zannettou S., Caulfield T., Setzer W., Sirivianos M., Stringhini G., Blackburn J.	2019	X						X	X			X	X				X								X		
27	Bevensee E., Ross A.R.	2019	X						X					X				X										
28	Haug M.	2019	X						X	X		X	X					X										
29	Zakharchenko A., Maksimtsova Y., Iurchenko V., Shevchenko V., Fedushko S.	2019	X		X				X	X	X								X			X						
30	Ribeiro B., Gonçalves C., Pereira F., Pereira G., Santos J., Gonçalves R., Au-Yong-Oliveira M.	2019	X	X				X		X								X	X									
31	Koidl K., Matthews T.	2017	X						X						X			X							X	X		X
32	Jones, Marc Owen	2019	X		X			X		X								X								X		
33	Kostakos, Panos; Nykanen, Markus; Martinviita, Mikael; Pandya, Abhinay; Oussalah, Mourad	2018	X	X					X	X								X					X			X		
34	Maddock, Jim; Starbird, Kate; Al-Hassani, Haneen; Sandoval, Daniel E.; Orand, Mania; Mason, Robert M.	2015	X	X					X			X			X			X				X				X		
35	Rajdev, Meet; Lee, Kyumin	2015	X		X				X	X		X						X					X					X
36	Amanda L. Hughes; Leysia Palen	2009	X						X									X									X	
37	Firoj Alam; Ferda Ofli; Muhammad Imran; Michael Aupetit	2015	X						X									X										X
38	Robert Thomson; Naoya Ito; Hinako Suda; Fangyu Lin; Yafei Liu.; Ryo Hayasaka; Ryuzo Isochi; Zhou Wang	2012	X						X						X			X				X						X
39	Oh, Onook; Kwon, Kyounghee Hazel; and Rao, H. Ragha	2010	X						X						X			X				X						X
40	Oh, Onook; Agrawal, Manish; and Rao, Raghav.	2013	X					X							X			X				X				X		
41	Cheong, France and Cheong, Christopher	2011	X						X									X						X				X
42	Bovet, A., Makse, H.A.	2019	X					X		X		X	X					X										

Table 4 - Concept matrix of 2020 articles

2.6 Summary of literature review

Following the subparagraphs below, the identified concepts are elaborated in detail based on the findings of the articles included in the review. For a model with a summary of these findings, see section 2.6.4.

2.6.1 Fake News

Of the 42 articles that make out the review, all 42 mentions fake news or one of the synonyms (fake stories, conspiracy theories, misinformation, false news/fake information, fake claims, disinformation, rumor cascades, tweet cascades, retweet cascades, and hoaxes). It also affects how this chapter is structured, as instead of dividing all the concepts within fake news, it is instead combined within the fake news paragraph with key characteristics identified in the literature review highlighted.

Fake News classifications and definitions

It is argued that the term's ambiguousness is one of its main problems (Budak, 2019; Burbach et al., 2019; Campan, Cuzzocrea, & Truta, 2018; Del-Fresno-garcía & Manfredi-Sánchez, 2018). It is the context in which it is used that defines its meaning couplet with the actor in-question (Burbach et al., 2019). The attempts at classifying all of its forms have been many (Campan et al., 2018; Ferreira, 2018) - just to mention some. One of the more extensive classifications comes from Zannettou et al. (2019), which in its full form can be read in section 2.2 and summarized here as the following: fabricated, propaganda, conspiracy theories, hoaxes, biased or one sided, rumors, clickbait, and satire (Zannettou, Sirivianos, Blackburn, et al., 2019). Furthermore, several of these elements can be further labeled as either positive or negative charged, e.g., the elements can be used in both a positive setting while at other times its intention is solely malicious (Babcock, Cox, & Kumar, 2019; Zannettou, Sirivianos, Blackburn, et al., 2019). With such ambiguousness, concerns have been raised whether fake news is a valid terminology for the phenomena studied within the academic (Vosoughi et al., 2018). Depending on the actor and context, fake news is classified differently, which points out this issue with consistency (Zannettou, Sirivianos, Blackburn, et al., 2019).

Fake News properties

The properties of fake news are also an area of study that has gained attraction over the last years. Fake news relies on playing on the emotions of the reader, often through feelings such as fear, disgust, and surprise, in the opposite of "traditional media/news items" (our words)" attempts to play at feelings such as trust, sadness, happiness, and anticipation (Vosoughi et al., 2018). Fake news often tries to top traditional news items sensationally, and commonly attempts are made to make the target of the news item a person, theme, problem area, etc., false and the other way around (Pal & Chua, 2019). False in this context means to affect the public perception of the target/theme of the news item presented. Patterns indicate that truthful news aims to enlighten the public while fake news seeks to mislead. An interesting observation is that fake news seemingly spreads more than traditional news (Pal & Chua, 2019). More about spread later. Fake news often holds little burden of proof and more than often, the claims of the news item are far off from what the traditional society and research advocates (Michela Del Vicario et al., 2016). A prominent example of this is the anti-vaccine group. When this type of news and/or supporting people is met with criticism, the argumentation is often turned around and returned, which leads to stern fronts and polarization (Michela Del Vicario et al., 2016).

Fake News is often mentioned in the same context as the post-truth era (Del-Fresno-garcía & Manfredi-Sánchez, 2018; Zimmer, Scheibe, Stock, & Stock, 2019b). Some of the arguments used in this regard are that human intuition alone is no longer enough to verify whether the news is

true or false (Burbach et al., 2019). It further complicates the case that echo chambers, and filter bubbles make this process more difficult, as people are more likely to accept fake news as a piece of truthful news when it conforms to their world views (Bastos & Mercea, 2019; M. Del Vicario et al., 2018; M. Del Vicario, Quattrociocchi, Scala, & Zollo, 2019; A. S. Ross & Rivers, 2018; B. Ross, Heisel, Jung, & Stieglitz, 2018; Zimmer et al., 2019b).
While filter bubbles and echo chambers share many similarities, there are some key differences in how people end up in one (or both) of them, as the figure 13 below demonstrates.

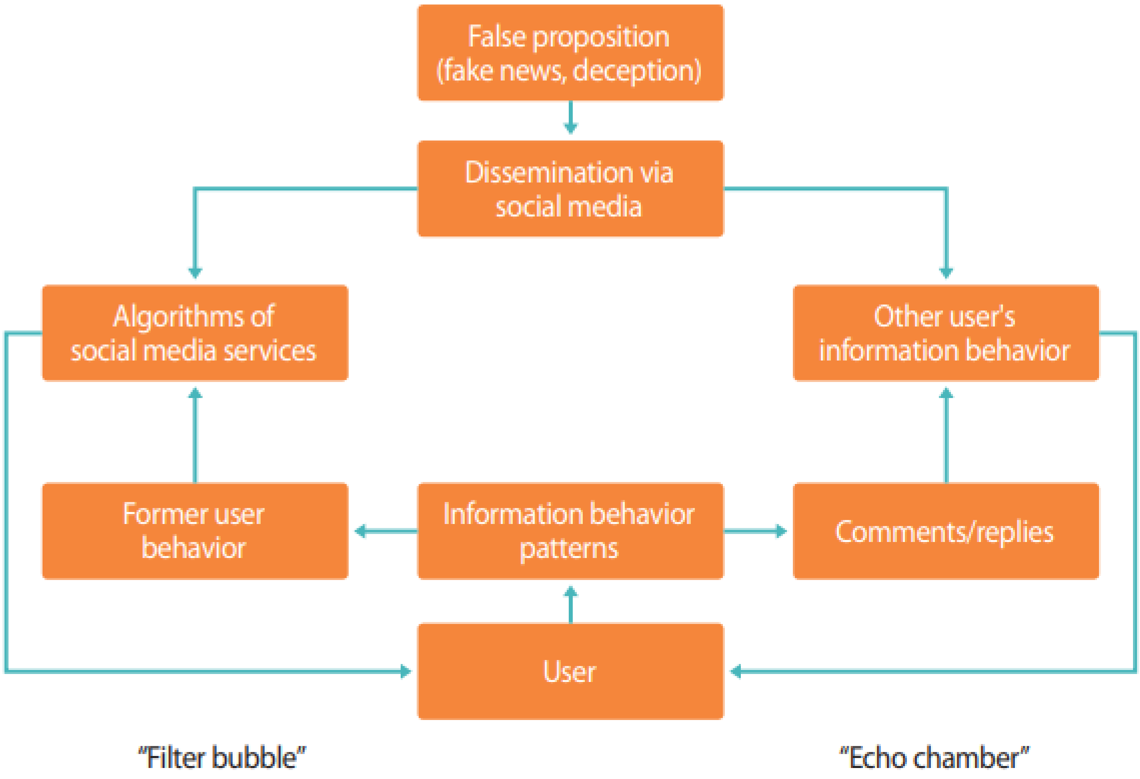


Figure 13- Filter bubbles and echo chambers (Zimmer, Scheibe, Stock, & Stock, 2019a, p. 43)

Filter bubbles, although triggered by the user's behavior, traps the user in a circle where algorithms present news that conforms to their interests and views. Echo chambers are the situation where people search for grouping (communities) of people sharing the interests and opinions of this person. While there is more to about echo chambers and filter bubbles, it falls more under the category of social media rather than fake news, as fake news is a byproduct of social media. Thus, there will not be further detailing the difference between the two in this text.
Related to this, understanding the difference between disinformation and misinformation is vital. Misinformation can be summarized as the act of believing that the piece of news is truthful, while disinformation is the act of spreading – by intention, a piece of fake news camouflaged as truthful (Zannettou, Sirivianos, Blackburn, et al., 2019). Often one sees that a piece of news content starts as disinformation, but as it is spread and accepted in different communities, it becomes misinformation (Zannettou, Sirivianos, Blackburn, et al., 2019).

Fake News diffusion

Previously in this thesis, it was argued that Twitter is the most prominent platform for the spread of fake news, something which case studies (related to the Catalonian election) also reflect (Del-Fresno-garcía & Manfredi-Sánchez, 2018). Studies from the American election shows similar signs (Budak, 2019). Furthermore, echo chambers and filter bubbles have a positive association with the dissemination of fake news (Michela Del Vicario et al., 2016).

Contradicting to what one might believe, studies have pointed on the fact that it is people – not automated processes such as bots, that spreads the majority of fake news (Burbach et al., 2019; Vosoughi et al., 2018). The same study from Budak shows that 1% of the total user mass accounted for 80% of the overall spread of fake news. Additionally, the origin of fake news seems to be primarily from traditional users of a social medium (Jang et al., 2018). Fake news appears to spread slower initially than truthful news, while over time has a longer life cycle and sees more alternation (Jang et al., 2018). Alternation here refers to different wording in the message, imagery used, author, and potential sources.
While the studies point in the direction of individual humans is the primary force of dissemination of fake news, one cannot underestimate the effect by bots and trolls. Following the American election, studies estimate that one-fifth of all the communication on Twitter related to the American election was produced by bots (Budak, 2019). One of the strengths of the automated processes such as bots is the ability to quickly establish networks of users that create rumor cascades to derail the public debate on a subject (Bastos & Mercea, 2019; Budak, 2019; Pal & Chua, 2019). A bi-effect of such cascades (often referred to as rumor cascades) is the origin of the news item being lost in this process, which can hide the actors with interest in spreading disinformation (Zannettou, Caulfield, et al., 2019; Zannettou, Sirivianos, Caulfield, et al., 2019).

Extensive amounts of work have been done on the classifications, properties, and spread of fake news. However, the empirical work is lacking, something which leads to gaps in the knowledge related to behavioral understanding around fake news (Al-Rawi et al., 2019).
We know how fake news can be created (Wilder & Vorobeychik, 2018; Zannettou, Caulfield, et al., 2019; Zannettou, Sirivianos, Caulfield, et al., 2019), we know the different types of it (Zannettou, Sirivianos, Blackburn, et al., 2019), we know how it spreads (Abokhodair, Yoo, & McDonald, 2015; Nied et al., 2017), yet – understanding the human mechanisms in-detail with credibility results is a challenging task (Lerbæk & Olsen, 2019b; Talwar et al., 2019).
The damage of fake news cannot be underestimated; it has the power to misguide the public perception, alter democratic processes and ultimately pose a threat to the stability of the society by setting up countries, people and cultures against each other's (Aigner et al., 2017; Budak, 2019; Davidson & Farquhar, 2020; Inglehart & Norris, 2016; Jang et al., 2018).

2.6.2 Social Media platforms

From the 42 articles that make the foundation of the literature review, various platforms (social mediums) have been in focus with Twitter dominates this by a large margin. Forty-one of the articles mention Twitter in one way or another in correlation to fake news. In comparison, the next most focused platform is Facebook, with 13 articles referring to it in one way or another. The reason why Twitter is such a platform of focus could be its own study, and it is not further explicitly discussed in the literature collected, other than being mentioned that it is often in focus (Al-Rawi et al., 2019; Bastos & Mercea, 2019). Assumptions can be formed that Twitter analysis tools offer a low entry-gate (skill-wise and effort-wise) to extract data from the platform, as almost all the articles that collected Twitter data used the same Twitter Streaming API. The Twitter API is also more open and accessible compared to other social media platforms, making it easier to collect and extract data for analysis.

2.6.3 Disruption type

As discussed in section 2.3, the terminologies are utilized differently across the literature; thus, this is presented "as-is" in the following part. This means that crisis, emergency, and disaster possible is used in a different context than defined in sections 2.3 and 2.4.

Disaster is a short- to a medium-time event with an immediate threat to human lives, dangers to infrastructure and challenges both the regional society alongside responders responsible for aiding the public (Alam, Ofli, Imran, & Aupetit, 2018; Oh, Kwon, & Rao, 2010). These types of events spur the public discussion and attract vast amounts of traffic in social media, and with the different intentions and actors, it also incubates fake news in one of its forms (Kogan et al., 2015; Rajdev & Lee, 2016).
The social media platforms have become the go-to arena for information. They can, in disaster events act as a disaster management platform to assess the damages and, at the same time, act as an information channel to the public (Alam et al., 2018).

Emergency shares several similarities to disaster, as it both operates within a compressed amount of time and imposes changes to routine life in a more or lesser degree (Hughes & Palen, 2009). As pointed out with mixing terminologies, an emergency is in the literature collected often similar to disaster, as the examples brought forth are wildfires and terrorist attacks (Hughes & Palen, 2009). As oppose to disaster, emergency also has a humanmade dimension to it (terrorist attacks), where a disaster has a solely natural cause (accident) trigger rather than being started by humans. Ultimately the literature collected covered this aspect in minimal degree.

Crisis is an event with a longer time horizon than those of disaster and emergency (Abokhodair et al., 2015; Andrews, Fichet, Ding, Spiro, & Starbird, 2016). The events within this category can be labeled as a natural crisis but are primarily used to describe humanmade (triggered) events (Aigner et al., 2017; Li et al., 2018; Nied et al., 2017). Often the crisis has roots from geopolitical actions and events, e.g., the crisis is a result of a previous event, or the event itself can be the crisis in-question (Jones, 2019; Zannettou, Sirivianos, Caulfield, et al., 2019).
As with disaster and emergency, social media plays a vital role by providing a platform to discuss the crisis at hand, while the messages that are being conveyed often emphasizes larger on political opinions and messages (Abokhodair et al., 2015; Li et al., 2018). The term, while being informative, can also be used as a label to fit a political narrative, like labeling *refugees* as a collective term under the *refugee crisis* in social media, to create a sense of urgency and crisis to the society (Aigner et al., 2017). While it is not unique to a crisis when speaking of disruption, a crisis is quite possibly the most receptive of the three in terms of the amount of fake news is disseminated related to an event within this category (Nied et al., 2017; Zannettou, Caulfield, et al., 2019).

2.6.4 Summary of the literature review table

To get a better understanding of the topic and to identify gaps in the existing literature, we performed a literature review. To summarize the literature, we created a model that can be seen in Table 5 below.

Concept	Findings	Literature
Fake News	Fake news can be defined as a deliberate presentation of false or misleading information. Although the term “fake news” is new, the act of spreading false information is not. The term “fake news” gained traction after the 2016 US presidential election. Before, the research literature used terms such as misinformation, disinformation, and rumors. The introduction of social media	(Babcock et al., 2019; Bastos & Mercea, 2019; Burbach et al., 2019; Michela Del Vicario et al., 2016; Del-Fresno-garcía & Manfredi-Sánchez, 2018; Pal & Chua, 2019; A. S. Ross & Rivers, 2018; Vosoughi et al., 2018; Zannettou, Sirivianos,

	platforms has made it easier for users to disseminate and consume fake news. On social media platforms, various actors with various motivations can disseminate fake news quickly to a broad base of users.	Blackburn, et al., 2019; Zimmer et al., 2019a)
Social Media Platforms	The use of social media platforms has drastically increased in the last few years. Social media provides a platform where users can share and receive information in a way that was previously not possible. This shift from a monolithic model to a decentralized model means users can choose what they want to read, see, and share. One of the most used social media platforms today is Twitter. Twitter is a microblogging service that lets users share short messages/tweets. These tweets can be shared on a public timeline for everyone to see.	(Al-Rawi et al., 2019; Bastos & Mercea, 2019; Heidemann et al., 2010; Nied et al., 2017; Zakharchenko et al., 2019)
Disruption Types	The three disruption types discussed are a disaster, emergency, and crisis. Disaster is a short-to-medium time event with an immediate threat to human lives and infrastructure. Similar to disaster, an emergency has a short-to-medium timeframe but also has a humanmade dimension (i.e., terrorist attack). A crisis has a longer timeframe than disasters and emergencies. A crisis can be natural or humanmade. During crises, social media can play a major role by providing a platform for information and knowledge sharing during a crisis	(Abokhodair et al., 2015; Alam et al., 2018; Andrews et al., 2016; Hughes & Palen, 2009; Jones, 2019; Kogan et al., 2015; Li et al., 2018; Nied et al., 2017; Oh et al., 2010; Rajdev & Lee, 2016; Zannettou, Sirivianos, Caulfield, et al., 2019)

Table 5- Summary of literature review

2.7 Limitations in the literature review

The literature review shows signs of immaturity in the literature, at least when looking at fake news on social media set in the context of crisis. Following the methodology provided by Kitchenham as well as Webster & Watson (Kitchenham, 2004; Kitchenham et al., 2009; Webster & Watson, 2002), couplet with elements suggested by Frank Danielsen (Danielsen, 2019), measurements have been made in an attempt to combat this. One such measurement is that we, in section 2.4, bridges fake news, social media, and crisis by arguing their connection as the literature is lacking.

Another of these measurements is a lengthy search string in combination with several filters and evaluation of the results this yielded. To highlight this, the Scopus stored searches list is attached below. Worth noting here that these are the final strings that were manually saved. A significant amount of time was used in the start solely on experimenting with what combination of strings and filters yielded the most precise results.

Query	Documents
TITLE-ABS-KEY ("Social Media" OR "Social Network") AND ("fake news" OR "false news" OR "false information" OR "false facts" OR "disinformation" OR "misinformation") AND ("refugee crisis" OR "refugee" OR "crisis") AND (LIMIT-TO (PUBSTAGE , "final")) ... View More	177
TITLE-ABS-KEY ("Social Media" OR "Social Network") AND ("fake news" OR "false news" OR "false facts" OR "disinformation" OR "misinformation") AND ("refugee crisis" OR "refugee" OR "crisis") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (SUBJA... View More	162
TITLE-ABS-KEY ("Social Media" OR "Social Network") AND ("fake news" OR "false news" OR "false facts" OR "disinformation" OR "misinformation") AND ("refugee crisis" OR "refugee" OR "crisis") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (SUBJA... View More	402
TITLE-ABS-KEY ("Social Media" OR "Social Network") AND ("fake news" OR "false news" OR "false facts" OR "disinformation" OR "misinformation") AND ("refugee crisis" OR "refugee") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (SUBJAREA , "CON... View More	23
TITLE-ABS-KEY ("Social Media" OR "Social Network") AND ("fake news" OR "false news" OR "false facts" OR "disinformation" OR "misinformation") AND ("refugee crisis" OR "refugee") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (SUBJAREA , "CON... View More	36
TITLE-ABS-KEY ("Social Media" OR "Social Network") AND ("fake news" OR "false news" OR "false facts" OR "disinformation" OR "misinformation") AND ("refugee crisis" OR "refugee" OR "crisis") CONF (icis OR ecis OR pacis OR mcis OR scis OR amcis OR... View More	18
TITLE-ABS-KEY ("Social Media" OR "Social Network") AND ("fake news" OR "false news" OR "false facts" OR "disinformation" OR "misinformation") AND ("refugee crisis" OR "refugee" OR "crisis") CONF (icis OR ecis OR pacis OR mcis OR scis OR amcis OR... View More	4

Table 6 - Search string evolution

Furthermore, the databases ISCRAM, Aisel, and Web of Science were utilized in combination with Scopus, to ensure the reach was sufficient. There also were five iterations of literature filtering through the usage of inclusion/exclusion criteria (see section 3.2 for more details). However, it is undoubtedly a mixed quality in the literature used, and there is a clear need to conduct more research on the subject area in the years to come.

Lastly, the process around peer-reviewing is a bit unclear, depending on the different databases. Although an active investigation was conducted on the various databases pages, they all operate differently in how they illustrate whether articles listed in the database were peer-reviewed. If articles utilized in this study, in turn, should show signs of lacking peer-reviewing in the literature used, this is then a result of misunderstanding rather than intention by the study group.

3 Method
This chapter presents the research approach, perspectives utilized, and measurements to address validity and reliability throughout this study. The method is grounded by the research questions and background literature on the subject, then fit-to-purpose (although following the methodology) to enable the group to be answering the research questions.

3.1 Philosophical perspective
As discussed within the introduction, this study has aimed to gain insights on the subject of fake news set in the context of the refugee crisis. By analyzing what is being said (what people do on the social medium Twitter) rather than asking people directly what they do through surveys and observations. As previous studies show, asking users their opinions and actions related to fake news gives inconclusive results (Lerbæk & Olsen, 2019b; Talwar et al., 2019), which suggest focusing on unmoderated raw data might yield more precise results. Unmoderated here refers to both self-moderated and moderated by the platform (Bastos & Mercea, 2019; Volkova & Bell, 2017).

The paradigm is set to a positivistic, as it is believed that events (fake news related to refugee crisis) can be observed empirically and explained with logical analysis (Kaboub, 2008, p. 343). It is worth noting that the group (self)identifies interpretivistic elements in this study. This stand is based on 1) the book by Oates (Oates, 2006), which points on that hypothesis testing is in the core of a positivistic paradigm, and that universal laws are current (such as gravity will lead to a certain outcome and is very observable) without the dependency of humans (Oates, 2006, p. 286). 2) As for interpretivism, it aims more on understanding the social constructs by studying people in their natural social settings (Oates, 2006, p. 293).
Points can be made that some elements of this study, to a certain degree, falls within the latter paradigm. Nevertheless, understanding complex social interactions and subjective realities miss the points of this study, even if one analyzes what people write on social media through document analysis. With the data we have studied, even if one does both quantitative and qualitative analyzes to it, it would be much too shallow to understand the social complexity in-details, as would be needed if following an interpretive paradigm. The last paradigm, critical (not dwelling into the sub-paradigm) is not current. This because it is by its very own definition, not the focus for the study, as its definition is quoted below.

Critical research in IS and computing is concerned with identifying power relations, conflicts and contradictions, and empowering people to eliminate them as source of alienation and domination (Oates, 2006, p. 296)

3.2 Research strategy
On an overall level, this chapter can be summarized as follows: the research strategy of choosing is a case study, data generated through documents, and the data analysis is done mixed. Below these choices will be discussed in detail.

Reasons for choosing case study is due to a confined timeframe of both the data collected (2016), the study period (January to end of May) and as well that the phenomena studied are of high complexity tied to a real-life context which suits case study as strategy well (Schell, 1992; Walshe, Caress, Chew-Graham, & Todd, 2004). Points can be made that what is mentioned above is too big to fit case studies, as they often hold key characteristics such as a focus on depth rather than breadth (Oates, 2006, p. 142). However, the setting is natural (it is not in a laboratory,

sterile setting), to some degree holistic (as it focuses on complexity rather than individual factors) and multiple sources and methods (Oates, 2006, p. 142).
Case studies come in three primary types: exploratory, descriptive, and explanatory. As discussed under paradigm, we argue that this study borrows elements from both exploratory and descriptive studies. While this study aims to help to understand a research problem through studying a real-life situation (fake news on Twitter in the context of the refugee crisis) which can help form questions and/or hypotheses for future studies, it also performs a detailed analysis of a particular phenomenon (Oates, 2006, p. 143). Ultimately, exploratory is the type of case study chosen in this study due to the immaturity of the literature and prior knowledge of the subject area.

The data generation method is Documents and defined as Research-generated documents rather than Found documents (Oates, 2006), as the data stems from previously collected data that were intended to (may)be analyzed in the future. The documents (CSV file) were evaluated to decide whether it held the needed quality and content and has been analyzed by the group before choosing to work with this data to ensure it adheres to the method defined by Oates (2006).
The data generation method was early in the study planning decided to be documents, of reasons mentioned under 3.1. Still, it was a process to determine whether to collect real-time data rather than using existing collected data. The group had assumptions beforehand the study that data related to the case needed to be recent, "fresh" if you like, as the user/system removes accounts and tweets over time. This assumption was later confirmed by the literature collected (Abokhodair et al., 2015; Bastos & Mercea, 2019). Using the same data collection method as what was done in 2016 today would lead to a lot of the data that makes out the current dataset to be missing, as it would have been deleted since the time of creation. This ultimately led to the choosing of the pre-existing dataset, even if it meant a lengthy application process and a range of ethical challenges. These are further discussed in-details under section 3.11 and limitations section of this thesis.

Analyzing the data is done by a mixed approach. As the data that makes out the study stems from both text and numbers with following quantitative analyzes to determine patterns and most mentioned phrases (to mention one of the analyzes, topic detection), this naturally falls within the quantifiable domain with further empathize on this is the positivist paradigm (although both interpretive and critical sometimes utilized this method as well (Oates, 2006, p. 245)). The different types of data within the dataset means there are different types of analyzes performed to return a measurable result. Counting occurrences of words are within nominal data, while ordinal data is data fields such as like and retweet count of a tweet, as illustrated in the image below (S = string, I = integer).

S Text	S Created At (UTC)	I Retwee...
	2016-07-12 04:03:49	7
	2016-07-12 04:04:57	93

Table 7 - Example of datatypes from the dataset

Several of the analyzes conducted do not operate with strictly numbers and counting of occurrences (although still possible (Oates, 2006, p. 266), instead they focus on working with language analysis and other text processing sub-analyses. These parts of the analyzes are more qualitative. Nevertheless, due to the sheer size, they fall within quantitative as the size meant we as a group could not delve deep into textual analyzing of each tweet to extract the exact meaning and interpretation or control check the analysis in detail.

These types of analyzes were done through KNIME Workflows (Berthold et al., 2009). This type of analysis is often referred to as computer-aided qualitative analysis (Oates, 2006), which we will elaborate on in-details on from section 3.8. While the group initially believed that Grounded Theory was the used data analysis method, Oates (2006) points out that what has actually been utilized within the analysis of an inductive approach, as categories identified stems from the research data collected – not the literature.

Some of the analyzes were done mainly qualitative. These were primarily the identification and classification of fake news within tweets and URLs, where the group relied on three fact-checking sources (PolitiFact, Snopes, FactCheck.org). This was then couplet by incorporate literature that previously had (if) covered relevant aspects of the item in question attempting to make sound argumentation as of why things were classified the way they were. The classification of fake news utilized a pre-existing framework to classify and was done entirely qualitative and manually (Zannettou, Sirivianos, Blackburn, et al., 2019).
This approach is further explained to some degree in the actual findings and discussion, as it is natural to discuss how these results were obtained; thus, the section is kept brief on the matter.

3.3 Validity and reliability

Validity and reliability are something that must addressed on both a paradigm level (general) and analyze level (specific) and is for this reason split into two subparagraphs and elaborated on what it means in the context of this study.

3.3.1 In paradigm, at an overall level

In positive paradigm one could say that validity is the overall quality of the study, e.g., it holds relevance in the topic being researched and what is presented is sound/valid, while reliability points to whether the results can be trusted and repeated in the future if a similar study were to be repeated (Jacobsen, 2016). As discussed under 3.1, to combat some of the noise (internal validity) that was visible in previous studies (Lerbæk & Olsen, 2019b; Talwar et al., 2019) and to fall more in-line with “The world exists independently of humans” (Oates, 2006, p. 286), the focus has been on what is written rather than what people say that they write on Twitter. There is no hypothesis testing (external validity) as the study is more exploratory in its shape and attempting to build up an understanding of human actions (it is inductive in nature), which later can be used to form new studies with more measurable variables. There also are some elements being borrowed from interpretivism, as it is a study of what people write, or act, in their natural environment (Oates, 2006). It is too shallow to be considered interpretivistic, yet worth taking into consideration as there are elements from its presence in the study. Validity and reliability concerning these elements covered by interpretivism are whether the study is successful in trustworthiness, confirmability, dependability, credibility and transferability throughout the study (Oates, 2006, pp. 294 - 295), in other words, a unique study with credible results that is possible to reproduce similarly in other studies by following the same method. As standardized tools and methods are used on a large dataset, with support of background literature to support the findings and in-detail followable steps throughout the study, this speaks to believe that the elements that in the study borrow some of its aspects from interpretivism as well as the overall positivistic paradigm, follow these ground pillars of validity and reliability.

3.3.2 In data collection and analysis

As discussed above, the validity relates to the quality of the study, and in this context, the quality of the collected data and whether it is of relevance to answering the research questions at hand. The sampling size, e.g., the number of collected tweets are of such a high quantity, couplet with

several synonyms to ensure collection of relevant discussions, means the selection that makes out the dataset holds satisfaction levels of validity. Also, as the data stems from an acknowledged research institution related to the RISE SMA project that has conducted similar projects and methods before (RISE_SMA, 2020).
The literature provided in this study shows signs of immaturity (Lerbæk & Olsen, 2019a; Scopus, 2020), meaning performing hypothesis testing would require a well-established understanding, both theoretical and empirical of the phenomena studied. Thus, this speaks for having a more exploratory approach where the data drives the analysis rather than hypothesis analysis to generalize the findings. The following subparagraph covers relatability.

The analyses were conducted by (mostly) in-built workflows in KNIME with a few exceptions – see section 3.4 and A.1.1 - A.1.8 for details regarding the software and explanation of workflow & nodes. In the exceptions, the workflows were tailored to support the analyzes conducted, with the inclusion of nodes that performed specific tasks that were needed to complete the analysis. All these custom nodes originate from the KNIME community and have been tested by other users before without (apparent) reports of misconfigured nodes. One such example is the use of Javasnippet nodes to extract URLs / mentioned links to other webpages from the text in the tweets collected(KNIME, 2020b). The strength of utilizing established workflows is that those particular workflows can be compared to peer-reviewed articles; they have gone through verification by the community to limit the shortcomings and errors in the workflow and are trusted to provide credible and repeatable results when data of sufficient quality as applied to it. Lastly, no workflow analyzes used online libraries to store data, e.g., network dependable, and was run in controlled local environments/instances with no external interference.

3.4 Software: KNIME Analytic

In this section, the details surrounding the software used in the study will be described. Note that not all aspects of the software are covered, though crucial components of it that are used in the analysis will receive extra attention where due.

KNIME – Konstanz information miner is a tool that is used in data analytics. It is designed as a teaching, research, and collaboration platform (Berthold et al., 2009). It is free, open-source software that is open for everyone to use. It was developed in 2006 by a team at the University of Konstanz, Germany. KNIME is used in several different industries, including financial services, retailers, manufacturing, government, and research (KNIME, 2020a). Under are some of the processes KNIME can be used for:

- (i) Extract, Transform, Load (ETL) processes
- (ii) Data cleaning (Preprocessing)
- (iii) Machine learning (ML)
- (iv) Deep learning
- (v) NLP – Natural Language processing
- (vi) Application Programming Interface (API) integration
- (vii) Data visualization

In KNIME *nodes* are the most general processing units (Berthold et al., 2009). They wrap all their functionality in a single unit or group. Nodes have several input and output instances. These inputs and outputs are how the data flows through the workflow, from node to node. A node usually has one single function/goal, e.g., filter, join, count, sort (Berthold et al., 2009). There exist a wide range of nodes in the library. Nevertheless, users can also modify and create personal nodes.

Workflows are a collection of nodes interconnected. Complex operations can be carried out by connecting nodes in a workflow. A workflow usually consists of several different parts, e.g., data reader, preprocessing, calculation, and visualization of results. A workflow can be altered by changing nodes (Berthold et al., 2009). KNIME workflows often start with a node that reads data from a data source. Data is stored in a table-based format in columns. KNIME supports several different data types (Integer, String, Image).

KNIME offers a vast library of nodes and a large amount of documentation. Users are actively creating new nodes and workflows for others to use. The modular pipelining concept (workflow) lets users use existing nodes or create their own based on their needs. This modular concept, along with the different modules, makes KNIME very suitable for our needs. Analyzing tweets require distinct functionality that KNIME can offer. With existing nodes as well as nodes you can modify/create yourself, KNIME provides much of the features this study needs.

KNIME was chosen due to both recommendations from the supervisors, functionality, and after some preliminary testing. The software is sophisticated with a high learning curve, yet, it offers the needed flexibility (to fit the purpose), support (through forums and information available), and is highly effective in both analyzing and visualizing the results.

KNIME has been used in several published research articles (Akhtar & Ahamad, 2017; Awrahman & Alatas, 2017; Baydogan & Alatas, 2018; Minanović, Gabelica, & Krstić, 2014). In these articles, KNIME has been used both as a support tool and as the primary tool for data collection and analysis. Minanović et al. looked at how text analytics and sentiment analytics could gain valuable business insights from social media streams like Twitter. Using KNIME, they performed data collection from Twitter, text analytics, and sentiment analytics (Minanović et al., 2014). Baydogan et al. also used KNIME on Twitter data. Ten thousand tweets were preprocessed, classified using sentiment machine learning algorithms, and the results accuracy analyzed and visualized (Baydogan & Alatas, 2018).

3.5 Data background

At the time the data was collected, the refugee crisis (see section 2, particularly 2.3) was a particular big topic of discussion on social media, which means the activity related to these tags / # / hashtags (same meaning) was rather extensive. Other events of interest within this year was Brexit (Bastos & Mercea, 2019; Inglehart & Norris, 2016), the American presidential election (Allcott & Gentzkow, 2017; Budak, 2019), amongst other events of interest (Lindsay, 2016). All in all, this most likely contributed to a very active year on Twitter based on the fact that people use social media increasingly more (Broadbandsearch.net, 2020; Statista, 2020).

3.6 Data collecting

The Twitter data was collected using the SMART tool – a Twitter data harvesting tool internally developed by the University of Duisburg-Essen. Duisburg-Essen developed the tool to collect Twitter data - as they are specialized in social media analytics. The University of Duisburg-Essen runs the tool to collect interesting events related to crises, political, or otherwise public events of interest. It is worth noting that due to the new GDPR regulations, the privacy concerns regarding this collection of data have both been addressed and approved by the University of Duisburg-Essen in advance of the collection (as well as the application related to this study for the usage of this data). The SMART tool utilizes the Twitter Streaming API to collect data from Twitter through an open developer platform/endpoint (Twitter, 2020c). The Twitter API lets people with sufficient technical competence search for tweets that contain specified search words, and possibly store them into files, databases, or through programs such as KNIME. The Twitter API makes it possible to collect publicly available tweets that meet certain predefined criteria, such as hashtags. Twitter users can choose to make their profiles and tweets private. These private tweets and profiles are not open to the public and are, therefore, not in the public timeline or dataset (Hughes & Palen, 2009). Tweets that contain the search word(s) can then be collected, and with necessary approvement be further analyzed (in our case NSD, with similar instances in other countries for this type of purposes).

The dataset used in this study was collected between January 1, 2016, and December 29, 2016. Twenty-two keywords were used to find relevant tweets, as seen in table 8 below. In addition to the keywords, a filter was applied in the data collecting (according to the information given the group) that made sure only to collect English tweets. The reason why the German wording keywords was because it was initially intended to be used to study the refugee crisis from the perspective of Germany.

Keywords			
asyl	flüchtlinge	migrantcrisis	refugeeswelcome
asylanten	freilassig	multikulti	schauhin
asylpolitik	grenzkontrollen	mundaufmachen	vielfalt
asylwahn	ichbineflüchtet	refugeecrisis	wehrdich
bereicherung	krimigranten	refugees	
deraustausch	marchofhope	refugeesnotwelcome	

Table 8 - Keywords used in data collecting

3.7 Dataset description

The raw dataset of tweets collected with the SMART tool has a size of 13.9GB before the structuring and cleaning. The file format is a Comma-separated values (CSV) file. It contains 14.3 million tweets posted by 2.5 million unique users. The dataset includes both tweets and retweets. 9.8 million retweets, 68.6% of total tweets, 4.5 million unique tweets, 31.4% of total tweets. Due to the large amount of data, the dataset can only be processed with the right tools. The tools must be able to handle large amounts of data efficiently and must be able to do several different computations. Due to an unknown error, the data for November and December is partly missing. This can be seen by the number of tweets posted in these months, which is significantly lower than in previous months. Using the Apache Tika library and Language Identifier class, we found that the dataset contained 41 different languages. English was the most dominant language with 98,8%, and German was the second most used language, with 0,5% of the tweets. It is worth noting that on short texts like tweets with a maximum of 140 characters, it is more challenging to perform language detection. If there are tweets with multiple languages, the Apache Tika library will return the most prominent language.

3.8 Data analyzing

As there are both quantitative and qualitative analyzes present in this study, the analysis at a general level is set to mixed. This due to the different analyzes requiring different data to conduct an analysis. Some analyzes were based on processing numbers (data) to provide a result, e.g., evidence (Oates, 2006), while others were conducted more qualitative by processing non-numeric data such words (Oates, 2006).

The quantitative data in the study was primarily nominal, as the data itself was not directly numbers that could be put on a scale to weigh against each other's to determine how much or how little one data point was compared to another data point. Instead, the focus was set on seeing the frequencies of the data within the dataset. This means analyzing the data for occurrences and provide tables based on the following results. An example of such can be viewed below on the Tweet Timeline figure 16, which stems from KNIME performing descriptive statistics on the number of tweets throughout 2016. While it primarily was nominal quantitative data in the dataset, there were also some ordinal data. Examples of this were *like*, and *retweet count*, which are datapoints that can illustrate the popularity of a tweet, and in itself can be viewed as / analyzed as nominal and ordinal. One could do analyzes to determine how many of the total tweets that have received a like or retweet to provide tables. At the same time, it is possible to sort the like or retweet number by itself to indicate the popularity of the particular tweets. The latter was the analysis form utilized in the study as it made more sense and provided less noise and chances for errors.

The qualitative data were a couple of text fields. One contained the raw text of what was written in the tweet while another held a location string. The latter was by many used as a humorous field, where they wrote irrelevant stuff such as *earth*, *universe*, *not home*, etc. Others had written their current location city. As for the text field, there were occurrences of performing quantitative analyzes by counting frequencies of tweets containing URLs (amongst others), but the primary analysis form was qualitative in nature. An example of this type of analysis was topic modeling. This analysis is further elaborated on in its paragraph below. As for location, this uses elements of both qualitative and quantitative, as it is analyzing words by comparing filter and grouping based on known cities in a document imported. Below a similar workflow of this analysis is attached to illustrate these steps (while not identical to those utilized by the group).

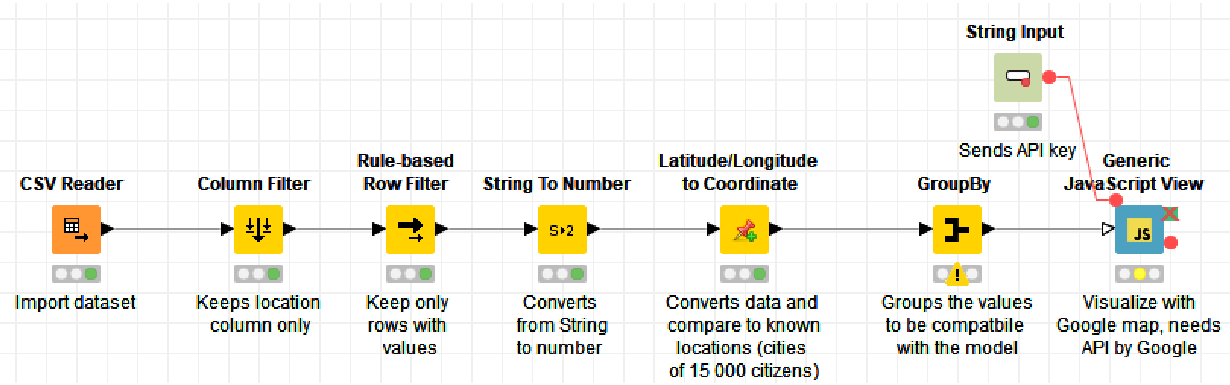


Figure 14 - Example of location workflow in KNIME

3.8.1 Preprocessing of the dataset

When collecting tweets through Twitter API, one can obtain a lot of data and a lot of noise. A lot of this information is very valuable, but there is also a lot of unwanted data that we do not need to analyze. If we do not preprocess the data or prepare it correctly, there is a chance that the analyzation and algorithms do not operate correctly or will report errors (García, Luengo, & Herrera, 2015). To remove the unwanted data, we do a process called preprocessing. Preprocessing can contain several different tasks, dependent on the type of analysis to perform. The aim of preprocessing is to remove and transform unwanted data and text from the raw dataset (Rajput, Haider, & Ghani, 2016). Raw data can also be unstructured and can make the analysis process difficult and resource-intensive. In the end, treated data that is clean, accurate, and scaled and is ready to be analyzed. In the preprocessing, the hashtags were not removed. Users often use hashtags as a replacement for a common word (e.g. "The current #migrantcrisis must be handeled"). Removing the hashtag could possibly take away key words in a sentence. Therefore losing the sentence essense and message.

(i) *Stemming* - Reducing a word to its root/base word, by cutting off the end or beginning. This is done with a list of common suffixes and prefixes.
(ii) *Case conversion* – Convert text to the same letter case. This can change the text into upper or lowercase.
(iii) *Punctuation Removal* – Remove punctuation characters from the text. Punctuation characters often do not give any useful information when analyzing text.
(iv) *POS tagging* - *Part of Speech tagging* – Each word in a text is assigned an appropriate part of speech. This label can be verbs, nouns, adverbs, adjectives, and more.
(v) *Lemmatization* – Reduce a word to its root/base word. This is done with a dictionary.
(vi) *Number filter* – Filter all numerical terms, numbers, and operators.
(vii) *Stop word filter* – Remove words that do not give much meaning to a text. E.g., the, a, in, an.

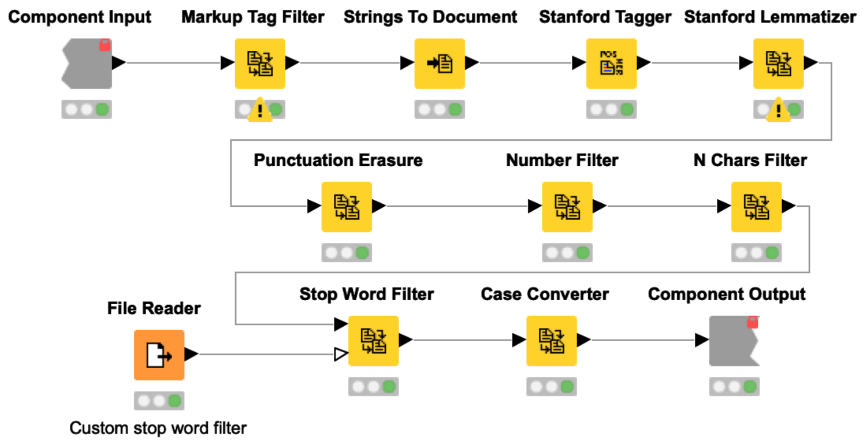


Figure 15 - Data Preprocessing

3.8.2 Coding of the dataset

As part of the agreement with NSD and UiA, several actions were performed to ensure that privacy was safeguarded, which meant the dataset underwent extensive coding. The ethical issues raised by this can be read in detail in the ethical paragraph below, while this part focuses on covering the technicalities related to coding. Six steps were performed.

1) The *authorID* column was replaced with a new authorID column containing a hashed representation of the original authorID. This action was performed through a workflow in KNIME, as illustrated below. The new authorID consists of strings with randomly generated 11-characters words.
2) Most of the analyzes presented do not operate with the text in a readable format. Instead, it is presented through topic detection, numbering, claims raised, etc., meaning the original sentences in most examples are not present (except for most liked and retweeted tweets).
3) *Longitude* and *latitude* columns deleted. In the dataset, there were about 10 000 users with such information available.

4) Deletion of *author name* and *author screen name*, only keeping the newly hashed authorID
5) Deletion of *author description*, as it is both was a concern privacy-wise, but also held little value to the study and its focus.
6) *Original author name / original screen name / original ID / original description / original longitude / latitude* removed. This to adhere to point 1-5, with the secondary focus of reducing the size of the dataset and reduce analyzing time (which could take up to three days for an analysis). Original was separate columns, and not referred to those of step 1-5.

3.8.3 Descriptive statistics

To get a better understanding of what the dataset included and how the data was distributed, we performed some descriptive statistics. Having such a large amount of data makes it hard to understand the data without descriptive statistics and visualization. To better understand how many tweets were posted within a timeframe, we created a timeline. See figure 16. The timeline shows the total number of public tweets posted with any of the keywords per week. The average number of tweets, without November and December, is about 1.4 million each month. The group and the supervisors contacted the University of Duisburg-Essen to investigate the reason for the sudden drop of entries in November and December, but it was unclear to those involved why this was the case. An assumption here could be bugs or errors in the collection or due to the massive American political focus caught in the dataset - or a combination (see findings). Although the numbers for these two months were lower than in previous months, these have been included in the analysis of the dataset.

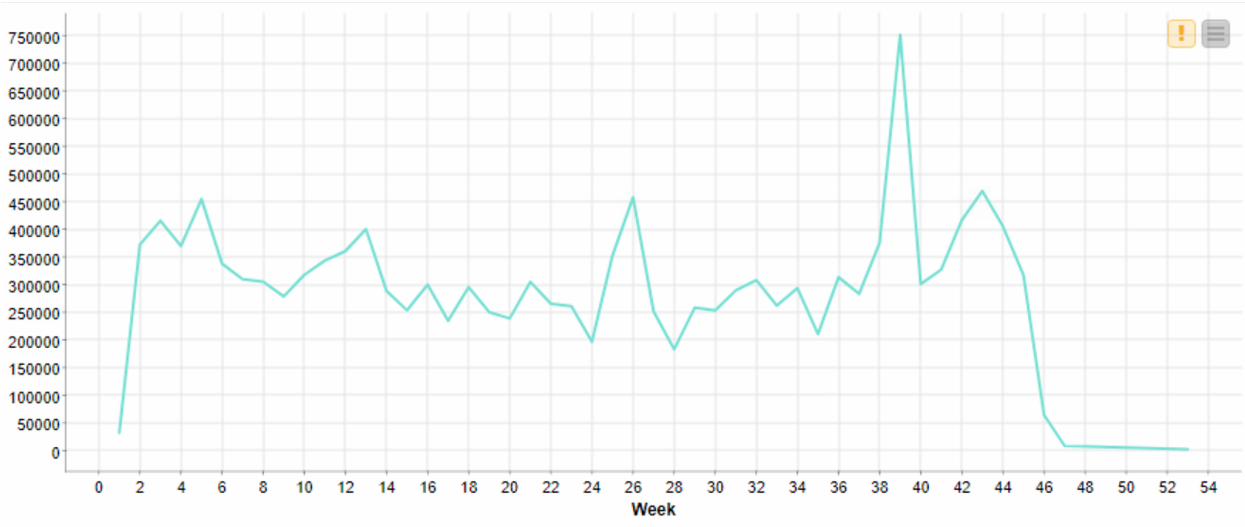


Figure 16 - Timeline over tweets per week

Topic modeling with hashtags

A dataset with 14.3 million tweets is quite large. It would be challenging and time-consuming to go through each tweet manually to perform the various analyses. The process of finding the main topics in a text is called topic modeling. “Topic modeling can be defined as an ability to capture different core ideas or themes in various documents”(Reese, Reese, Kaluza, Kamath, & Choppella, 2017). To find the most discussed topics in the dataset, we utilized LDA, Latent Dirichlet Allocation algorithm. LDA, which is a well-known topic modeling technique for doing NLP, natural language processing The LDA algorithm is unsupervised, which means the algorithm will look at a text and find topics by itself, without any instructions. The LDA algorithm looks at a given text and finds semantically related words and group them together. Topic modeling helps us understand and summarize extensive collections of textual information. With topic modeling, we aim to discover hidden topical patterns that are present across different times of the dataset (Alam et al., 2018). Below is an example of experiments with the software and dataset to illustrate how this topic modeling can be done if one sets the tool to focus on hashtags only to define the topics.

Figure 17 - Initially hashtag word cloud

Rank	Hashtag	Frequency	Rank	Hashtag	Frequency
1	Syria	142917	11	news	30455
2	migrantcrisis	97811	12	UNHCR	25468
3	EU	92703	13	UNGA	23583
4	tcot	88471	14	cdnpoli	22059
5	auspol	77951	15	refugeesGr	20303
6	Syrian	67326	16	rapefugees	18528
7	migrants	66668	17	UN	18177
8	MAGA	55696	18	pjnet	17402
9	SyrianRefugees	46291	19	immigration	17278
10	UN4RefugeesMigrants	43220	20	ccot	16726

Table 9 - Most used hashtags on a portion of the dataset (search keywords from data collecting excluded)

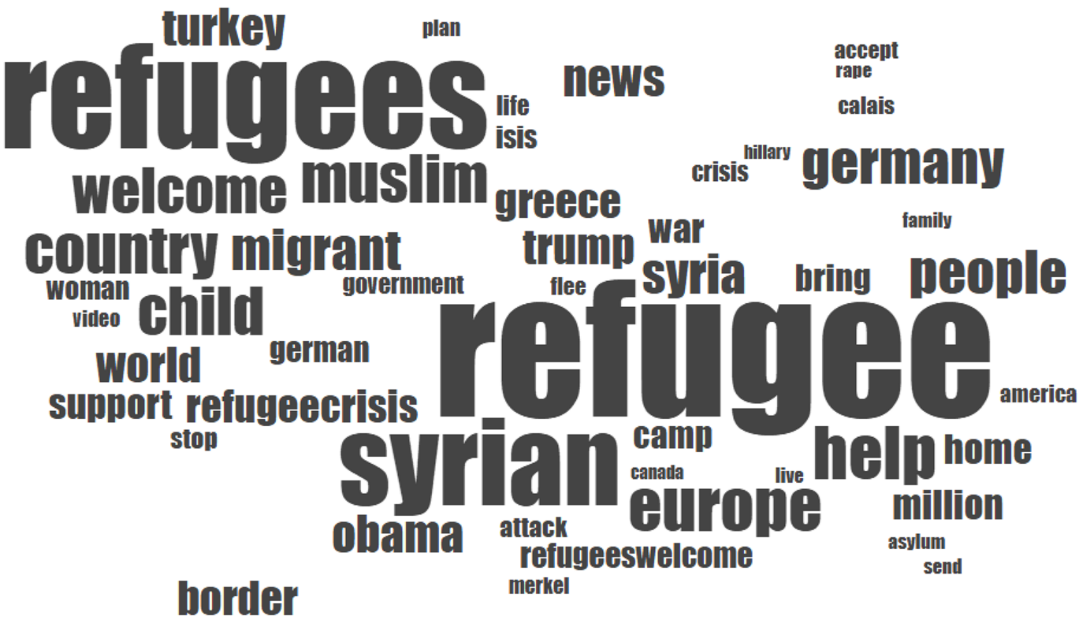


Figure 18 - Topic detection on a portion of the dataset

Rank	Word	Frequency	Rank	Word	Frequency
1	refugee	6064152	11	germany	325920
2	refugees	2198178	12	migrant	297045
3	syrian	1220961	13	syria	286392
4	help	444919	14	news	263976
5	europe	413857	15	obama	258104
6	muslim	372839	16	turkey	254476
7	child	368813	17	world	248321
8	country	368478	18	trump	248160
9	welcome	334281	19	border	239951
10	people	328748	20	greece	236612

Table 10 - Most used words in tweets on a portion of the dataset

Topic modeling with full text

The central part of this study has been to gain a preliminary understanding of some of the topics that make out the public discussion caught in the dataset related to the refugee crisis and try to determine whether these topics are prone to attract fake news in one way or another. A first step was to identify the most prominent issues and to accomplish this, KNIME was used with the customized workflow fit-to purpose. The workflow (figure 19) was initially a workflow created to summarize books (Thiel, 2014). It is worth noting that this original workflow has found its place into the standard library of KNIME, something which suggests it holds certain credibility. As both source material and goal were a bit different, this workflow was altered into the following setup. Note that figure 20 relates to the grey "pre-processing" node and is the actual steps performed within.

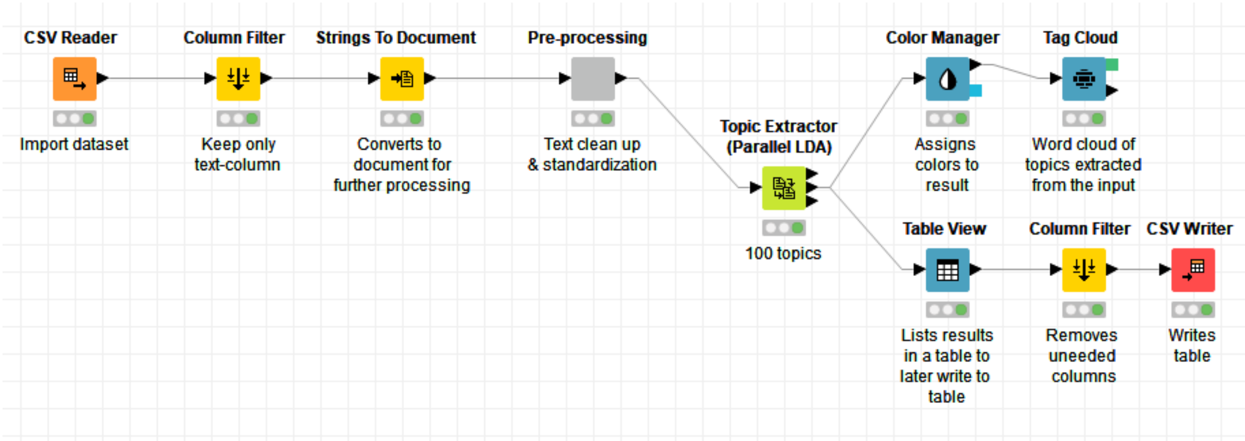


Figure 19 - Customized workflow to perform topic detection

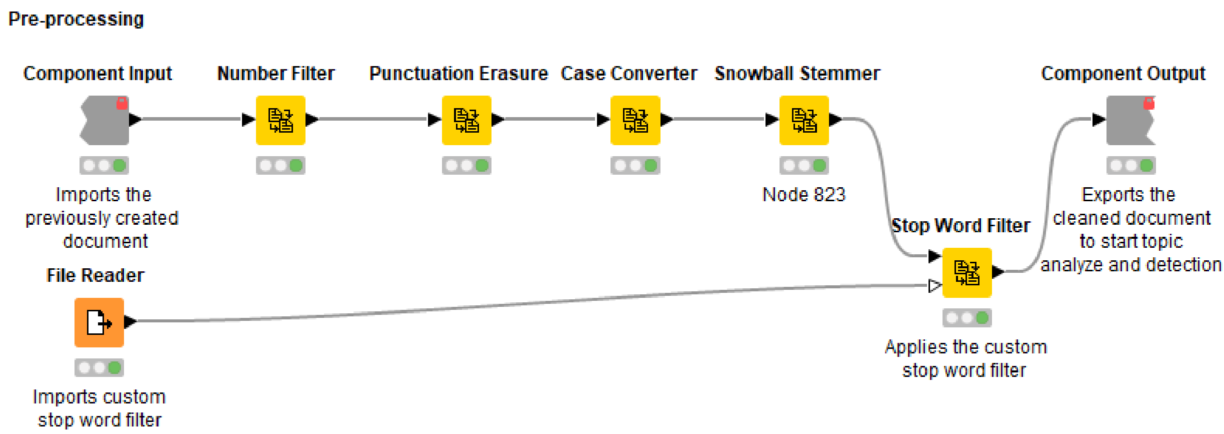


Figure 20 - Pre-processing node in-details

The key in this workflow is the topic extraction (parallel LDA) node. The node consists of algorithms and methods based on literature related to topic extraction (Newman, Asuncion, Smyth, & Welling, 2009; Yao, Mimno, & McCallum, 2009). Furthermore, the node utilizes a topic modeling library, which stems from machine learning provided by the University of Massachusetts Amherst (Mallet, 2018). For the study, the limit was set to 10 topics with 10 terms. These settings are in line with previous research methodology in similar studies (Zannettou, Caulfield, et al., 2019; Zannettou, Sirivianos, Caulfield, et al., 2019).

3.9 Compiling list of websites for fake news detection

60,1% of the tweets in the dataset have at least one URL. These URLs point to different websites. These websites could contain information that can be considered fake news. To be able to find and detect these websites, we compiled a list of URLs that point to a site classified to be producing fake news stories. There exist several different lists of websites that are classified as fake news sites. To be able to detect as many URLs as possible and, at the same time, reduce sample selection bias and ensure correct classification of the websites, we used four different lists. These lists come from different sources with different selection criteria. Most of the websites in these lists overlapped each other, indicating that the lists have undergone a robust verification.

i) Paper by Grinberg et al. (2019) – 490 URLs
ii) Paper by Allcott et al. (2019) – 672 URLs

iii) List by Politifact (2017) – 327 URLs
iv) List by OpenSources (2017) – 1001 URLs

The paper by Grinberg et al. 2019, (Grinberg, Joseph, Friedland, Swire-Thompson, & Lazer, 2019) contains a total of 490 URLs. The URLs are classified into three groups, black (entirely fabricated stories), red (extremely high), and orange (high) based on the likelihood to publish misinformation.
The paper by Allcott et al. 2019 (Allcott, Gentzkow, & Yu, 2019) contains 672 URLs. The authors compiled the list from five different sources, two academic papers, and three from fact-checking sites like Politifact and FackCheck.org. The list by Politifact (Politifact, 2017) contains 327 unique URLs. The list classifies websites into fake news, some fake stories, or imposter. The list has been created in collaboration with Facebook. The list by OpenSources (OpenSources, 2017) contains 1001 unique URLs. The list classifies websites like fake news, extreme bias, conspiracy, and rumor mill. The list is created for assessing online information sources and is available for public use. The research team behind the list is maintained by Melissa Zimdars of Merrimack College. The team utilized a six-step process for analyzing potential fake news websites, 1) title/domain analysis, 2) about us analysis, 3) source analysis, 4) writing style analysis, 5) aesthetic analysis and 6) social media analysis.

After compiling the four different lists together, for the purpose of this study, we generated a new list with 1387 unique URLs. The newly compiled list is not by any means a complete list of websites that produce and share fake news. There are many websites that are not included, and that are hard to find. It is also worth noting that not every news item or post on these websites are fake. Some of the things on the websites classified as fake may be clickbait, rumors, hoaxes, satire (lesser forms of fake news), general political, or even true. Even though this is the case, we include them and classify them as fake news due to the broader framework provided by Zannettou et al. (2019).

3.10 Fake news detection

In this study, it has been used both automated detection of fake news as well as manual detection. The automated detection was done by utilizing preexisting lists of websites (URLs) provided by different sources compiled together. Previous research has also used the URL method to identify and extract fake news (Bovet & Makse, 2019; Budak, 2019; Zannettou et al., 2017). Bovet & Makse looked fake news on Twitter during the 2016 U.S presidential election. To identify fake news, they used the list of classified websites by OpenSources and matched them with URLs extracted from tweets (Bovet & Makse, 2019). Budak also looked at Twitter during the 2016 U.S presidential election but focused on the prevalence of both fake news and traditional news mentioning either Clinton or Trump. This paper used the same method as Bovet & Makse by identifying fake news at the domain level (Budak, 2019).

As there is a lack of if research performing manual fake news detection through utilizing Expert-based Fact-checking Websites (Zhou & Zafarani, 2018, pp. 7-8), the method of choice has been manually fact-checking. Through the webpages and services listed at Social Cybersecurity Working Group (SCWG, 2020), most of the manually fake news detection has been done through webpages and services listed here. Throughout the fake news detection process, we have provided source citation to the service utilized. In those instances where alternative manually methods not reliable on the webpages listed at Social Cybersecurity Working Group (SCWG, (Group, 2020)) have been utilized, these have been explained couplet with why these are deemed acceptable for the purpose.

Tweets are quite short messages. In 2016 (until November 2017), the year the dataset was collected, the limit for one tweet was 140 characters (Twitter, 2016). Many users, therefore, use links (also referred to as Uniform Resource Identifier, or URL short) to other websites to back up their claims or share a news story or other media. The challenge of retrieving URLs from the dataset was because all URLs shared on Twitter is encoded into smaller versions of the original, to both reduce the length of URL in the message as well as provide statistical data to Twitter (Twitter, 2020a). Through our workflow in KNIME, we were able to retrieve the original URLs. The workflow, in its entirety, can be viewed under section A.1.7. An abstract level of this process can be viewed in figure 21, focusing on the first four “boxes” of activity that extracts the complete URL.
Creating this workflow in KNIME was hugely beneficial as this would then make it possible to check whether the URLs provided in tweets pointed to webpages containing fake news. The original dataset contained 14.3 million tweets. 8.6 million of them contained at least one URL. With the workflow, we were able to analyze the complete dataset and identify every URL.

The workflow starts by importing the dataset. The dataset is then cleaned and filtered to make sure the data is ready to be analyzed. Filtering the data keeps the rows we are interested in analyzing, the tweets itself, and removes everything we do not need. After filtering, we use regular expressions to identify and extract URLs. Many of the URLs are shortened, e.g., t.co, shorturl, TinyURL, to keep the tweet under the maximum number of characters. We take the shortened URL and expand it to the original full-length URL and extract the hostname using Java URL class. After we have fetched the complete URL, we look at each URL in the dataset and loop through the list of classified fake news sites. If there is a match, we mark the tweet as fake news. After the loop is made, we have a list of tweets that are classified as fake news. The list was then sorted, filtered, and each URL was counted by the number of occurrences.

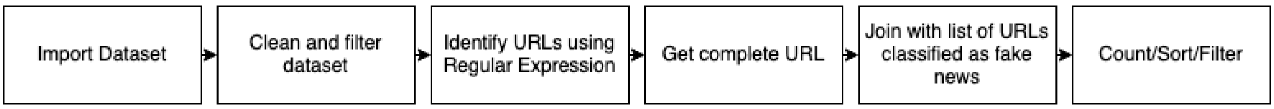


Figure 21 - Fake news detection - URL method

3.11 Ethical challenges

Throughout the study, different ethical challenges have arisen that needed actions from both the study group, the University of Agder (UiA), and Norsk Senter For Forskningsdata (NSD). To better illustrate these challenges, what they relate to, and what was done to adhere to them, these have been split into the pre-study period and within study period paragraphs.

3.11.1 Ethical challenges, pre-study period

While a master thesis, project or entity that process personal data in one way or another in the context of research for the most parts over the years have had to deal with strict regulations, the new General Data Protection Regulation (GDPR) further enhance these regulations (NSD, 2020b). As the data utilized in the study were collected before GDPR implementation (collected in 2016, GDPR fully implemented in 2018 (European Commission, 2020)), this raised concern for NSD as using the data meant to use data that had not been cleared with the new GDPR regulations in mind. Even when the University of Duisburg-Essen has strict self-regulations and addressed privacy concerns before collecting data that has been cleared with the European Commission, the new regulation means that when the context changes, e.g., a new study analyzing the data, in this case, new applications must be filled. This meant that the group had to both apply with a typical application to NSD, as well as a Data Protection Impact Assessment (DPIA) document (NSD, 2020a). One of the main factors that lead to the requirements of this application other than the changed context of data was that political opinions are listed as sensitive privacy information (regardless if one could say that most Twitter data relating refugee crisis is political opinions). This comes in play even if the usernames and identifiers are removed/coded from the dataset, and users would be practically impossible to identify. Furthermore, as the new GDPR regulations set guides stating that those involved in the study (approx. 2.5M users) should have the right to remove their consent (and thus data), the need to argue why this was not the case in this study had to be done in great details. Ultimately NSD concluded that as the number of participants was over 100 000, it would be impossible for the group to ask every participant directly whether they wished the data to be deleted. Additionally, the data would be coded, which removed all traces of the origin of the tweet and mitigated the privacy concerns. Lastly, the data collected was essential for the completion of the study, coupled with the benefit such a study held both on a societal and research level led to the study being approved between NSD. When NSD had given their approval, the study details were sent to the Data Protection Official at UiA, as UiA, in the end, is responsible for the data in the study. This process took quite the time, and the final approval by UiA was given in March after further details had to be cleared in terms of storing and processing of data. Latter involved working on machines with 3-factor login: restrictive access to files through both password protection on the drive as well as mobile phone authentication, in combination with password protection of the files in question (on opening).

3.11.2 Ethical challenges, within the study period

Firstly, the initial datafile contained datapoints that in themselves did not pose a significant threat in identifying their origin (person posting the information). Still, the main problem starts when one groups them and makes “sense” of the data. One example where the grouping of data was abused quickly comes to the mind; the Cambridge Analytica Case, where personal user behavior patterns were being sold for money (Isaak & Hanna, 2018). Although the case is not directly comparable, it shares many similarities relating to ethical challenges when structuring and using data.
To combat this, the data was cleaned and coded, by removing any references to usernames or general identifiable data (such as username, screen name, user ID, location), to ensure that the focus is what is being said on Twitter. The scope of the study is to cover what is said and establish whether the content is truthful or falls within fake news. Who says what in this context is not relevant and thus provides no direct weakness to the study by removing this information.

Secondly, how the data is being linked and presents is a big point of concern. Linked here refers to how the data is combined and structured to give meaning, e.g., user X lives in location Y and has posted content Z on social media.
With this vast amount of data, the group could focus on whatever aspect of interest. Not all findings are of interest; thus, the grouping of the data to fit a narrative or case was very much up to the group to decide. This meant an extensive amount of time was used in the early parts of the study to determine what the focus should be, how to present it objectively, and to be self-reflective in our choices and process.

Thirdly, processing, and storing of data is one of the core concerns. As agreed between the authors, UiA, and NSD, all the data was to be stored at secure data locations that only the two main group members had access to (as well as supervisors). The platform of choosing was a shared Office 365 – One Drive location with strict user access - only the authors and the supervisors could access this.

Fourthly, Non-Disclosure Agreement. Even if the data is stored at a secure location and coded, the group inhabits an extensive amount of knowledge and information on the subject. Being self-aware on the matter in meeting with the public is essential, and something the group both have discussed and at best of our abilities adhered to under and after the study period.

Lastly, adhering to the application. Throughout the study, the study by nature iterates and sees subtle changes. While the application to NSD and UiA covers some flexibility in what types of analyses are to be performed, there are limits to this flexibility. Where one starts and where one ends are different things, thus there was a process of self-reflect throughout the study and continuous evaluation with the supervisors.

4 Findings

In this chapter, the focus is to present the findings stemming from the data analyzes. The results are based on approach and workflows utilizing as described within the method section, 3.8 and onwards. For readability, this is divided into subparagraphs with a focus on tables and figures. As the analyzes are data-driven, the various findings are presented and later discussed under the discussion section. As will be discussed further within the discussion, the structure deviates a bit from the traditional format, wherein some situations such as fake news detection need a certain discussion related to it to be able to classify the tweets accordingly to the framework used.

4.1 Contextual findings

As this study focus on analyzing a previously not studied dataset, there is a need to understand more of the data caught in it. To do so, we focused on a handful of contextual findings. Some of these findings, such as most liked messages and most shared URLs, are the basis for fake news detection and classification, which is the main focus of both the analyzes and discussion. Also, the contextual findings aim to address RQ1 and partial RQ2. Knowing contextual information relating to the tweets helps to understand the characteristics of them, as well that it helps to find potential traces of fake news in or surrounding the tweets.

4.1.1 Locations of activity

Understanding where the traffic in the tweet material stems from is beneficial to gain a preliminary understanding of the dataset at hand as well as who the most vocal groups of actors are. As seen in previous studies, actors such as trolls / state-sponsored trolls, bots and various subgroups of people with different goals and motivations partake in the public debate, which poses challenges in terms of the spread of fake news (B. Ross et al., 2019; K. Starbird, Dailey, Mohamed, Lee, & Spiro, 2018; Zannettou, Caulfield, et al., 2019; Zannettou, Sirivianos, Caulfield, et al., 2019). These were the results of the 50 most self-mentioned (users must add a location on the Twitter profile) locations of the users in the dataset. These findings will be discussed under discussion but note that by intention, we have kept the values as given by the users and only removed those with empty values or non-valid characters (in UTF-8 character standard), hence the similar names.

Rank	Location	Count	Rank	Location	Count
1	United States	266428	26	New Jersey, USA	22274
2	London	143522	27	Toronto, Ontario	21698
3	USA	141805	28	World	21161
4	London, England	112608	29	Florida	20953
5	Australia	96267	30	Toronto	20910
6	Washington, DC	86396	31	Sydney	20647
7	United Kingdom	69580	32	Germany	20281
8	UK	68694	33	Chicago, IL	20092
9	Canada	65386	34	Scotland	19728
10	Florida, USA	55912	35	Pennsylvania, USA	19562
11	New York, NY	53814	36	Melbourne, Australia	19351
12	England, United Kingdom	52400	37	Sydney, Australia	19306
13	Texas, USA	50755	38	Brussels, Belgium	18925
14	California, USA	50363	39	Melbourne, Victoria	18876
15	Europe	42808	40	Brussels	18829
16	New York, USA	36423	41	Everywhere	18584
17	New York	34010	42	Nigeria	18197
18	Texas	33832	43	Nairobi, Kenya	18112
19	Worldwide	33571	44	Melbourne	17700
20	Los Angeles, CA	32814	45	North Carolina, USA	16951
21	London, UK	29528	46	Georgia, USA	16750
22	Earth	27782	47	California	16657
23	India	27343	48	Ireland	16595
24	England	24845	49	Paris, France	16580
25	Global	22647	50	Planet Earth	16387

Table 11 - top 50 location of Twitter users

Glancing at these results, there seems to be a clear trend towards an overweight of American users. It also highlights the policy of Twitter regarding this datapoint: Users must add a location to their profile, but the location can be anything they want as it is a free text. *Worldwide*, *Global*, *World*, *Everywhere*, and *Planet Earth* has a substantial number of entries when looking at the whole.

Furthermore, to highlight the reach of the dataset and visualize the entries in it, figure 22 illustrates this by having a square dot on the map for each of the entries (locations with several entries is represented with one square only) in the dataset where the location is present. Also, the results in the table 11 are checked against a world city list (cities in the world of above 15 000 citizens), and if match - plot 1 square on the map.

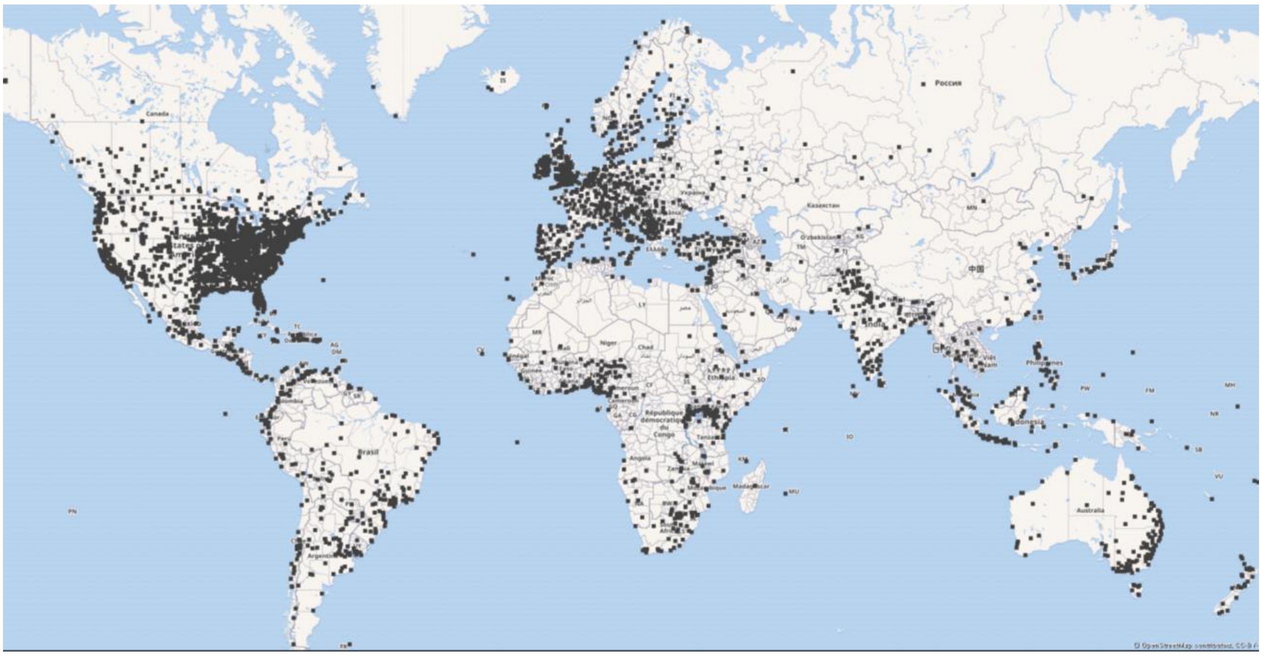


Figure 22- locations worldwide, one square for each location

There are a total of 508 250 different locations added. The top 1000 locations have an average of 5036 users tied to it, and the dataset contains 5 035 989 total tweets where users have added a location. Of the 5 035 989 total tweets with a location, 2 380 726 matched with the list of cities and is the foundation of figure 22.

4.1.2 Most used Hashtags in tweets

A hashtag on Twitter is a word with the prefix "#" in it, e.g., #fake news. This tag is used to categorize information on Twitter (also used on services like Instagram as well, to serve a similar purpose). It can act as a guidance to partake in the public debate on a particular topic due to the way the platform handles these tags. It is possible to search by these tags on Twitter to return tweets that contain these mentioned tags (Twitter, 2020b), which aids categorizing discussions within the same topic. One example below is when the hashtag "fake news" is used, which returns the following results (bear in mind that results vary from time, place and person as well that this is a small portion of the returning tweets):

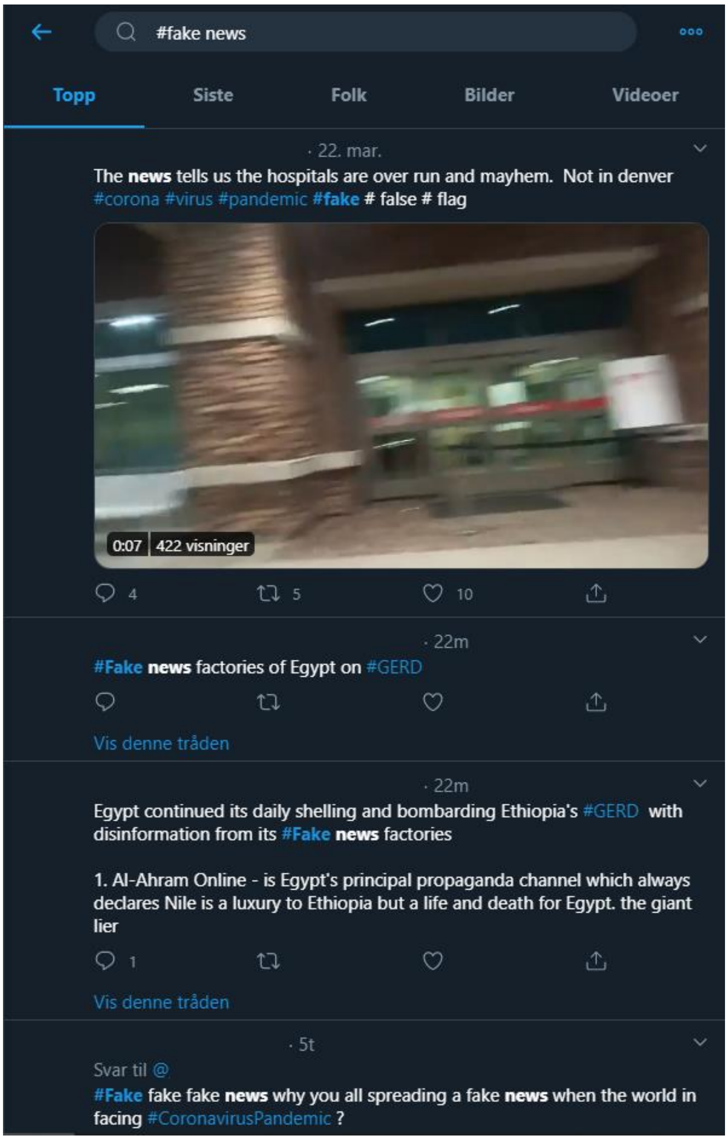


Figure 23 - Using the hashtag "fake news" - all persona info edited out

Analyzing the dataset, it was possible to extract the hashtags used in the tweet collected. For readability and scope, the results are limited to the 20 most mentioned hashtags and their frequencies:

Rank	Hashtag	Frequency	Rank	Hashtag	Frequency
1	Syria	142917	11	news	30455
2	migrantcrisis	97811	12	UNHCR	25468
3	EU	92703	13	UNGA	23583
4	tcot	88471	14	cdnpoli	22059
5	auspol	77951	15	refugeesGr	20303
6	Syrian	67326	16	rapefugees	18528
7	migrants	66668	17	UN	18177
8	MAGA	55696	18	pjnet	17402
9	SyrianRefugees	46291	19	immigration	17278
10	UN4RefugeesMigrants	43220	20	ccot	16726

Table 12 - Most used hashtags

4.1.3 Most liked tweets

When users are presented with various tweets, they have the option to either reply to it, retweet it (share), or like it (also called the *favorite*, although *like* will be the word of choice in this thesis). These options are not exclusive, which means one can do all three actions on the same tweet if one desires. For all practical purposes, a *like* can be seen as an approvement of the content of the tweet; thus, it is of relevance to see what type of tweets gained the most traction in the event window (entirety 2016). As with hashtags, the 20 most liked tweets are summarized below in their original form. Username is removed, due to the privacy agreement in the study. Frequency refers to like counter.

Rank	Tweet	Frequeny
1	Hillary's refusal to mention Radical Islam, as she pushes a 550% increase in refugees, is more proof that she is unfit to lead the country.	52501
2	Refugees are not numbers, they are people who have faces, names, stories, and need to be treated as such.	44012
3	A vote for Clinton-Kaine is a vote for TPP, NAFTA, high taxes, radical regulation, and massive influx of refugees.	39486
4	ISIS has infiltrated countries all over Europe by posing as refugees, and @HillaryClinton will allow it to happen here, too! #BigLeagueTruth https://t.co/U5hDdIc4rC	34304
5	The treatment by some towards these young refugees is hideously racist and utterly heartless. What's happening to our country?	29034
6	Black Lives Matter. All lives matter. All? All. Syrian refugees? Well.. https://t.co/gNxYUFaRSF	26326
7	Crooked Hillary wants a radical 500% increase in Syrian refugees. We canâ€™t allow this. Time to get smart and protect America!	24341
8	CLINTON REFUGEE PLAN COULD BRING IN 620,000 REFUGEES IN FIRST TERM AT LIFETIME COST OF OVER $400 BILLION. https://t.co/COZQNt6KVs	23474
9	@TarukMatuk: @CNN @FoxNews @realDonaldTrump @RogerRice10 Refugees from Syria over 10k plus more coming. Lots young males, poorly vetted.	18524
10	there are 2,500 life jackets used by refugees outside Parliament today, representing those who died on way to Europe https://t.co/cuRchnmJo1	16595
11	Here is a new poem entitled Refugees. Please bear with it. https://t.co/hREWTO6DrU	15425
12	What we're witnessing in coverage of Lily Allen and Gary Lineker is an attempt to make compassion towards refugees socially unacceptable	11438
13	Trump lies about refugees. Here is the YEARS-LONG VETTING PROCESS refugees go through before being resettled into the United States. https://t.co/cERC2oU2wb	10763
14	Now is not the time for demagoguery and fear-mongering. We will not turn our backs on refugees fleeing violence.	10049
15	The Republican Party could stand to remember who founded this country, and stop demonizing immigrants and refugees at every turn.	8959
16	#alllivesdidntmatter when y'all made up 10 million excuses as to why syrian refugees couldn't enter your countries	8717
17	.@brunelldonald: We want a president that puts Americans first - not illegal aliens, not refugees. https://t.co/xNzraMVx6D	8634
18	Lindsay Lohan has been doing charity work with Syrian refugees idk why this makes me so happy to see https://t.co/9GyBZHApPS	8059
19	American privilege is thinking you can move to Canada if Trump is elected, yet being skeptical of refugees seeking safe haven from war.	7664
20	I challenge President Obama to share with us a serious plan to screen out terrorists from the Syrian refugees.he is risking American lives	6463

Table 13 - Most liked tweets

4.1.4 Most retweeted tweets

Following the steps of hashtags and liked tweets, it is natural to look at the tweets that have been most retweeted, e.g., shared across the Twitter network at the period studied. Retweets are similar to likes, with a few key differences. Firstly, a user can retweet something with or without additional personal text as illustrated here from one of the groups personal (Tor Ole) Twitter account:

Figure 24 - An example of Retweet with a comment

Secondly, retweets are not limited to agreeing or disagreeing with the tweet, as the option for retweeting with comment opens (and is used) various use and intention of this function.

Analyzing the dataset, table 14 was produced, giving an overview of the most retweeted message throughout the year of 2016 (within those tweets caught in the data collection as described under method).

Rank	Tweet	Retweet count
1	RT @realDonaldTrump: TODAY WE MAKE AMERICA GREAT AGAIN!	352643
2	RT @realDonaldTrump: Such a beautiful and important evening! The forgotten man and woman will never be forgotten again. We will all come to…	225405
3	RT @Harry_Styles: Take a stand with us & @savechildrenuk: help make #RefugeesWelcome. http://t.co/pXUymBKqEn	124156
4	RT @FuchsOfficial: CHAMPIONS!!!! https://t.co/pFtvo5XUNx	123676
5	RT @LoganPaul: HUGE iPhone 7 GIVEAWAY! Just RETWEET this tweet & watch this video ?? ??https://t.co/lTfB5lLAXx https://t.co/dlpboygylh	118662
6	RT @StacyOnTheRight: Stop blaming white people for Trumps win last night. America voted for actual change. https://t.co/UIlSJcOIlg	105892
7	RT @blxcknicotine: So cute, it hurts. https://t.co/O2YbHAV9Rf	85464
8	RT @RFCdan: To people blaming refugees for attacks in Paris tonight. Do you not realise these are the people the refugees are trying to run…	79337
9	RT @realDonaldTrump: Just had a very open and successful presidential election. Now professional protesters, incited by the media, are prot…	71025
10	RT @MADBLACKTWINK: White people: All lives matter Syria: we got refugees in danger White people: new phone who dis	54826

Table 14 - Top 10 retweeted tweets

In the entirety of the dataset, 9.8 million (68.6%) of all the tweets are retweets. When compared to the entire dataset, these ten messages accounted for 8.64% (1235194) of all tweets and compared to the number of retweets (9.8 million), these accounts for 12.6% of all retweets. In comparison, these ten tweets account for 0.00006% of the number of tweets (14.3 million) within the dataset.

4.1.5 URLs within tweets

Many of the tweets collected contain URLs, pointing to other tweets, webpages, images, or general external material outside the original tweet. The reason for this can be many, for example, to save space in limited message length or to push various agendas (Burbach et al., 2019; Cui, Zhang, Liu, & Ma, 2011). Iterating through the dataset the following small summarize table 15 was created:

Total number of tweets	Tweets containing URLs	Unique URLs
14.3 million	8.6 million / 60% of all tweets	3.1 million / 21.7% of all tweets

Table 15 - Number of tweets containing URLs

We find that the majority (60%), or 8.6 million of the tweets caught in the dataset contains a URL. 4.4 million of the URLs pointed to a URL outside of Twitter. Of the tweets containing URLs, 36% of all the URLs were unique (roughly 3.1 million). Unique in this context means no duplicates or retweet of the same URL with or without a comment. When looking at the entire dataset, this means that 21.7% of all tweets contained a unique URL. In the entirety of the dataset the following table 16, consisting of the top 10 posted URLs is constructed:

Rank	URL	Frequency
1	https://twitter/***	26385
2	https://twitter.com/***/status/779554383256137728/photo/1	22510
3	https://twitter.com/***/status/712696517698723841/photo/1	11936
4	https://twitter.com/***	9559
5	https://twitter.com/***	8634
6	https://twitter.com/***	7915
7	https://huffingtonpost.com/entry/ted-cruz-syrian-refugees_us_564d279ae4b031745cefd25f	7879
8	https://twitter.com/***/status/728981004019486772\0/photo/1	7230
9	https://twitter.com/***	6560
10	https://cnn.com/2016/03/14/middleeast/syria-aleppo-behind-rebel-lines/index.html	5896

*Table 16 - 10 most tweeted URLs. *** = removed usernames*

The list is kept to ten in to increase the readability. While it would be of great interest to check the URLs in Table 16 above to give a contextual understanding of them, the focus related to URLs is in paragraphs 4.3 at attempting to find whether it contains fake news and, if such, classify them accordingly. The asterix (***) sign means the username has been removed (as those pointed directly to Twitter profiles).

4.1.6 Topic detection

Utilizing the workflow described in section 3.8.4.2 on the dataset provided, the following tag cloud, figure 26, was constructed and consisted of the various terms related to the different topics. The size of the term indicated the frequency (higher = more significant) and colored the corresponding topic in the color scheme below. The tag cloud is based on the LDA workflow (see 3.8.4.2), with 10 topics 10 terms. Attempts at other numbers yielded similar results, and one such can be seen in appendix F with 20 topics 10 terms. 10 topics, 10 terms became the final number utilized in this study.

Figure 25 - Tag cloud color scheme

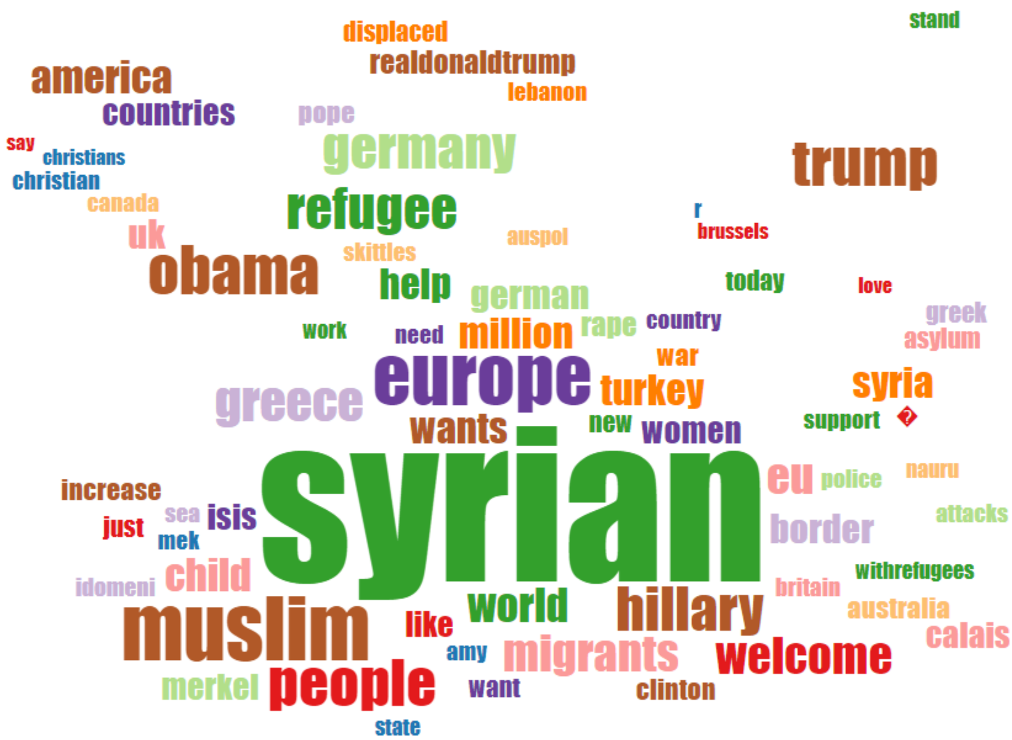


Figure 26 - Tag Cloud derived from workflow and based on the complete dataset

While the workflow for most parts worked as expected, some anomalies were present. One is found in topic 2, and the term � as well as topic five term *r*. The reason why these were present one can only speculate, but foreign characters that are not compatible with the KNIME workflow or a bug in one or more of the text processing nodes are likely candidates. All things considered, the impact of these two words amounts to 2% of the total number of terms and is thus considered of low impact on the overall results. The overview of the topics and their related terms can be viewed in table 17. Please note that an extended version of this with all term weighting within the corresponding topics can be viewed in the appendix G.1.

Topic	Terms	Frequency
topic_1	syrian, help, refugee, world, new, support, work, stand, withrefugees, today	2498219 | 8.23%
topic_2	welcome, syrian, people, ?, love, today, just, say, like, brussels	1910637 | 6.29%
topic_3	syrian, hillary, trump, wants, america, muslim, obama, realdonaldtrump, clinton, increase	4786138 | 15.76%
topic_4	people, europpe, like, want, country, countries, women, need, isis, syrian	2270108 | 7.47%
topic_5	syrian, obama, muslim, christian, amy, mek, america, r, state, christians	3094041 | 10.19%
topic_6	syria, million, syrian, world, war, displaced, turkey, people, countries, lebanon	3252414 | 10.71%
topic_7	germany, muslim, german, merkel, rape, europe, women, isis, police, attacks	3590176 | 11.82%
topic_8	syrian, australia, trump, canada, skittles, auspol, nauru, people, just, refugee	1974739 | 6.50%
topic_9	eu, child, uk, calais, europe, turkey, migrants, refugee, asylum, britain	3624764 | 11.93%
topic_10	greece, border, migrants, europe, syrian, pope, greek, refugee, sea, idomeni	3371960 | 11.10%
Total		30373196 / 100%

Table 17 - Topic and term overview

Topic 3 has the highest frequency in the dataset (15,76%), while topic 2 is the least present (6,29%). As for terms, *syrian* is the most present term (2,77%), while *brussels is* the least present (0,29%) when all topics are combined. We refer to the table in appendix G for the overview of all term loadings. As tweets could contain several topics, the total amount is above the 14.3 million. The finding above couplet with the findings previous contextual suggests that American political discussions are the most significant topic, although future research to verify if such a connection is present, and if so - to what degree.

4.2 Fake news detection & classification in most liked tweets

The structure of this section is based on the claims and links provided by the most liked tweets in section 4.1.3. Then based on classification from Zannettou et al. (2019), attempts are being made to classify what the type of fake news the claims fall within. Firstly, we will present our argumentation and, at the end of this section, table 14, to illustrate the classification. The tweets that act more as a personal opinion rather than claiming to be a news item/opinion of others are not considered, such as rank 2, 5, 6, 10, 11, 12, 13, 14, 15, 16, 18, 19 and 20. These tweets act more like a person's opinion rather than telling other people's agendas. The URLs in all tweets are manually checked against the webpages Snopes, Politifact, Factcheck.org, as well as identified literature to see whether it contains fake news in one of its forms. The ranks below refer to table 13 in section 4.1.3 and are kept here, meaning rank refers to rank in table 13 and not as in order they are listed in this section. The purpose of this section (as well as 4.2.1 and 4.2.2) is to address RQ2. It relies on open-source manual fact-checking to determine whether something is fake (and why) both from the perspective of the fact-checking services but also through the extended classification framework utilized in this study.

Rank one and the highest liked tweet in the dataset: We find two prominent claims here; the refusal of acknowledge of radical Islam and a 550% increase in refugees. As for refusal to mention radical Islam, this is both propaganda and biased. Without knowing for sure who posted this tweet, it is clear that it is within the political realm, as it intends to put Hillary Clinton (Clinton from now on) in a bad light to favor Donald Trump (Trump from now on). The tweet

also offers little to no balance in its presentation and focuses on one narrative/view only. PolitiFact has classified this statement as mostly false and points out that the origin of this claim is from a speech performed by Trump in 2016 (Sherman, 2016).
The next claim in this tweet is the 550% increase in refugees. While it is true that Clinton suggested that America should increase the number of refugees they let in at the time (from 10 000 to 65 000, (Jr., 2016)), it is presented very one-sidedly. Couplet with the claim from before about radical Islam, it gives the impression of Clinton aiming to flood the country with potential terrorists. The claim originates from the same speech as the previous claim.

Rank two and the third-highest liked tweet in the dataset: There are several claims within this one, and while they are not directly false, they are like the previous claims presented out of context or in such a way that it does not give a correct image for the actual situation (Contorno, 2016; Emery, 2016a; Farley, 2016; Jr., 2016). It (the tweet) is therefore classified as biased or one sided. One could go into great length here elaborating of why these are only partially true, yet the point is highlighting the classified. As with the other claims, many of these stem from one of the speeches by Trump (Contorno, 2016).

Rank four: contains two claims; that ISIS has infiltrated Europe by posing as refugees and that Clinton will allow this to happen in America. The message is actually identical to a tweet by Trump, and the URL points to this original tweet (when it is decoded from the Twitter URL shortening algorithm): https://twitter.com/realDonaldTrump/status/788930678255517696. The hashtag BigLeagueTruth (which explains why the message is identical to that of the above) refers to a social media group that spreads pro-Trump support on social media (Jamieson, 2016). As for claims at hand, these are classified as fabricated, propaganda, and conspiracy theories. From what we can find, there is very little concrete information regarding ISIS being camouflaged as refugees (or at least in such a way it can be generalized), and those stories that actually surface in the media seem to be entirely fabricated (Lacapria, 2015). The second claim regarding Clinton has little roots in the real world. As pointed out in the paragraph above, she advocated for an increase in refugees American should aid, yet generalizing a whole population/religion is at best of times misleading.

Rank six: Checks were conducted on the URL https://t.co/gNxYUFaRSF, which points to a tweet containing an image. The image's agenda speaks for itself (although debatable whether it should be within satire category of fake news framework - yet ultimately ruled out of this as it is not directly targeting an individual or politician), and needs no further elaboration:

Figure 27 - Image related to tweet, personal info removed

Rank seven: relates to the nickname "Crooked Hillary" as well as the same claims as third-ranking tweet, which has been discussed previously. The focus of this paragraph will then be Crooked Hillary. The term stems from Trump's ad campaign, and Trump regularly used the term as a collective term for news items or debate with or about Hillary and was not limited to one particular case. It intended to frame Hillary as dishonest (Bond et al., 2017). A previous study found that associations to Hillary on Twitter often came in terms such as "emails," "FBI," "WikiLeaks," "#DrainTheSwamp" amongst others, which contains an array of various rumors and claims related with them (Darwish, Magdy, & Zanouda, 2017).
As for classification, this tweet inherits propaganda and biased or one sided from the 550% refugee increase claim, as well as fabricated and conspiracy theories related to Crooked Hillary. Arguments can be made that rumors are more suitable than conspiracy theories as a classification here. However, the argument from the group is the severity and impact of the rumors/theories in combination with the groundless claims in many of the conspiracy theories tied to Crooked Hillary (such as Hillary signing uranium deals on behalf of America (FactCheck.org, 2017a; Putterman, 2018)) speaks for this more severe classification.

Rank eight: contains two claims; that Hillary will allow 620 000 refugees to America at the costs of 400 billion dollars. The URL (https://t.co/COZQNt6KVs) points to a session in the Senate, which later has been made inaccessible to the public for unknown reasons. The claim (although due to pseudonymization, it is not possible to say conclusive who posted the tweet) stems from Trump's ad campaign. As pointed out previously, the actual number suggested in Hillary's campaign was about 65 000 refugees each year, thus rendering this claim entirely false. Costs related to refugees over such a period of time is hard - if not impossible, to calculate correctly (Valverde, 2016). The tweet is classified as both fabricated and propaganda, as it is focused on political issues and fabricates the claims.

Rank nine: contains three claims; 10 000 refugees (and counting) as well as generalizing of the population (of refugees) and their screening process. As for numbering, according to statistics made available after 2016, the number of (Muslim) refugees America admitted was 38 900, and is thus not classified as fake news (Krogstad, 2019). Next is the generalization of the population, claiming mostly men are the population of the refugees. This is not correct, as studies show that various numbers of women and child (from roughly 50% to three-quarter) makes out the population of Syrian refugees (Binkowski, 2016; Valverde, 2017). Furthermore, the vetting process is insufficient. This, too, is wrong according to fact checks done at two different periods (Nicols, 2017; Quu, 2016). The origin of this claim seems to stem from yet again Trump and the congress. These two latter claims are thus classified as fabricated.

Rank ten: A Tweet with a URL at the end (https://t.co/cuRchnmJo1) and points to a tweet containing an image. The image information was checked against image checking webpages such as TinEye (TinEye, 2020) to verify its authenticity. The image is authentic, and the story related to it has been featured in a couple of articles in Washington Post (amongst others).

Rank 11: Tweet contains an image attached to it. It reflects the opinion of the user through a witty poem. It contains no content suitable for being classified as fake news.

Rank 12: claims that Trump lies about the refugee vetting process. Tied with the tweet is an image of the vetting process, and an image stamp claiming the image posted comes from the White House. By reverse-searching the image couplet with manually checking the webpages of the White House, the image is authentic and originates from the White House website (whitehouse.gov, 2015). Coupled with the fact-check conducted on rank nine, the claim holds water and is thus not classified as fake news.

Rank 17: claims that other presidents than Trump are more focused on aiding refugees than aiding the country. Tied with the claim is a URL to a video provided by Fox News, which contains the opinions of citizens (which contains a statement that makes out the claim). The various claims provided within the video are not considered, or fact-checked, only the statement above and how it is presented. The claim ultimately ends up being classified as biased or one sided, as it is very generalizing, provides little evidence to back up the claim, and does not discuss the topic from different perspectives (it solely focuses one narrative).

Rank 19: claims regarding Lindsay Lohan - an American actor, doing charity work with Syrian refugees, tied with an URL. The URL contained in the tweet has since it was posted been removed from the internet, and the focus is therefore on verifying whether such a story took place in the first place. While nothing can be said conclusive (as we do not know what the original URL contained), several webpages in the timespan of the data collection covered Lohan performing charity work related to Syrian refugees (Gordon, 2016; Moore, 2016). While the webpages hold various quality and trustworthiness, the amount of results (when limiting to 2016) seems to indicate this claim is true, thus prevents the claim from being classified as fake news.

Summarizing this manual check of all the claims provided in the top 20 liked posts within the dataset and by following the framework provided by Zannettou et al. (2019), table 14 is created. Note that due to the tweets containing various claims, with various characteristics, multi-classification is used.

Tweet	Fake news category within Tweet						
	Fabricated	Propaganda	Conspiracy Theories	Hoaxes	Biased or one sided	Rumors	Clickbait	Satire
Hillary's refusal to mention Radical Islam, as she pushes a 550% increase in refugees, is more prood that she is unfit to lead the country		x			x			
A vote for Clinton_kain is a vote for TPP, NAFTA, high taxes, radical regulations, and massive influx of refugee					x			
ISIS has infiltrated countries all over Europe by posing as refugees, and @HillaryClinton will allow it to happend here too! #BigLeagueTruth https://t.co/U5hDdIc4rC	x	x	x					
Crooked Hillary wants a radical 500% increase in Syrian refugees. We can't allow this. Time to get smart and protect America!	x	x	x		x			
Cliton refugee plan could bring in 620 000 refugees in first term at lifetime cost of over $400 billion. https://t.co/COZQNt6KVs	x	x						
(...) Refugees from Syria over 10k plus more coming. Lots of young males, poorly vetted.	x							
(...) We want a president that puts Americans first - not illegal aliens, not refugees. https://t.co/xNzraMVx6D					x			

Table 14 - Fake News classification of the most liked tweets

Reading these results, we find that of the 20 most liked messages, 7 (35%), contains fake news in one of its forms. The distribution of the classification is as following (% within the identified fake news messages (7) / % of the total number of tweets (20)):

Fabricated: ~ 57% / 20%

Propaganda: ~ 57% / 20%
Conspiracy Theories: ~ 28% / 10%
Biased or one sided: ~ 57% / 20%
Hoaxes / Rumors / Clickbait / Satire: 0% / 0%

While one could argue that the support multi-classification of the messages advocates that hoaxes, rumors, clickbait, and satire also could be checked in some of these messages, the approach has tried to distinguish them based on severity and focus of the message. The messages above are primarily driven by negatively targeting other individuals and have, in most cases, a political tone to it rather than societal tone. One example: If the tweet that stated Hillary Clinton would allow ISIS to get a foothold in America instead focused on ISIS potentially infiltrating countries as refugees. That would mean its focus had shifted away from the political over to a more questionable (of the situation) tweet, which then would be classified as rumors or biased, as the political agenda was not in direct focus.

4.2.1 Fake news tweets timelines

To get a broader understanding of the characteristics around the propagation of the 7 top tweets identified to contain fake news, we studied the occurrence of the original tweet throughout the year. This was done by tracking likes, retweets, and general occurrence of the original text through the entirety of 2016 (regardless of actors).

The first tweet with the wording *Hillary's refusal to mention Radical Islam, as she pushes a 550% increase in refugees, is more proof that she is unfit to lead the country*, saw the following propagation pattern:

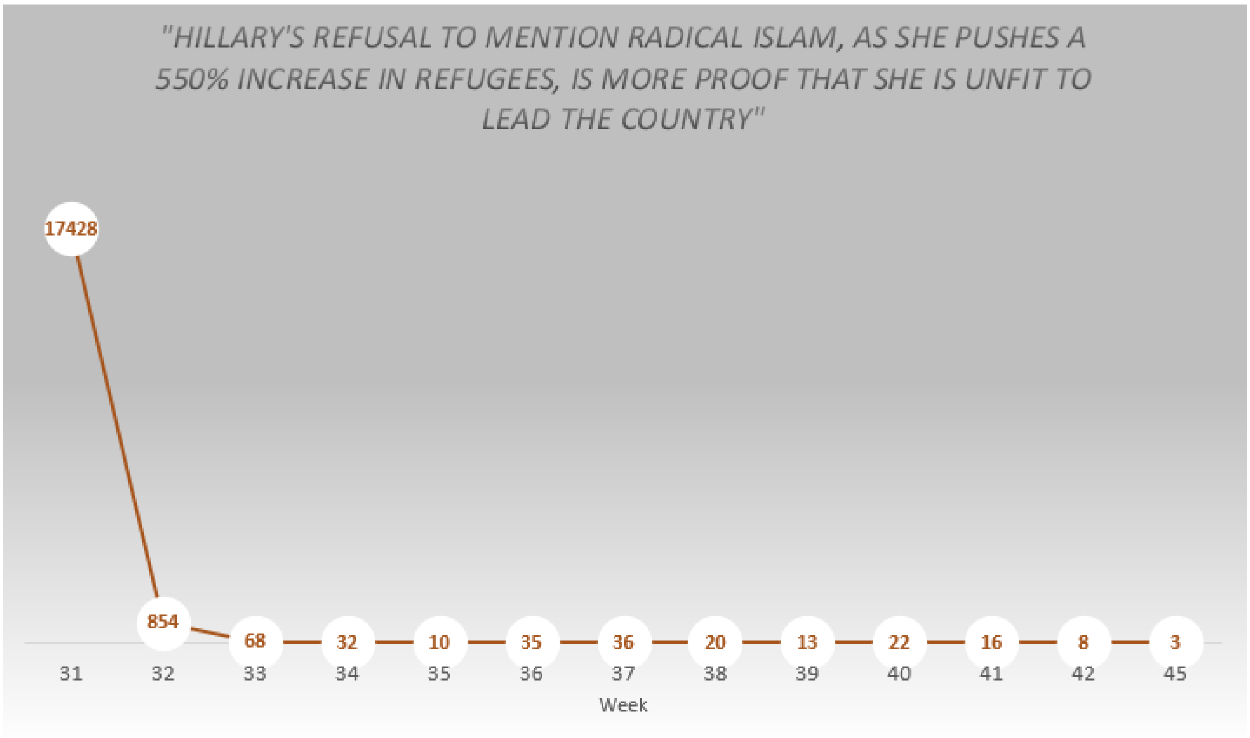


Figure 28 - Timeline propagation in most liked tweets #1

A finding here is that fact-checking was conducted on these claims prior to the tweet, as highlighted under section 6.3.1, but judging the spread in week 31, this held little effect on the spread. 93.98% of the total spread of this tweet happened during week 31. It had a rather long lifespan with 13 weeks of diffusion. As previously pointed out, the data from November and December is lacking in the dataset. Thus chances are the lifecycle might be longer than found here.

The second tweet is *A vote for Clinton-Kaine is a vote for TPP, NAFTA, high taxes, radical regulation, and massive influx of refugees*, with the following propagation pattern:

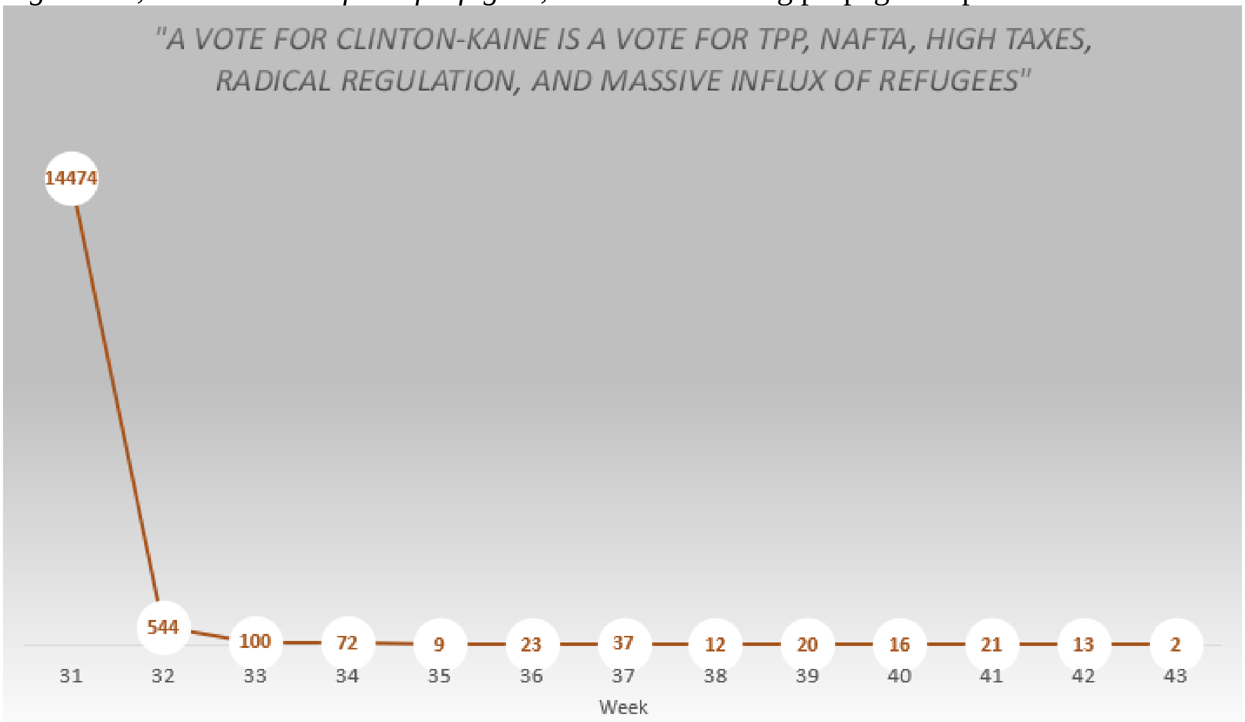


Figure 29 - Timeline propagation in most liked tweets #2

As seen with the first tweet, there is a clear overweight of the spread being centered around the first week of creation, with 94.33% of the diffusion happening that week alone. This too, has a long life cycle, as its diffusion continued for 14 weeks.

The third tweet is *ISIS has infiltrated countries all over Europe by posing as refugees, and @HillaryClinton will allow it to happend here too! #BigLeagueTruth* https://t.co/U5hDdIc4rC. Its diffusion pattern was the following:

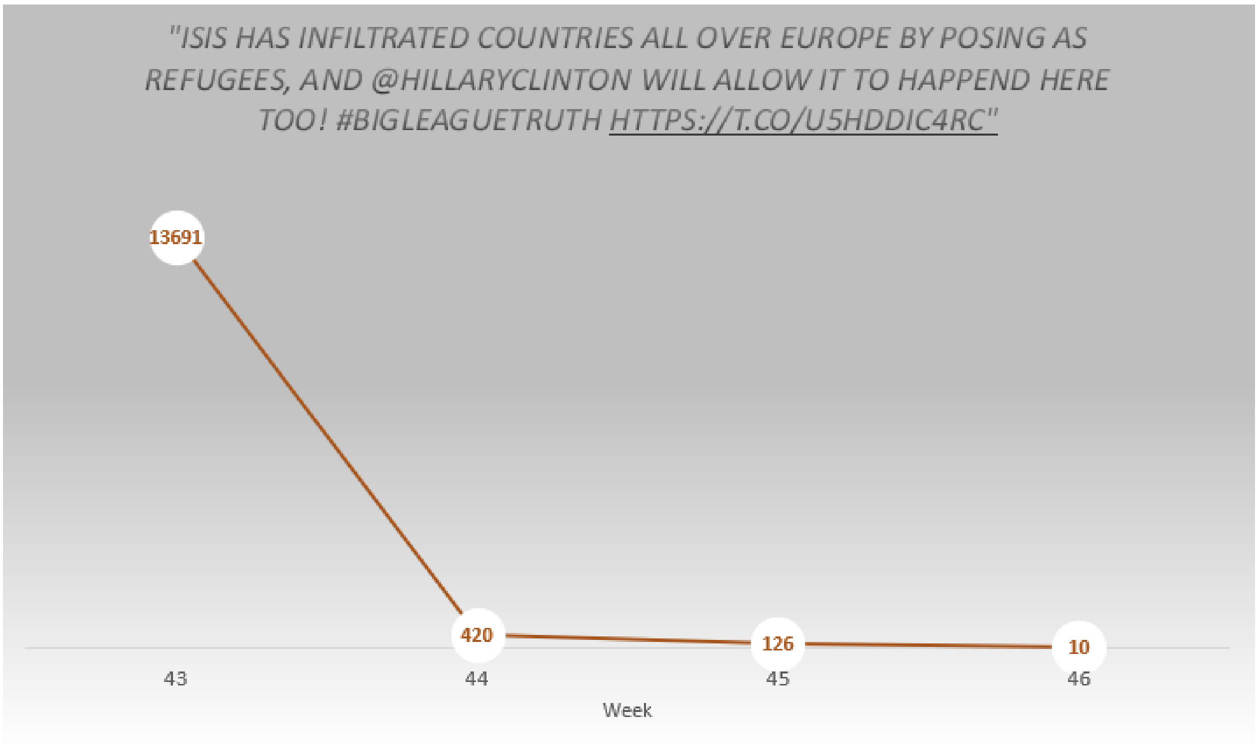


Figure 30 - Timeline propagation in most liked tweets #3

This tweed also shares a diffusion pattern where most of the diffusion happens the first week, with 96.1% of the diffusion the first week alone. As opposed to what we saw in the first two tweets, this has a shorter lifecycle with its four weeks in circulation.

The fourth tweet is *Crooked Hillary wants a radical 500% increase in Syrian refugees. We can't allow this. Time to get smart and protect America!* The propagation pattern looked like the following:

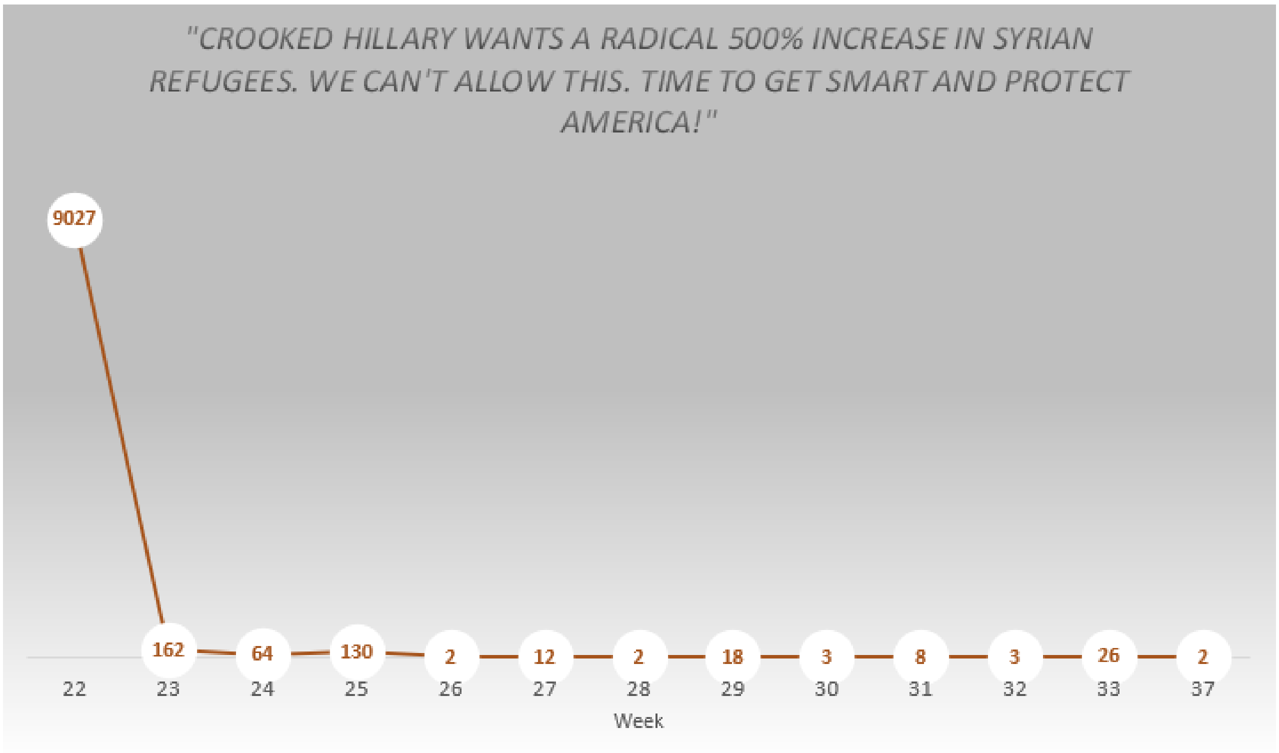


Figure 31 - Timeline propagation in most liked tweets #4

This tweet also had a high first-week distribution with 95.33% of the total diffusion in this week alone. Additionally, it has one of the most extended life cycles with its 17 weeks of diffusion. Its life cycle overlapped with the fact checks that happened at the time (see section 6.3.1), yet it continued its life cycle far beyond this point.

The fifth tweet is *Clinton refugee plan could bring in 620 000 refugees in first term at lifetime cost of over $400 billion. https://t.co/COZQNt6KVs*. The diffusion was as follows:

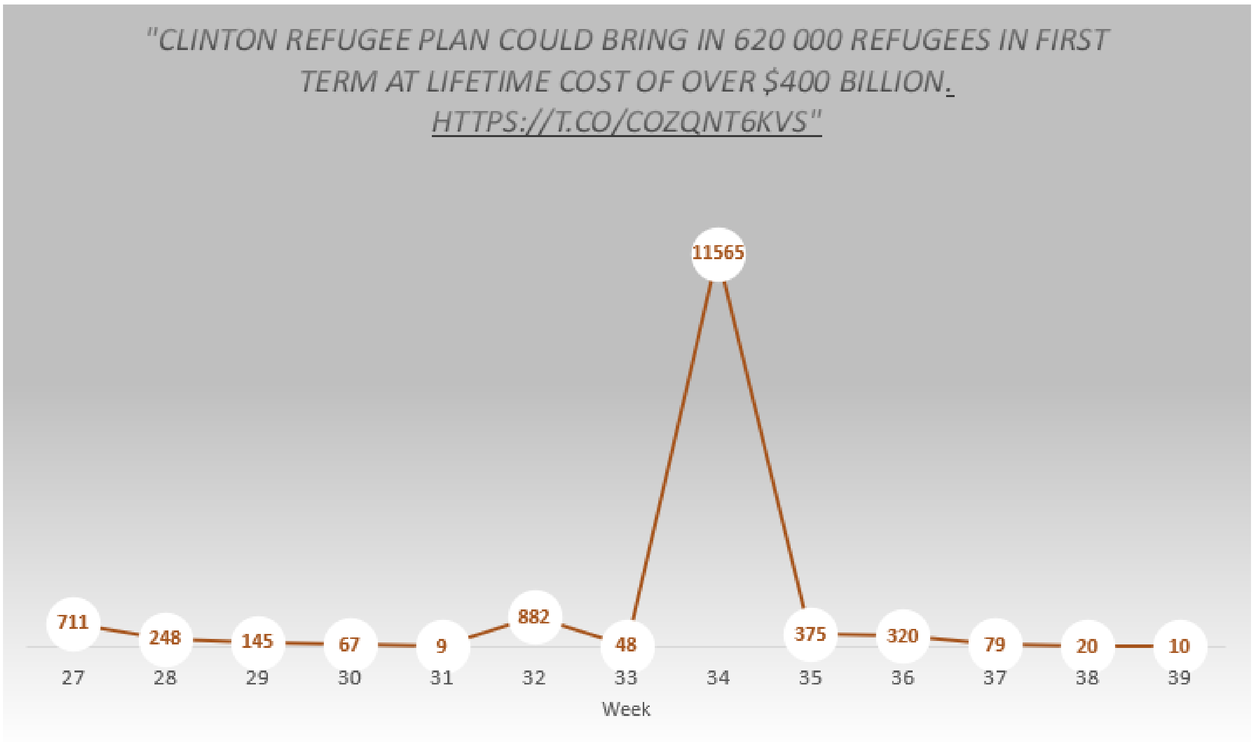


Figure 32 - Timeline propagation in most liked tweets #5

As opposed to the other claims, the distribution is more spread. The first week of spread sees frequency of 4.88%, while week 34 contains 79.40% of the total diffusion. Furthermore, it has the most extended lifecycle of the tweets checked in-details with 19 weeks of diffusion. The diffusion overlaps the fact checks (see section 6.3.1) and continues until week 46 - as many of the other tweets.

The sixth tweet is *(…) Refugees from Syria over 10k plus more coming. Lots young males, poorly vetted.* The snipped section contains only @ tags, which are a reference to other Twitter users. As we focus on the actual content of the tweet, this has been removed. Its diffusion was as follows:

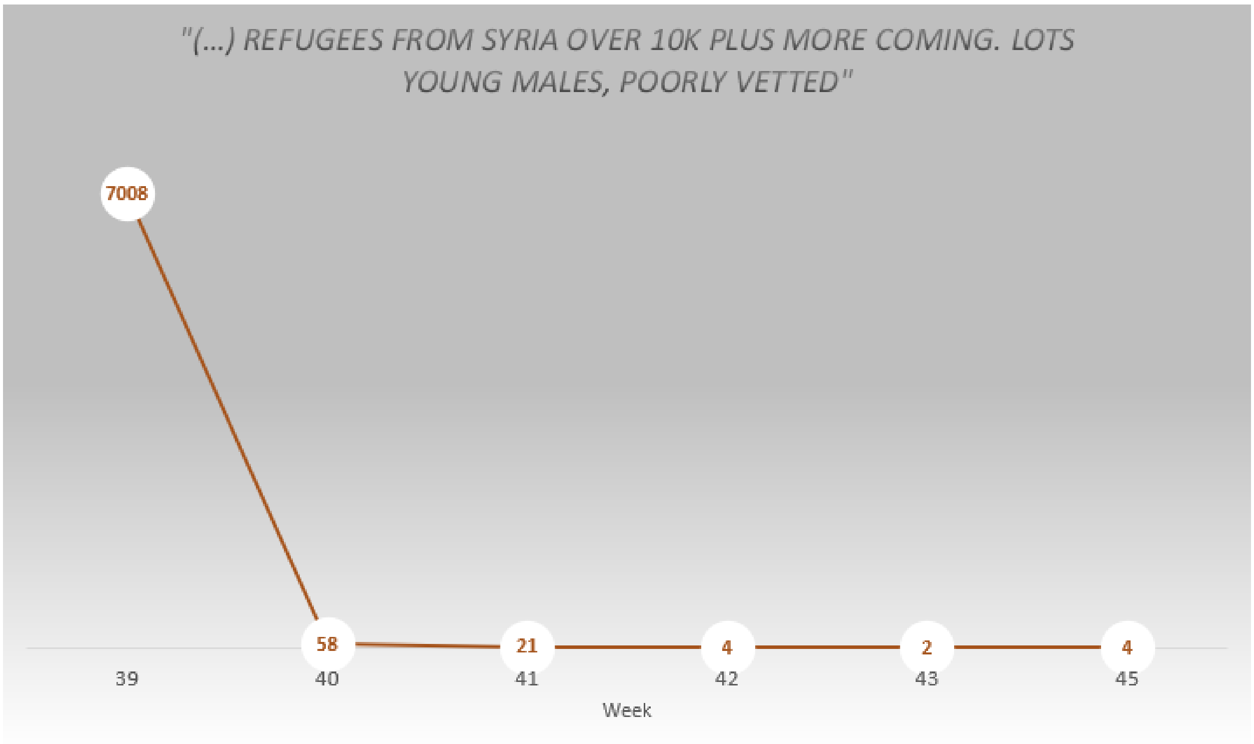


Figure 33 - Timeline propagation in most liked tweets #6

The first week of diffusion saw an overwhelming distribution, with 98.75% of all the occurrences of the tweet that week alone. Its lifecycle was six weeks. From what we can tell looking at the fact checks, these came afterward (see section 6.3.1) the claims in the tweet, thus rendering it of little effect on diffusion.

The seventh tweet is: *(...) We want a president that puts Americans first - not illegal aliens, not refugees* [*https://t.co/xNzraMVx6D*](https://t.co/xNzraMVx6D) resides. The excluded parts of this tweet are @ tags, and as spot six, these were removed. The diffusion was as follows:

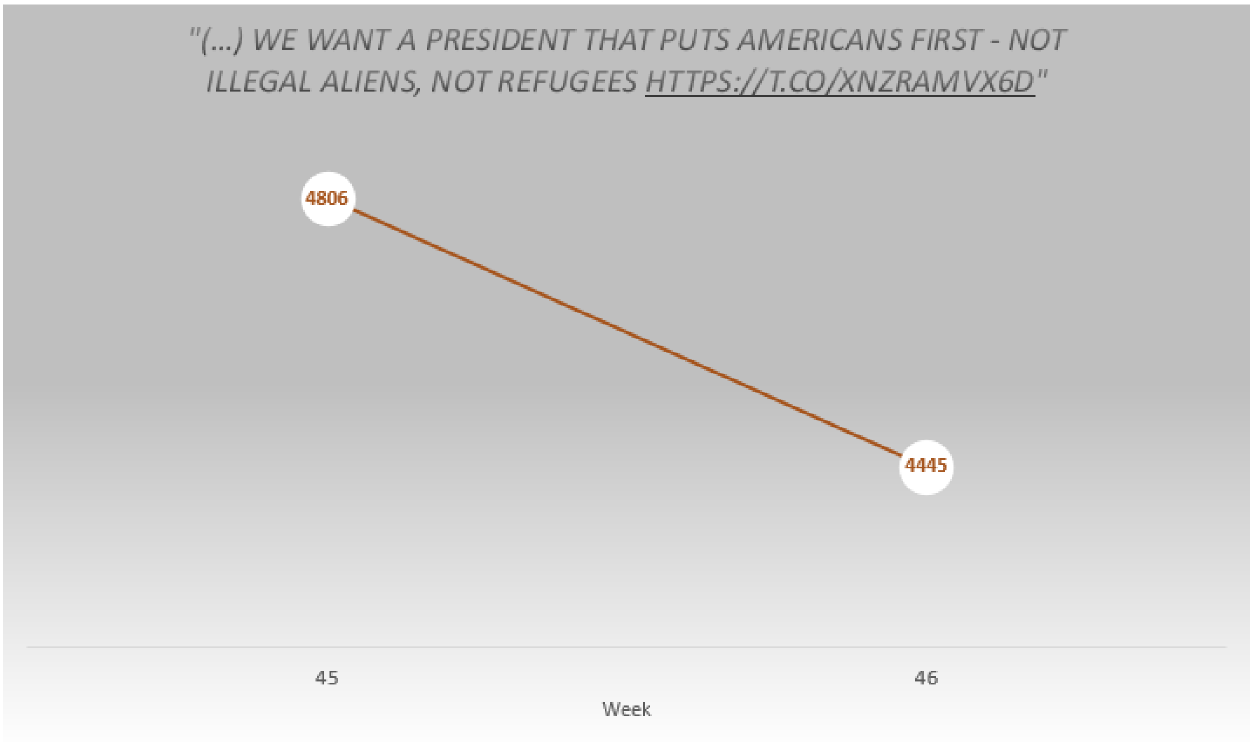


Figure 34 - Timeline propagation in most liked tweets #7

As opposed to the other tweets, this both held the shortest lifecycle and most spread frequencies. The first week of diffusion sees 51.95% of the distribution. As pointed out in section 6.3.1, there are no direct fact checks that are suitable to check against the claims in the tweet. Thus it should hold a little effect on the diffusion.

4.2.2 Summarizing fake news tweets timeline

Summarizing the timelines for all the seven messages identified to contain fake news in one of its forms can be done with the following graph:

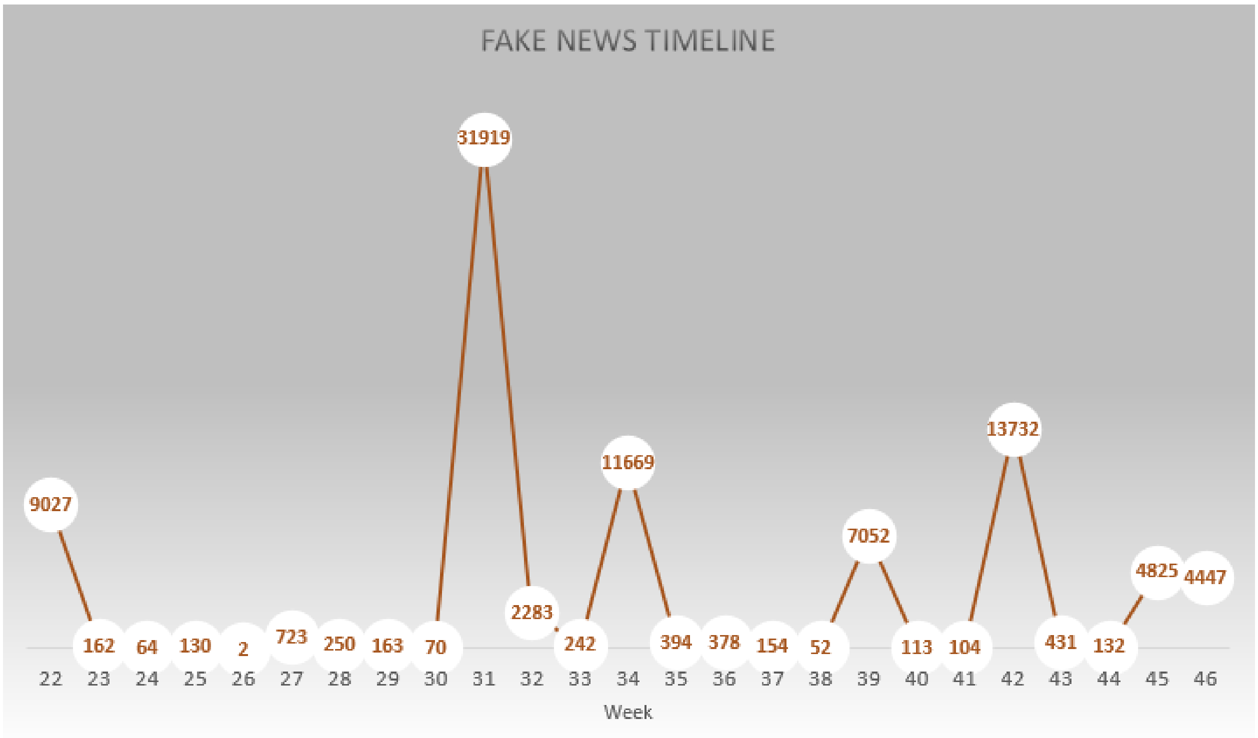


Figure 35 - Summarizing timeline propagation in most liked tweets

There are spikes in week 22 (10.20%), 31 (36.06%), 32 (2.58%), 34 (13.18%), 39 (7.97%), 42 (15.51%), 45 (5.45%) and 46 (5.02%). These spikes in occurrences make out 95.97% of the total number of times (88518) one of the seven messages were tweeted. Comparing these seven tweets towards the total tweets collected in the dataset, these messages make out 0.62% of all the entries in the dataset. This is an extensive amount when these tweets account for 0.00015% of the unique tweets within the dataset.

4.3 Fake news detection & classification of URLs

As done in the section dwelling into the most liked tweets, the same was conducted for URLs (URL and webpages refer to the same in this text, and both will be used). This process was done both automatically and manually. Using custom-made KNIME workflow, the URLs were checked against a list of 1387 entries known to spread fake news in one of its forms. As the analysis in this part focuses on the webpages in general (the URL domain), so is the classification, as we did not check each individual claim posted by the webpages. Based on the approach above, table 15 was produced, that checks whether the URLs was present in the dataset and if such the frequency (left side). It also lists the most shared webpages regardless if classified as fake news (right side).

The purpose of this section (as well as 4.3.1, 4.3.2, and 4.3.3) is to address RQ2. It follows a similar structure and approach as what is done in section 4.2. The various webpages will be briefly discussed and classified based on fact-checks performed on it, literature identified, and according to the framework provided by Zannettou et al. (2019).

Top 20 identified fake news URLs			Top 20 identified URLs		
Rank	Count	URL	Rank	Count	URL
1	82124	breitbart.com	1	95992	theguardian.com
2	37394	express.co.uk	2	82124	breitbart.com
3	36428	rt.com	3	80799	gu.com
4	20302	infowars.com	4	75338	youtu.be
5	18830	dailycaller.com	5	75036	nyti.ms
6	12411	truthfeed.com	6	69496	shr.gs
7	12323	zerohedge.com	7	56028	independent.co.uk
8	9114	jihadwatch.org	8	50652	dailym.ai
9	8699	sputniknews.com	9	46624	youtube.com
10	7250	wnd.com	10	44443	cnn.it
11	7206	barenakedislam.com	11	43418	amp.twimg.com
12	6517	shoebat.com	12	37394	express.co.uk
13	6326	therealstrategy.com	13	36428	rt.com
14	6167	conservativetribune.com	14	35459	ind.pn
15	5910	100percentfedup.com	15	32865	on.mash.to
16	5706	pamelageller.com	16	32417	aje.io
17	4900	thegatewaypundit.com	17	31139	bbc.in
18	4678	weaselzippers.us	18	26364	huffingtonpost.com
19	4496	wikileaks.org	19	25557	unhcr.org
20	4481	jewsnews.co.il	20	25396	nytimes.com

Table 15 - A selection of fake news webpages within the dataset (left) and overall top URLs (right)

For readability and scope, the list is limited to 20 most frequent results. The list of fake news URLs consists of 484 unique entries in total, with an average frequency of roughly 730. Of the entire dataset consisting of 14.3 million tweets, 3% of all tweets (435.099) contained a reference to one of the fake news URLs. As with the classification of most liked tweets, there is provided a table at the bottom of this section, table 16. As this proved to be an extensive amount of work, the top 16 results (left side) were manually evaluated. Its structure is also more in the shape of a discussion, as the group argues that it is natural to briefly discuss how the URLs are classified based on the fact checks and literature identified. Thus, much of the discussion relating the classification happens within this chapter as the analyzes performed by the fact-checking webpages (Snopes, PolitiFact, Factcheck.org) are used to determine the final classification. The right side of the table shows that 3 out 20 (15%) of the most shared URLs in the dataset are classified as fake news according to the list used. Rank in the text below refers to the rank on the left side of table 15.

Rank one is breitbart.com. Breitbart sees more than double the amount as the second entry in the list (82124 versus 37394). Breitbart seems to focus its effort around American politics and offers a range of news pieces with different topics with a tendency of being right-leaning. All three fact-checking sources have, on numerous occasions, checked the content of this webpage, and while there are stories with truth to them, the majority of the news items has been checked has been classified as fake in one way or another (FactCheck.org, 2020b; PolitiFact, 2020c; Snopes, 2020c). In the literature, numerous research has pointed to a similar classification (Bovet & Makse, 2019; Jang et al., 2018; Murungi et al., 2018).
Breitbart is, due to its sheer number of new items on its webpage, fact-checks (aimed at it) and literature, classified rather uniquely as all of the eight categories.

Rank two is express.co.uk, which is referred to as Daily Express (Evon, 2020). The webpage is by Snopes referred to as a British tabloid paper, and a handful of claims have been checked by the service. Of those checked, one returned to be a valid claim, one unverified (rumor) and one as clickbait (Snopes, 2020d). In the literature, it is also referred to as tabloid news, and it has been seen to attract the interest of account with dubious intentions (Bastos & Mercea, 2019, p. 45). The classification of this webpage is then both rumors and clickbait, in line with the classification of Snopes (2020).

Rank three is rt.com is to be found. rt.com refers to Russian Today, a fact-checking webpage that is seemingly focusing its efforts at cases that aligns well with Russian interests (Yablokov, 2015; Zannettou et al., 2017). Poynter.org did case-by-case research on 16 of the most prominent claims (Poynter.org, 2018), and landed on the conclusion that the webpage contains both factual fact-checking but also biased opinions with shortcomings in its methods. Poynter warns that the most significant issue here is the implications dubious fact-checking webpages have on the public perception of them, as it undermines the public trust of these (Valeeva, 2017). Previous research has classified the webpage as a conveyer of fake news in one of its forms (Maddock et al., 2015; Zannettou et al., 2017; Zannettou, Caulfield, et al., 2019).
Classification of rt.com is then propaganda, conspiracy theories, and biased or one sided, based on the extensive fact-checking performed on this webpage.

Rank four is infowars.com. The webpage focuses on right-leaning news and conspiracy theories in the context of American politics. The webpage has, on numerous occasions, been fact-checked and received multiple classifications such as propaganda and conspiracy theories in particular (FactCheck.org, 2020d; PolitiFact, 2020d; Snopes, 2020a). In the literature, similar classifications have been given (Bevensee & Ross, 2019; Bovet & Makse, 2019; Maddock et al., 2015; Zannettou et al., 2017). Our ruling is that the webpage is classified as fabricated, propaganda, conspiracy theories, and biased or one sided.

Rank five is dailycaller.com, an online publication that seems to lean towards the conservatives in American politics (PolitiFact, 2020a). By fact-checking webpages standards, it has been evaluated to post fabricated stories with little to no roots to the reality ("Daily Caller," 2020; PolitiFact, 2020a) and from the literature, it has been classified as a webpage that spreads (or is categorized as) fake news (Agrawal et al., 2019; Bovet & Makse, 2019). Classification for dailycaller.com is therefore fabricated, propaganda, and biased or one sided.

Rank six is truthfeed.com. The webpage is no longer operational. When it was active, it did focus its stories on conservative-leaning stories (FactCheck.org, 2017b). Of the fact checks that have been conducted on this webpage, 3 out of 4 have been fabricated claims with a political agenda (Evon, 2016a; PolitiFact, 2020g). In the literature, the webpage has been known for spreading fake news (Bovet & Makse, 2019). The webpage is classified as fabricated, propaganda, and biased or one side; this due to its political sway, coupled with groundless claims in those claims that have been checked.

Rank seven is zerohedge.com, which is known to spread groundless conspiracy theories in the context of American politics (Palma, 2017). Previous research has come to similar conclusions, with the addition of seeing a clear tendency for a right-leaning focus (Bevensee & Ross, 2019; Bovet & Makse, 2019). The webpage is therefore classified as fabricated, propaganda, conspiracy theories, and biased or one sided.

Rank eight is jihadwatch.org. The webpage focuses on news items concerning negative viewed Islamic news. Snopes (2018) has, with two occasions, checked claims originating from the webpage, whereas one was conspiracy theories and the other as false, fabricated information (Lacapria, 2018; Snopes, 2018). While these classifications came after the dataset studied, it stands to believe the focus of the webpage has not changed significantly over time. Literature has classified the webpage as a spreader of islamophobia (Awan, 2014). The webpage is classified as fabricated, conspiracy theories, and biased or one sided, due to the way it both fabricates its stories as well as how they are presented.

Rank nine is sputniknews.com (often referred to as just Sputnik). The webpage has similar to rt.com ties to the Russian government, both through its focus and support (Salvo & De Leon, 2018; Watanabe, 2018). PolitiFact has indirectly covered it on several occasions in relation to Trump, which mostly has turned out to be fabricated fake news (PolitiFact, 2020f). Furthermore, it is in previous studies classified as a noticeable fake news actor (Zannettou et al., 2017). As several of the fact checks have shown that many of the claims are entirely fabricated, couplet with the ties to the Russian government, its classification is both fabricated and propaganda.

Rank ten is wnd.com (also referred to as WorldNetDaily). The webpage focuses on American politics, with clear tendencies of a conservative lean (Orso, 2016). From the fact-checks that have been performed, these ranges from being fabricated stories, conspiracy theories, or strongly biased (FactCheck.org, 2020c; PolitiFact, 2020b; Snopes, 2020b). Literature has found that the webpage is both a spreader of fake news and susceptive to being utilized by bots (Narayanan et al., 2018; Shao, Ciampaglia, Varol, Flammini, & Menczer, 2017). The webpage is classified as fabricated, propaganda, conspiracy theories, and biased or one sided.

Rank 11 is barenakedislam.com. The webpage focuses on anti-Muslim news with a clear right-wing focus (Palma, 2018). There are not many fact-checks performed on the webpage that fits the timeframe, yet, one claim was found to be a fabricated story concerning American flags on a Democratic National Convention (Evon, 2016b). Literature has marked the webpage as a conveyer of junk news (Narayanan et al., 2018). The classification of the webpage is fabricated, propaganda, conspiracy theories, and biased or one sided. The reasoning behind this is due to its fabricated stories with an apparent political sway to it, with little to no balance in its presentation.

Rank 12 is shoebat.com. There are limited fact checks performed on this webpage, while the one we did find is by Snopes that debunked one of its claims that Hillary denied Christians in America to practice their faith (Emery, 2016b). The same goes for literature, which turns back empty. Judging from the webpage’s front page, it is however pretty clear that the webpage’s focus on rumors and conspiracies:

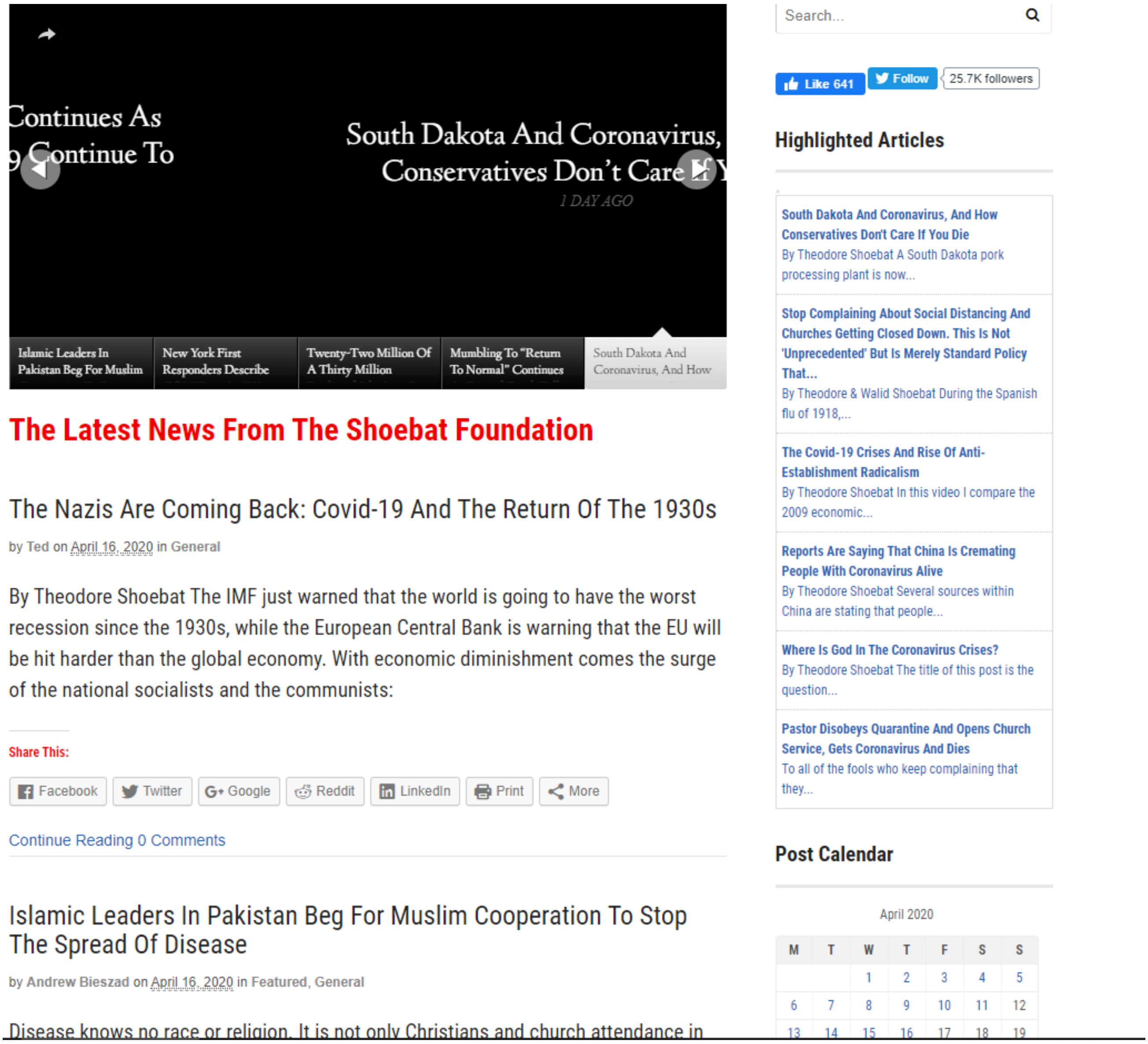


Figure 36 - Frontpage of shoebat.com

Based on the one fact check performed by Snopes (2016) as well as observation by the study group, the webpage is classified as fabricated, conspiracy theories, rumors, and clickbait. To elaborate further, most of the stories do not contain any sources for the claims, and those stories that have sources often point to YouTube videos (Shoebat, 2020). Furthermore, language (poor language in the articles, headers, and general) is of such level that its credibility is questionable.

Rank 13 is therealstrategy.com. The webpage is no longer accessible, as well that there are no fact checks performed by the three services utilized in this study. We find traces of it in the literature, where it sees a connection to social botnets (Nied et al., 2017), with a focus on alternative (rumor) news items (Zannettou et al., 2017). The classification is solely based on the two articles found and is classified both as rumor and fabricated, as it was pointed out by Nied et al. that the news items related to this webpage saw several alterations to the content over time (news item is altered to fit the narrative).

Rank 14 is conservativetribune.com. As for fact-checking, this webpage does not seem to have been extensively investigated, but for those few cases where it has been checked, indications of it spreading fake news in one of its forms is present (Jacobson, 2016; Robertson;, Farley;, Kiely;, Gore;, & Schaedel;, 2016). In the literature, it also seems to be classified as above (Zang & Pottle, 2016). The webpage has since merged with The Western Journals and is now under the domain westernjournal.com (Journal, 2020). By its name, and by manually checking the webpage, tendencies of being strongly favorable of conservative politics is present, which also can be illustrated by today (16.04.20) frontpage:

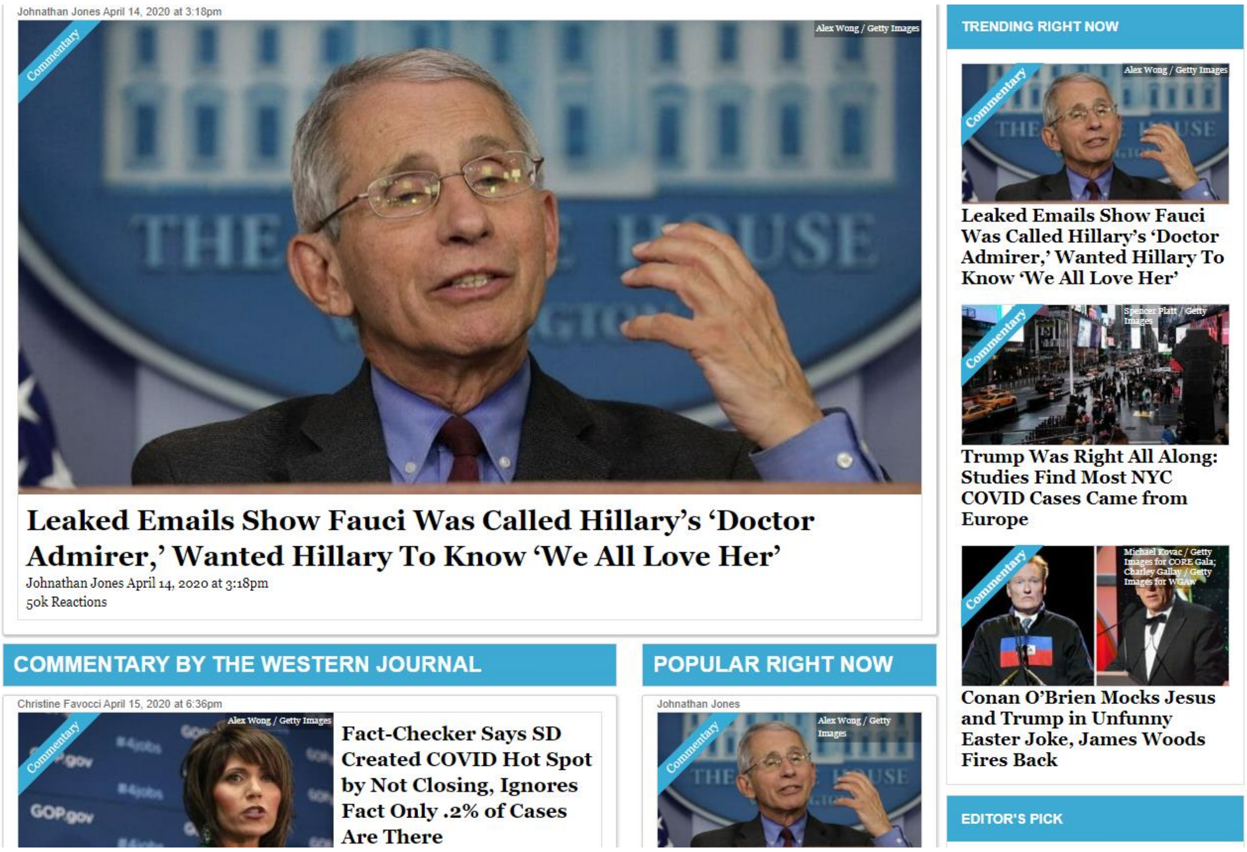


Figure 37 - conservativetribune (now westernjourmal) frontpage

Bear in mind that we do not have a copy of the design and statements of the webpage in 2016. As for classification, conservativetribune.com is both propaganda and biased or one sided, due to its hefty biased conservative lean as well as focus on discrediting of non-conservative with a questionable hold in reality.

Rank 15 is 100percentfedup.com. The webpage is, like several of the others, focuses on political stories. It has been in the focus of several fact-checking articles, which has turned out to be a mixture of fabricated political stories or conspiracy theories (Emery, 2017; Fichera, 2019; Sherman, 2018). The view of the webpage is shared by the literature (Faris et al., 2017). The webpage is by these reasons classified as fabricated, propaganda, conspiracy theories, and biased or one sided.

Rank 16 is pamelageller.com (which points to https://gellerreport.com/), a webpage that seemingly focuses on the news often related to Islam. This based on the info provided on the actual webpage below in quotation. Bear in mind that we do not have a copy of the design and statements of the webpage in 2016:

She is a foremost defender of the freedom of speech against attempts to force the West to accept Sharia blasphemy laws, and against Sharia self-censorship by Western media outlets (Geller, 2020)

The fact checks that have been performed at pamelageller.com classify the webpage as a spreader of fake news by fabricating rumors (FactCheck.org, 2020a; PolitiFact, 2020e; Snopes, 2020e). From the literature, the webpage also has gained the attention of being both classified as *rightwing* and *junk news* (Narayanan et al., 2018; Stein & Salime, 2015). As the webpage has a clear political tendency, couplet with several pieces of conspiracy theories, it is classified as fabricated, propaganda, conspiracy theories, and biased or one sided.

Summarizing the sixteen webpages above, table 16 has been produced. Take notice of this table focuses on a general level of the webpages which has been mentioned in the tweets (which explains the heavy use of multi-classification). The most prominent webpages (through the number of occurrences in the dataset) are being studied in section 4.3.1.

URL	Fake news category within URL							
	Fabricated	Propaganda	Conspiracy Theories	Hoaxes	Biased or one sided	Rumors	Clickbait	Satire
breitbart.com	x	x	x	x	x	x	x	x
express.co.uk						x	x	
www.rt.com		x	x		x			
infowars.com	x	x	x		x			
dailycaller.com	x	x			x			
truthfeed.com	x	x			x			
zerohedge.com	x	x	x		x			
jihadwatch.org	x		x		x			
sputniknews.com	x	x						
wnd.com	x	x	x		x			
barenakedislam.com	x	x	x		x			
shoebat.com	x		x			x	x	
therealstrategy.com						x		
conservativetribune.com		x			x			
100percentfedup.com	x	x	x		x			
pamelageller.com	x	x	x		x			
% of pages within this category	75%	75%	62.5%	6.25	75%	25%	18.75%	6.25

Table 16 - General fake news classification on webpages

The distribution of the classification is as follows:
Fabricated: 75%
Propaganda: 75%
Conspiracy Theories: 62.5%
Hoaxes: 6.25%
Biased or one sided: 75%
Rumors: 25%
Clickbait: 18.75%
Satire: 6.25%

4.3.1 Fake news URLs timeline

As done under section 4.2.1, to get a broader understanding of the characteristics around the propagation of the most shared fake news URLs, we studied the occurrence of the URLs (domains) in tweets throughout the year. Timeline in this context refers to occurrences of identified fake news webpages throughout the year, regardless of actor posting them in their tweets. The first figure (figure 38) shows the occurrences of the webpages individual, while the second (figure 39) and third figure (figure 40) shows them combined.

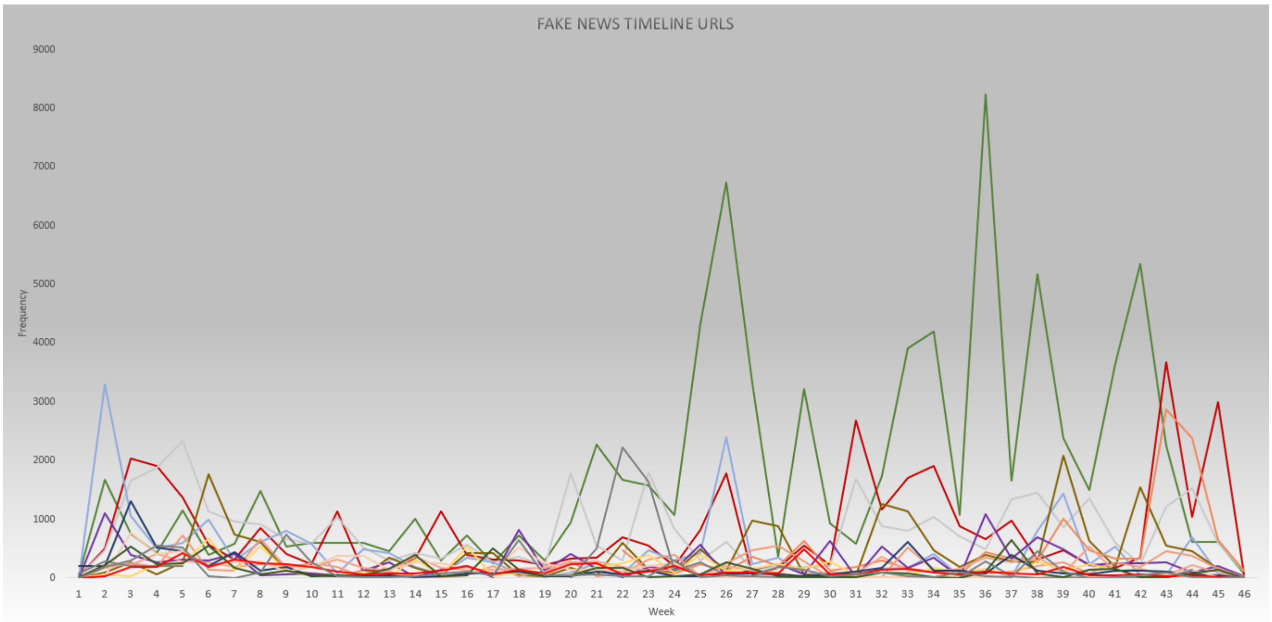


Figure 38 - Prevalence of the top 16 webpages categorized as fake news

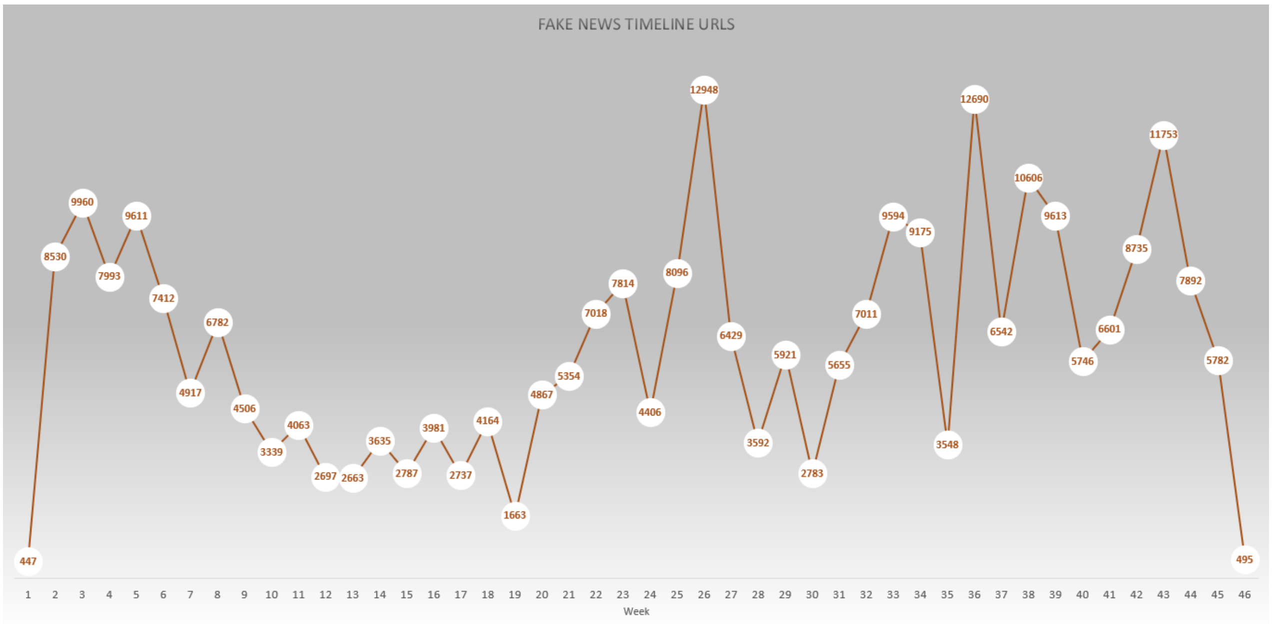


Figure 39 - Prevalence of the top 16 webpages categorized as fake news, combined

Converting the figures above to percentages of the total number of tweets containing these 16 webpages (282708), the figure 40 was created. Note that results from week 47 - 53 (no week 54 entries in the dataset) were removed, as these accounted for 0.06% (155) of the data and acted more as a clutter to the model than providing meaningful information (especially as the data collection from week 47 and onwards is incomplete in the dataset).

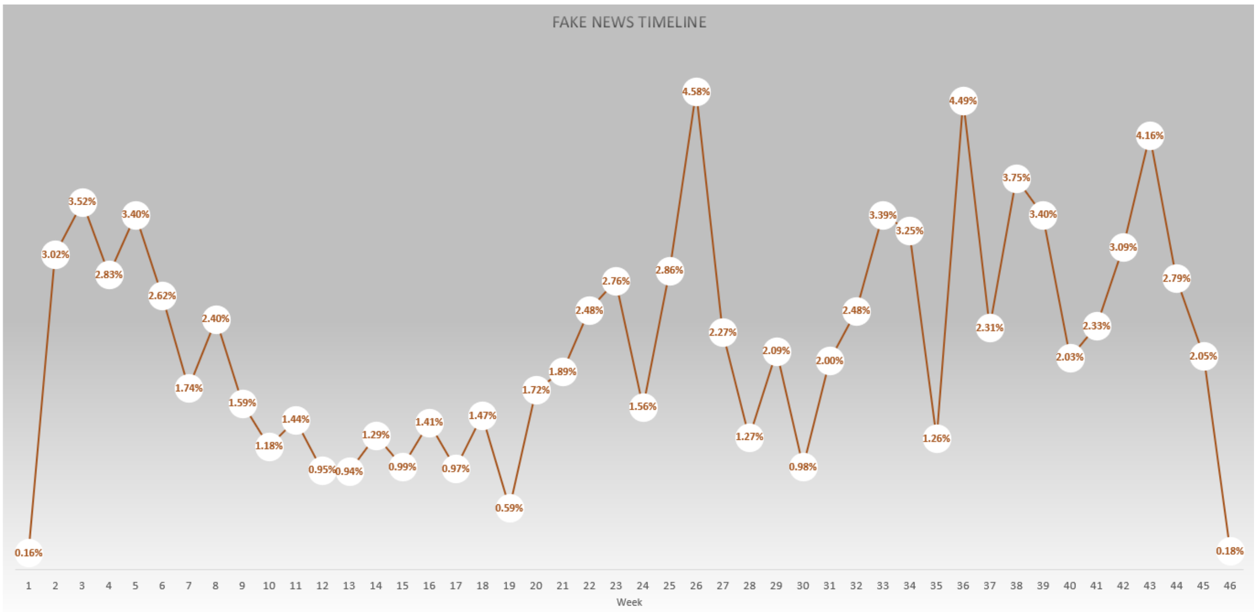


Figure 40 - Prevalence of the top 16 webpages categorized as fake news, percentages

The average is 2.17% propagation per week, while the median is 2.07%. Spikes in weeks 26, 36, and 43 (4.58%, 4.49%, 4.16%), although what is considered to be a spike in this model, is to be discussed later.

4.3.2 Fake news prevalence and concentration

To automate the detection of fake news, we used the compiled list of URLs classified as fake news. The list contains 1387 unique URLs. The list of fake news classified URLs was checked against the total dataset of tweets, 14.3 million tweets.

Of the total dataset of tweets, 14.3 million, there are 8.636.696 tweets with at least one URL. By joining the 8.636.696 tweets with URLs with the list of URLs classified as fake news, we found 435.099 tweets with URLs classified as fake news. Of all tweets containing URLs, 5,03% of them contained links to websites classified as fake news. Figure 41 shows the trends of fake news sharing. The x-axis is time in days, and the y-axis is the number of tweets that have shared fake news. The fraction of links to fake news varied day by day. Some periods we can see a high spike in content from fake news sources. The spikes are up to 4.5 times higher than on an average day. The average number of fake news content shared is 1351 per day (standard deviation of 877), or 36.258 per month (standard deviation of 18.938).

In the dataset, 113.700 unique users have shared or tweeted URLs that are classified as fake news, of the 435.099. The user that has shared most URLs has shared 645 tweets with URLs classified as fake news. The average user has shared 3,8 URLs in their tweets. The top 10% of users who shared fake news, shared 56% of all accounted for fake news shares.

56% of tweets with URLs classified as fake news came from 10 websites. If we look at the top 20 websites, the share has increased to 69%. In other words, only 20 websites stood for 69% of the URLs shared on Twitter that is classified as fake news.

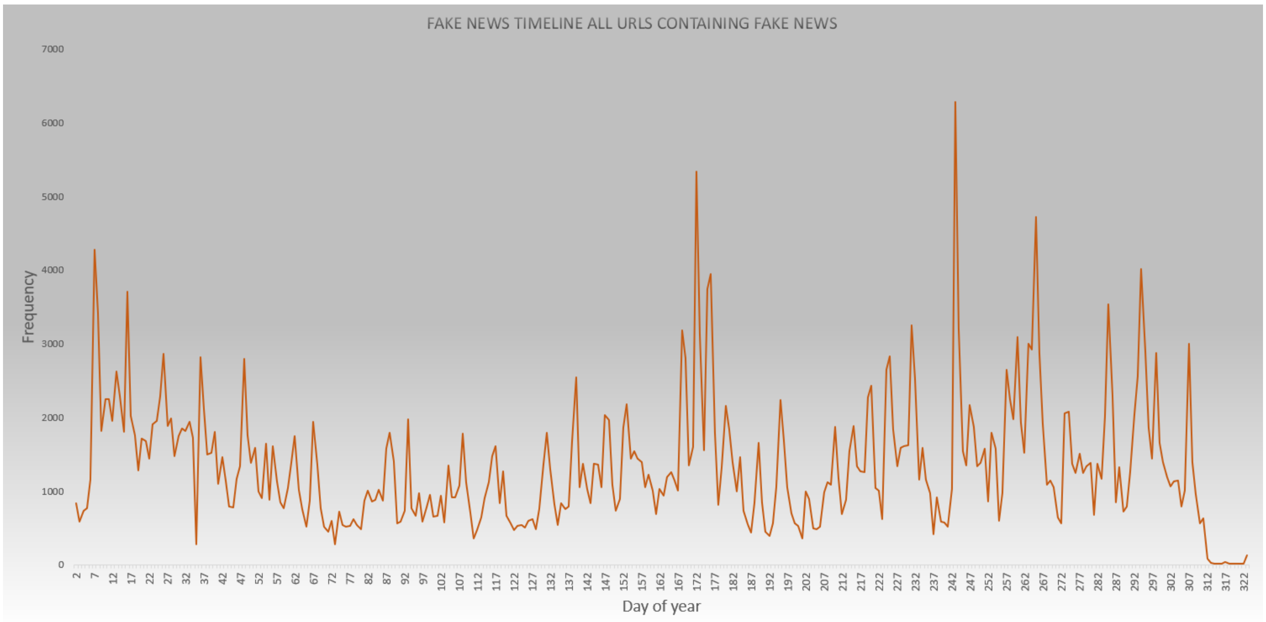


Figure 41 - Fake news timeline, all URLs. X-axis - time in days. Y-axis - number of tweets containing URL classified as fake news

4.3.3 Top fake news stories

As a result of fake news detection with the compiled list of fake news classified websites, we also found the most tweeted fake news stories. These fake news stories are individual URLs that point to single stories posted on external webpages. The websites identified are sorted by how many times they are tweeted (and retweeted, thus the total number of occurrences). These webpages are derived from utilizing the same fake news detection document as what was done in section 4.5.1

Rank	Fake news story URLs	Count
1	http://www.breitbart.com/2016-presidential-race/2016/06/23/cnns-tom-foreman-caught-lying-in-trump-refugees-fact-check/	5772
2	http://www.breitbart.com/2016-presidential-race/2016/08/29/bill-clinton-calls-for-rebuilding-detroit-with-syrian-refugees/	5359
3	http://www.breitbart.com/jerusalem/2016/08/18/hacked-memo-george-soros-behind-obama-decision-accept-100000-refugees-per-year/	3668
4	http://truthfeed.com/video-huma-abedin-on-hidden-cam-admits-hillary-will-open-u-s-borders-to-syrian-refugees/30772/	3514
5	http://www.express.co.uk/news/world/727862/Migrant-crisis-club-wielding-refugees-running-battles-Stalingrad-Metro-Paris	2598
6	http://dailycaller.com/2016/02/04/refugees-go-clubbing-in-russia-harass-girls-wake-up-in-hospital-the-next-morning/	2398
7	http://www.infowars.com/report-three-syrian-refugees-rape-little-girl-at-knifepoint-in-idaho/	2254
8	http://therealstrategy.com/came-saw-conquered-syrian-refugees-recruited-police-officers-bovaria/	2221
9	http://www.breitbart.com/big-government/2016/06/16/441-syrian-refugees-admitted-u-s-since-orlando-attack-dozens-fl/	2046
10	http://100percentfedup.com/?p=13671	1730
11	http://www.breitbart.com/big-government/2016/10/08/wikileaks-hillary-clinton-jordan-cant-vet-refugees-syria-jihadists-coming-along-legitimate-refugees/	1708
12	http://www.breitbart.com/big-government/2016/06/28/seven-refugees-active-tb-sent-idaho/	1683
13	http://www.breitbart.com/big-government/2016/05/17/22-resettled-refugees-minnesota-tested-positive-tuberculosis/	1611
14	http://waterfordwhispersnews.com/2016/06/24/thousands-of-british-refugees-make-dangerous-journey-across-the-irish-sea/	1571
15	http://www.express.co.uk/news/world/681614/Calais-migrants-refugees-Britain-UK-EU-referendum-Brexit-Euro-2016	1507
16	http://www.breitbart.com/2016-presidential-race/2016/10/11/trump-pushes-extreme-vetting-hillary-says-vetting-impossible/	1471
17	http://www.theamericanmirror.com/obama-summer-jobs-program-devotes-least-2-million-refugee-youth/	1407
18	http://dailycaller.com/2016/10/10/hillary-worried-about-jihadists-entering-with-refugees-in-private-speech/	1395
19	https://wikileaks.org/imf-internal-20160319/	1344
20	http://toprightnews.com/dalai-lama-says-shocking-thing-about-muslim-refugees-furious-obama-attacks-him/	1320

Table 18 - Top fake news stories

Eight of the 20 (40%) most shared fake news stories relate to Breitbart.com. One also sees that the list contains new URLs not previously covered under section 4.3. waterfordwhispernews.com, theamericanmirror.com, wikileaks.org, and toprightnews.com entered the list for single shared stories. Based on observation of the name of the stories, a clear (negative) focus on refugee-centered narratives emerges. The stories are not further explored.

4.4 Summarizing findings

To illustrate the findings from the sections above, figure 43 was constructed with the core findings listed up as a flow diagram. The arrows indicate the flow and dependencies of the findings. Example: *Fake news detection & classification in most liked tweets* were performed after the contextual *most liked tweets*, and the result from the latter was used to progress to the next activity (the former). Color of the boxes refers to which RQ the findings relate to.

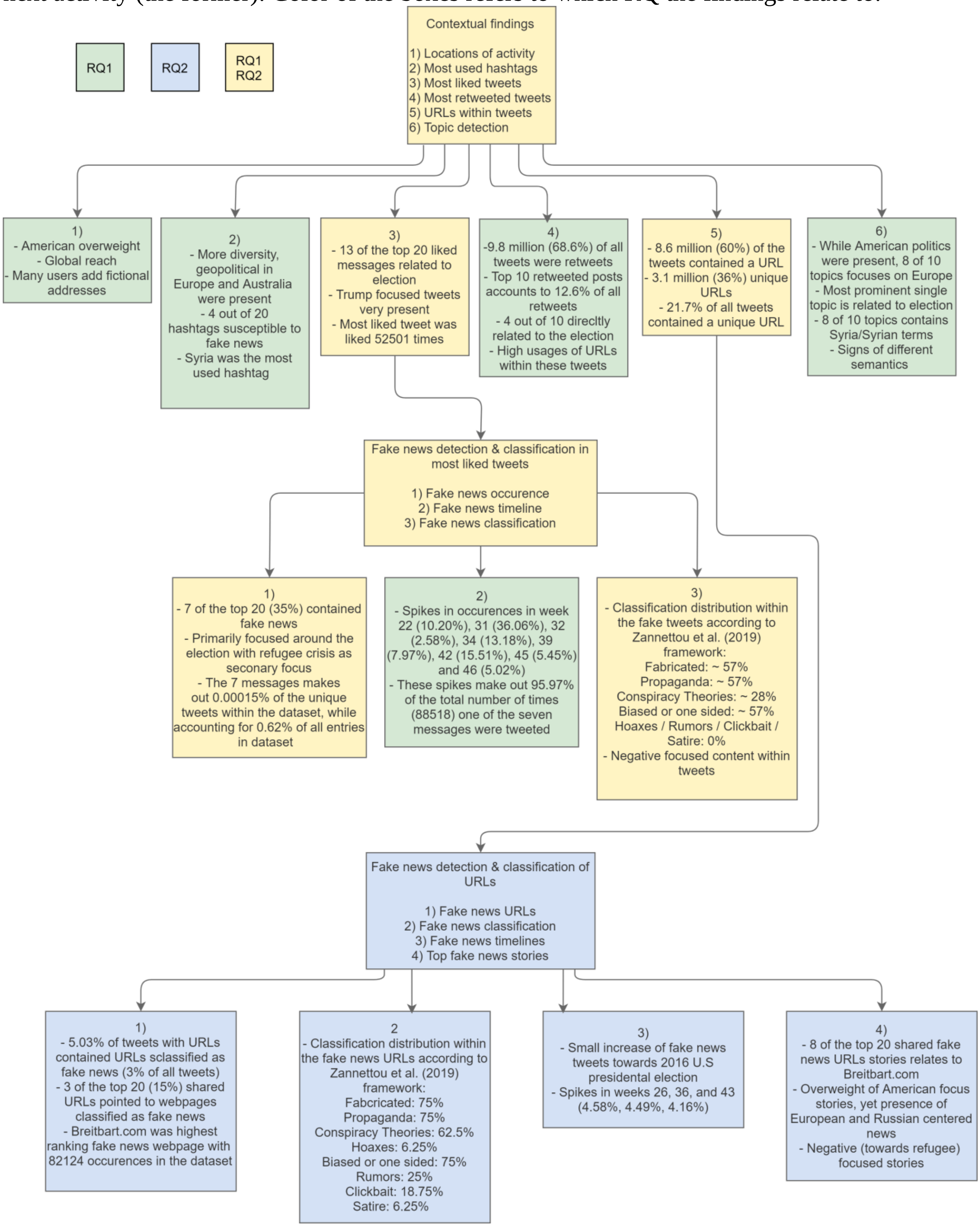


Figure 42 - Summarizing findings model

5 Discussion

The purpose of this section is to discuss the findings derived from the analyzes performed on the dataset against the literature collected and utilized in this study, moving towards a conclusion.

From presenting the challenge imposed by fake news on social media, the following research questions were formed: *What were the most liked and shared tweets in a dataset related to the 2016 refugee crisis, and what were the characteristics of tweets?* And *Using open-source services and a mixture of manual and automatic fake news detection, is it possible to find the presence of fake news in the tweets, and if so, what type of fake news is present?* Through the literature review conducted in the previous semester in combination with this study, we identified that there was a need for bridging fake news with social media and then set the context to the refugee crisis of 2016. The context did not derive from the literature alone. Rather it was based on the dataset available for the group to utilize in a combination of other factors (see the introduction, motivation for choosing this topic). As a result of this, table 5 was created, which focused on the findings from the literature review.

Judging the literature collected, different points of focus were present. While the American presidential election had, to some degree, been studied, there were mixtures of different geopolitical situations covered in the literature, such as the Ukrainian election, Brexit, Gulf Crisis, Fukushima disaster, the hurricanes Sandy, Harvey, Irma, and Maria - just to mention some. This highlights further how the literature uses crisis, emergency, and disasters intangible and that the literature in general at this part still is a bit young (more on this in section 5.1). This not to be mistaken as a sign of weakness - the research conducted in the literature collected is of high quality, yet, it is still in its early stages as there are many different focus areas and small degrees after trailing previous work.

To be able to answer the research questions, we focused the efforts on identifying several contextual areas that could aid in answering these questions. The contextual information had three purposes: 1) Give a preliminary understanding of the dataset at hand, 2) To be able to find the most liked and shared content caught in the dataset (RQ1) and, 3) Be able to give a broader insight into some of the characteristics of the content classified as fake (RQ1 and RQ2).
We wanted to be able to tell why something was fake news, as the word *fake news* holds great ambiguousness (Vosoughi et al., 2018). Pointing to one of the findings in isolation to say whether it answered the research questions or not is too narrow in this context, rather the sum of the contextual findings helps to answer these. The contextual findings were also needed to be able to conduct the main part of this study (RQ2), such as identifying the most liked tweets and most shared URLs within tweets.
In the following sections, we will cover and discuss the findings in-details and relate these to the literature identified and utilized in the study, coupled with discussions on how the results derived from the analyzes compares to the literature identified.

As we will discuss further in the following sections, the findings show an overweight in American political focus caught in the findings.
As of why this was the case, there are several possible factors that, in isolation and coupled together, might explain some of this phenomenon. One is that due to the extensive amount of fake news present in the American presidential election (Budak, 2019; A. S. Ross & Rivers, 2018), often related to political spamming (Al-Rawi et al., 2019; Bastos & Mercea, 2019), means that this extensive traffic affects the dataset and skews the balance of the data. A other factor is that the refugee crisis was adopted into American politics (as opposed to being a focus area of Europe in isolation), which means the ties to the election (and thus the spamming) were possibly

tight (Bhatia & Jenks, 2018; Galán-García, 2017). Usage of Twitter as a social media platform around the world might also affect this, and Statista (2020) summarized the adoption of Twitter, as shown in figure 43.

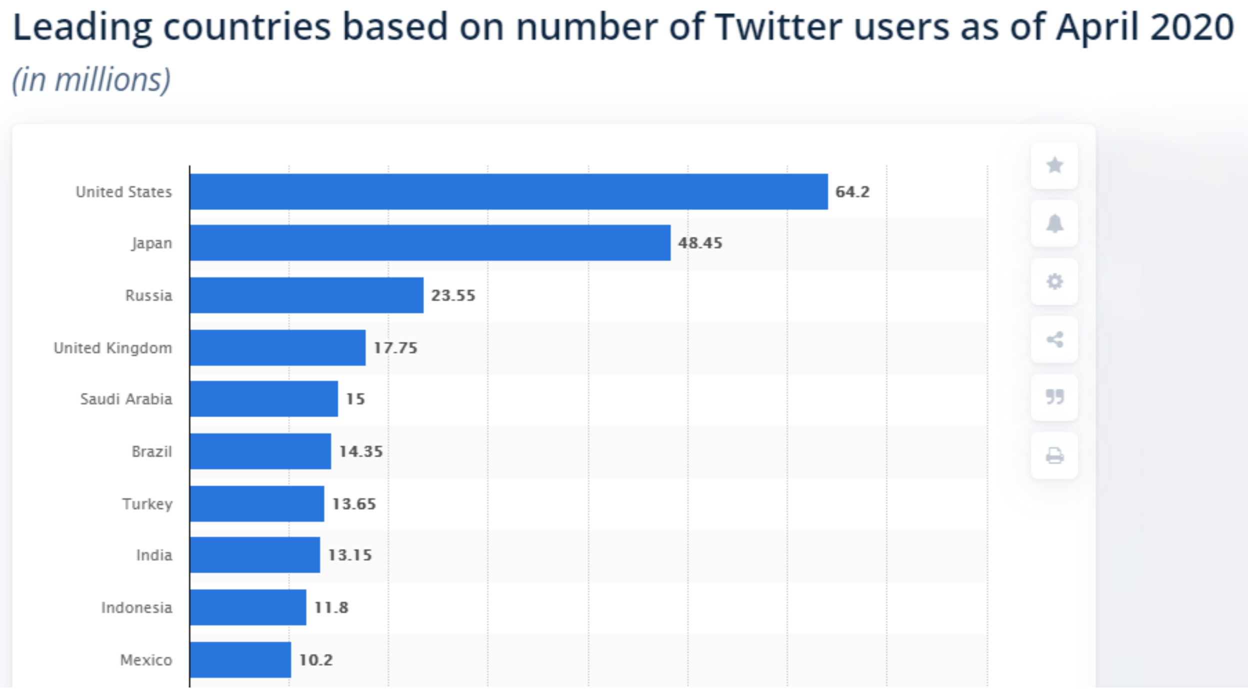


Figure 43 - Twitter usage around the world (Clement, 2020)

Ultimately, this study was exploratory in its nature and focused on exploring the dataset aiming (at a general level) to 1) give insights on the public discussion, "as was," 2) the fake news present, and 3) the characteristics of such content. With such an approach, it was decided not to clean and remove much of the data, as it would then violate the intention of looking at the dataset with fresh eyes and present the results derived from this endeavor. The knowledge of American overweight in the content presented is a result derived from the analyzes and not as a result of weakness in the method or approach.

5.1 Theoretical Contribution

As the literature related to fake news in social media is still young in academia, there is a need for more knowledge on the subject in terms of academic usage (Jang et al., 2018, p. 112). Additionally, the context of the crisis sees even more signs of immaturity (see the section above). This is something the group shares view on, based on the following: Of those results that delved into fake news on social media with a context of crisis (see table 5), these were not directly comparable to this thesis, as neither focused directly on the refugee crisis of 2016. Furthermore, there is great ambiguousness of the terminologies of both fake news (Budak, 2019; Burbach et al., 2019; Campan et al., 2018; Del-Fresno-garcía & Manfredi-Sánchez, 2018), as well as the vague borders between crisis, emergency, and disaster, as highlighted under section 3.5.4. In appendix E, figure 53, it is also included contextual information in relation to the concept matrix to illustrate the focus areas of the literature identified in the literature review.

We saw that the terminologies were used differently based on those using it, which is especially visible in the fake news detection of URLs that will be covered in 5.2.5. Many of the entries in the lists we combined used terminologies such as bias, hate, rumor, satire, conspiracy theory, clickbait. In reality, the webpages were advocating fake news in one of its forms when the framework by Zannettou et al. (2019) was applied to them. It also empathizes the wording from

Gelfert in the introduction; that fake news is misleading by *design* (Gelfert, 2018), and webpages and content in general that falls within one of the categories should be classified accordingly as they intend to mislead the reader one way or another, regardless of the severity.

The study group shares the view by Vosoughi et al. that fake news (in isolation) is a term not suitable for academia due to its ambiguousness (Vosoughi et al., 2018), and that further sub-classifying the term accordingly to a framework such as the one provided by Zannettou et al. (2019) is crucial for clarification. As a result, we believe that the thesis derived from the study addresses some of the gaps in the literature by providing a novel study in a new context (the refugee crisis), utilizing a little-used research approach (manual fake news detection), as well as clarification on the terminology used (using Zannettou et al. framework). It is worth noting that the manual approach to detect fake news is a method suggested to be utilized within the research field on this matter (Zhou & Zafarani, 2018, pp. 7-8).

5.2 Contextual findings

As the research approach is exploratory, so did the analyzes, and the result reflects this. Since it was the first time the dataset had been used, there was a need to understand more about what the dataset contained. As mentioned at the end of section 5, to be able to address the RQs, we needed several contextual findings. Thus, there are six contextual findings within this study that helps to give insights into the data collected. In the following subparagraph, the specific findings will be discussed and compared against the previously collected literature to see similarities and new findings.

5.2.1 Locations of activity

We saw a hefty overweight of American-based users caught in the dataset. It is important to notice that while Twitter requires users to add a location, it is a free text area, which means users can plot in whatever address they like (with whatever granularity) as well as changing this as many times they would like.
Previous research has suggested that malicious actors (such as trolls, bots, and actors with various agendas) seem to target countries outside their origination in an attempt to increase the perceived credibility of the accounts (Bastos & Mercea, 2019, p. 44; Zannettou, Sirivianos, Caulfield, et al., 2019, p. 220). Manipulation of location is also seen in other contexts than elections, with similar aims, which suggests it is problematic to say anything conclusive about findings related to locations on Twitter (Hughes & Palen, 2009, p. 253; Rajdev & Lee, 2016, p. 17). However, when looking at the findings from both most liked messages and most shared URLs, we believe that while the location of activity cannot be looked at in isolation, it helps to confirm the overweight of American focuses content in the dataset when looking at the findings together. It would be of interest to check whether the dataset contains known actors (such as state-sponsored trolls and bots), to see the influence these held on the data collected and ultimately see the similarities from previous research on this subject (Zannettou et al., 2017; Zannettou, Caulfield, et al., 2019; Zannettou, Sirivianos, Caulfield, et al., 2019). However, scope constraints, DPIA agreements, and focus of the study meant this was not a viable option to undertake.

5.2.2 Most used hashtags in tweets

Hashtags are susceptible to being hijacked by bots and malicious actors (same as the section above), and it is therefore of interest knowing the most mentioned ones caught in the dataset (Bastos & Mercea, 2019; Jones, 2019).

As seen in Table 12, we summarized the 20 top mentioned hashtags. There are, however, details that need to be taken into consideration when interpreting the results. The hashtags that were used to collect the tweets were removed (see section 3.8.1), meaning hashtags that very well might have been the most occurring hashtag were removed from the analysis. The reasoning behind this was because these hashtags would be occurring in every tweet, creating too much noise in the analysis.

As opposed to what was seen in 4.1.1, the hashtags were not solely revolving around American politics. Sure enough, #MAGA was the 8^{th} most used hashtag, yet there are other themes present here. Firstly, we saw that synonyms to refugees were used, through the usage of tags such as #migrants (7^{th} spot), #immigration (19^{th} spot), and #refugeesGR (15^{th} spot).
The hashtag #EU (3rd spot) and #migrantcrisis (2th spot) suggests the results derived from studying the hashtags seem to be more balanced (in terms of America versus Europe/World). It was a surprise to see that #Brexit was not present in the top list of hashtags, as it was believed the hashtag hijacking would mean these tags would gain more attention, as shown in previous research (Bastos & Mercea, 2019).

We found that the hashtags #tcot (4^{th} spot), #maga (8^{th} spot) #pjnet (17^{th} spot), and ccot (20^{th} spot) were present in the dataset, which are hashtags shown to be susceptible being used in to spread fake news (Zannettou, Caulfield, et al., 2019, p. 221; Zannettou, Sirivianos, Caulfield, et al., 2019, p. 358). The other hashtags are not explicitly studied in-details in the literature collected; thus, the following utilized various external sources to try to explain their meaning.

> #auspol was related to the Australian federal elections (Semmens, Moon, Bolliet, Amarasekara, & McKinnon, 2016).
> #UN4RefugeesMigrants was related to political usage concerning Belgium politics (Van Leuven, Deprez, Joye, & Ongenaert, 2018).
> #UNHCR was related to the organization United Nations High Commissioner for Refugees (Loescher, 2001).
> #UNGA was related to organization United Nations General Assembly (Brazys & Panke, 2017).
> #refugeesGR was a collective term related to Greece refugees (Twitter, 2020d).
> #raperefugees was related to an anti-refugee slogan (Gallego, Gualda, & Rebollo, 2017).

Hashtags such as #syrian, #migrantcrisis, #eu, #migrants, #news, #UN, and #immigration were not further looked into, as these were either too broad/general or needed no further explanation (such as UN). It is, however, worth noting that the three hashtags #UN, #UNGA, #UNHCR tells something about the topics of discussion in this period.

5.2.3 Most liked tweets

As this study heavily focused on what people wrote and discussed on Twitter related to the refugee crisis of 2016, a natural step was to identify the most liked tweets. Similar research has been conducted before, although the focus of efforts was at the American presidential election in isolation and random selection of tweets in-question (Pal & Chua, 2019, p. 271).
While the group had assumptions beforehand the analyzes that the American presidential election would be present, the sheer number of tweets related to this acted as a surprise. As seen in Table 13 that lists the 20 most liked tweets in the dataset, most of these are focused on American politics. Related to the assumptions from the group, it was believed that geopolitical situations such as Brexit and the European Refugee Crisis (at a general level and not tied to American politics) would be more present, yet, compared to the overweight of American politics little

traces were found - at least in the top 20 list. Of the 20 tweets identified, these have previously not been covered or discussed by the literature collected, although method choices and collection might explain this differential (Pal & Chua, 2019).
The frequency, e.g., *like* counter of the tweets, visualize a skewed balance. The most liked tweet was liked 52501 times, while the 20th most liked tweet was liked 6463. The dataset consists of 4.5 million unique tweets (not counting retweets and duplicates), meaning the median number of likes is significantly lower.

In terms of the literature, these findings are unique, as previous work has focused its effort on the actors, general topic analysis of collected data, or a random selection of tweets studied (Budak, 2019; Pal & Chua, 2019; Zannettou et al., 2017). It is - by the opinion of the group and when compared to previous literature, not a weakness that the results mostly focused on the American presidential election. Also, with more time, it could have been interesting to exclude all the results that focus on the election in isolation, yet, scope constraints made this not a viable option for the study. Exclusion criteria, workflows as well as sound reasoning based on both the research approach in combination with a highlight of literature to argue the case (of why removal of the American election content) would have to be presented at sufficient quality.

5.2.4 Most retweeted tweets

The analyzes have shown that the majority of the tweets caught in the dataset are not unique. 68.6% of all tweets were retweets of other people or organizations tweets. This result can be skewed further: the top 10 retweeted tweets, which amount to 0.00006% of all tweets in the dataset, amounts to 8.64% of all collected tweets when looking at the occurrences of these. High amounts of spam, through the usage of retweets (which reduces the number of unique tweets), is not a new phenomenon at Twitter (Al-Rawi et al., 2019; Bastos & Mercea, 2019). Although it is important to keep in mind that while these studies focused on investigating the reach and effect of bots, the results are appliable to this context as we did not filter out actors from the dataset. It stands to believe that caught in the dataset, there are actors, such as bots, that has affected the debate on Twitter (and thus data collected). Previous research has shown values of as low as 2-4% bots in a population within a discussion can affect this discussion (B. Ross et al., 2019). Nevertheless, more research on the actual actors caught in the dataset is needed to determine the reach and prevalence.

Of the top ten retweeted tweets four related to American politics in general (3 addressing Trump directly), three concerning attitudes toward refugee, one promoting to win a phone, one was related to sports events, and lastly, one miscellaneous. We did not delve into these in particular depth, but preliminary indications of these tweets showed only 2 addressing European politics, which empathizes the American overweight seen previously. There was also a clear usage of URLs and otherwise external material in these tweets. The motivation for such might be to alter the perception of credibility in the tweets (Aigner et al., 2017), especially as it also is linked to both political statements and refugee related statements.

5.2.5 URLs within tweets

It has in previous research been shown that a vast amount of fake news is spread through the usage of links to external material (URLs) in Twitter tweets (Bovet & Makse, 2019). Therefore, a natural step of understanding more of the dataset was to identify the extent of URLs in the dataset, followed by assessing the degree (what classification) of fake news within these URLs. Assessing the fake news occurrence is done in section 5.4, leaving this section to focus on the general discussion relating to URL occurrences in the dataset.

By counting the occurrences of URLs within tweets, we were able to find that 60% (8.6 million) of all the tweets in the dataset contained a URL. Although this is a high amount, the high frequency of URLs in tweets is not uncommon (Aigner et al., 2017, p. 257; Rosa, Shah, Lin, Gershman, & Frederking, 2011, p. 4). Furthermore, 36% of the tweets shared were unique. The number of unique URLs was seemingly high, though compared to previous research, similar high numbers have been seen (Zannettou et al., 2017, p. 407). When looking at the entire dataset, we saw that approximately one out of five (21.7%) tweets contained a unique URL. This finding was interesting, as the group believed due to the massive retweeting and liking of certain messages, it would also mean there were less unique URLs present in the dataset.

Table 16 shows that eight of the ten most shared URLs are pointing to Twitter related content, while 5 of these are URLs pointing solely to Twitter accounts. These occurrences happened in all types of activity on Twitter, either through a retweet with an explicit reference to an account or a new tweet with reference to the profiles. With more time available, it would be of interest to study how many of the URLs shared points to Twitter material and profiles, as opposed to external websites.
We choose to pseudonymize the results in this table, to prevent the usernames from being visible. The argumentation by the group was that while public profiles by prominent actors such as politicians should expect to be a target for study, the same does not apply to individuals outside the light of the public. This also falls in line with the DPIA agreement, where usernames should - if within reasonable measures, be removed. Furthermore, when people retweet or tag people or organizations in their tweets, their chosen screen name - not username, is shown (which also explains why screen names are visible under sections such as most liked tweets section).

5.2.6 Topic detection

To understand more about the different public discussion related to the refugee crisis in the data collected, we performed topic detection with the same methodology as previous research on the field (Zannettou, Caulfield, et al., 2019; Zannettou, Sirivianos, Caulfield, et al., 2019).
The strength of this type of analysis (workflow) is that its focus is on the related discussions, which meant it showed a more granular result when the American presidential election was grouped into its own topic (primary topic 3 in our LDA analysis as seen in table 17) rather than overflowing into every other interconnected discussion and topic.

Studying table 17 in the combination of what was seen in the other contextual findings, the overweight of the American presidential election focused content was ever-present. This finding is in line with previous research that has shown the election to be a very focused and discussed topic on Twitter (Budak, 2019, p. 143) - although worth noting that this particular paper focused on the election in isolation. Compared to other studies that did not exclusively delve into the American Election (in isolation), there were some similarities in the topics, such as topics containing the terms Trump and Hillary, yet ultimately the topics identified in our research differed a bit from these (Zannettou, Caulfield, et al., 2019, p. 360; Zannettou, Sirivianos, Caulfield, et al., 2019, p. 223). The differential was likely due to different focus, dataset, methods, and/or software used (an example of this is how Zannettou et al. utilizes Hawkes processes in their research and different data collecting methods).

As was seen in section 5.2.2 related to the hashtags, and due to the increased granularity, we were able to identify additional topics outside the American focused content. Of the ten topics, two addressed American politics, while the remaining eight seemingly focused on Europe and Australia. In the center of the topics was *syria/syrian*, with 8 of the topics containing this term as one of the ten highest occurring terms within this topic. It would be of great interest to delve

deeper into the topics to look at the semantics caught within these topics, to see whether the results derived from this dataset falls in-line with similar research surrounding Syrian refugees (Nied et al., 2017), as it was hard judging these by looking at the LDA results in isolation. For example, topic 1 contains the terms *syrian, help, refugee, world, new, support, work, stand, withrefugees,* and *today.*
In contrast, topic 8 contains the terms *syrian, australia, trump, canada, skittles, auspol, nauru, people, just,* and *refugee.* The first do by the glance of it indicates more positive semantics, while the second seems more negative. The keyword here is skittles, which was a negative remark Trump made towards Syrian refugees (BBC, 2016a). Before these have been studied in greater detail, it is hard to judge the semantics and, ultimately, the susceptibility of fake news within the topics. Nevertheless, it opens up paths for future studies as we now know that the dataset contains different geopolitical situations that can be investigated further as they have not in previous studies been discussed in lengths.

Drawing on the parallels to the hashtags in section 5.2.2, we saw that the term syria/syrian were prominent topics. This finding is in line with prior research that has shown that topics addressing Syria seems susceptive in the creation of echo chambers and, ultimately, fake news in one of its forms (Kostakos, Nykanen, Martinviita, Pandya, & Oussalah, 2018; B. Ross et al., 2018; Zannettou, Sirivianos, Blackburn, et al., 2019).

5.3 Fake News detection and classification in most liked tweets

As what was done under section 4.2, extensive work was conducted at manually classification of the claims within the most liked tweets. In term of approach - manual and not automated fake news detection, this is suggested as an approach in need of more research (Zhou & Zafarani, 2018, pp. 7-8). Thus, this approach was chosen (both at a methodical level as well as analyze/result-oriented level) instead of strictly automated detection and/or of utilizing machine learning and similar options that have been used by the literature prior (Bastos & Mercea, 2019; M. Del Vicario et al., 2018; Rajdev & Lee, 2016; Zannettou, Sirivianos, Blackburn, et al., 2019, p. 16).

Much of the discussion around the fake news classification is performed within section 4.2 and is not repeated here. This deviation from the usual structure of a thesis is due to the approach, as findings were based on arguments utilizing fact-checking webpages and literature for classification, which makes sense to discuss where applied. We refer to this section for general discussion related to the aspects around classification.
One interesting element is that the seven (out of 20) tweets identified to contain fake news in one of its forms, is the sheer amount of occurrences these tweets holds in the data collected. These tweets, while accounting for just 0.00015% of the unique tweets within the dataset, amount to a total of 0.62% of all tweets in the dataset. High amounts of fake news in Twitter messages are results in-line with previous research (Bovet & Makse, 2019; Budak, 2019; Rajdev & Lee, 2016). The fake news tweets identified are solely focused around American politics, and most of them were in the end classified as fabricated, propaganda or biased or one sided (and usually a combination of these three as tweets could receive multiple classifications due to the various claims within). An alternative approach here would have been to deconstruct the tweets and focus solely on the claims within the tweets one by one to reduce the multi-classification. However, the group argues that the RQs, in combination with streamlines in the analyzes, speak for such an approach (the same approach was used in URL classification, where the main focus was at an overall and not specific news item level).

5.3.1 Fake news tweets timeline and properties

As the tweets focused around the American presidential election, so did the propagation pattern reflects this. The earliest sign of one of the seven tweets occurred in week 22 (end of May 2016), and last entry in week 45 (mid-November 2016). The election took place in week 45 (November 8). As the dataset is incomplete for November and December, the following sentence is speculations that would need further research to verify its credibility. Nevertheless, prior research saw a buildup followed by a drastic drop in fake news propagation from week 45 in 2016, which suggests there might be some hold in this speculation (Budak, 2019; Grinberg et al., 2019, p. 2; Pal & Chua, 2019).

Week 22, 31, 32, 34, 39, 42, 45, and 46 saw spikes in the propagation in the dataset, and 95.97% of the total propagation of the tweets happened within these weeks. The propagation pattern for each tweet is covered in their individual paragraphs below, with a small discussion of the properties of the tweets.

The first tweet is worded the following:

> *Hillary's refusal to mention Radical Islam, as she pushes a 550% increase in refugees, is more proof that she is unfit to lead the country.*

Its propagation pattern shows that over 90% of its spread happened the first week after its creation. Previous studies show that the propagation pattern of fake news tends to increase over time, which we saw here as well, but comparing these results with such studies has some limitation due to 1) the timeframe in the reference study was shorter and 2) our study does not study the alternation of the tweets over time, meaning we do not know how the wording of the tweets evolved over time (Jang et al., 2018; Pal & Chua, 2019). *Radical Islam*, in combination with *increase in refugees* plays on the negative emotions of the reader, which saw similarities to previous research (Vosoughi et al., 2018).

While the usernames of those creating the tweets have been removed from the study, it stands to believe that the tweet originates from Trump. This based on three observations. 1) The message is highly liked and shared across the network, which means it most likely originates from an author with an extensive network of followers. Furthermore, his account has previously been noted to have an extensive network (cluster) related to it (Bovet & Makse, 2019, p. 4). 2) It targets Hillary specific and the word cues such as "unfit to lead," which was a phrase Trump used throughout the election campaign. 3) The claims of both *refusal to mention Radical Islam* and *a 550% increase in refugees* originate from Trump (Sherman, 2016). These three arguments are reused as argumentation when discussing the following tweets, though not written explicitly to make this section more readable.

The second tweet is worded the following:

> *A vote for Clinton-Kaine is a vote for TPP, NAFTA, high taxes, radical regulation, and massive influx of refugees.*

As was seen in the first tweet, over 90% of the occurrence of this exact worded tweet happened the first week after creation.

This tweet also highlights how the extended classification of fake news in our study affects the classification of such messages. Its message is not entirely fabricated with no grounds in reality, but it is a simplification of the actual situation and very biased in its presentation (Jr., 2016). As was seen in the first tweet, it spread quickly, and couplet with its wording, we utilize the three same arguments in this one to assume it originates from Trump. Further strengthening this assumption is the fact that Trump did tweet an exact copy of this message around the time of the data collection (Trump, 2016c). The properties of the tweet (as one will see in the following tweets studied) are playing on negative emotions of the reader, such as increased economic costs (high taxes), lack of freedom (radical regulations), and destabilization of society (massive influx of refugees). These types of properties are common surrounding fake news in social media (Vosoughi et al., 2018).

The third tweet is worded the following:

> *ISIS has infiltrated countries all over Europe by posing as refugees, and @HillaryClinton will allow it to happend here too! #BigLeagueTruth https://t.co/U5hDdIc4rC*

Its propagation pattern matches that of the two tweets above; however, it has a much shorter lifecycle (four weeks). The author of the tweet is unknown, yet the hashtag at the end reveals that whoever it was, this actor was related to the BigLeagueTruth group who posts and retweets pro-trump content on social media (Jamieson, 2016).
Around the peak time of the propagation, Trump tweeted an almost identical message (Trump, 2016a), which speaks for its origination to be from him as well. The message plays at the fear of the reader (ISIS infiltrating Europe with the same fate for America if Hillary would win the election), and its properties are in line with previous research (Vosoughi et al., 2018).

The fourth tweet is worded the following:

> *Crooked Hillary wants a radical 500% increase in Syrian refugees. We can't allow this. Time to get smart and protect America!*

The propagation pattern shares similarities to those from before, and it had one of the most extended life cycles. Again we believe this tweet originates from trump, of the same reasons as before in combination that he posted an identical tweet around the time of the data collection (Trump, 2016b). Its wording cue "protect America" playing at the negative emotions of the reader through the fear of what Hillary would mean for the security of the nation in this context. These properties are in line with previous literature (Vosoughi et al., 2018).

The fifth tweet is worded the following:

> *Clinton refugee plan could bring in 620 000 refugees in first term at lifetime cost of over $400 billion. https://t.co/COZQNt6KVs*

The propagation pattern for this particular tweet is unique when compared to the other tweets studied. The primary propagation (79.40%) happened in week 34, 8 weeks after its first occurrence in our dataset. Furthermore, it had the most extended lifecycle of the tweets studied, with a total of 19 weeks of diffusion (27 - 46). As discussed briefly under findings, its first conception is hard to pinpoint, yet, that it originates from Trump's Twitter or (in combination with the) add campaign seems plausible. Further strengthening this is the fact that Trump tweeted an identical tweet at least with one occasion (Trump, 2016d), in combination with the arguments utilized in the first tweet in this section. As for the properties of this tweet, one sees similarities to the other tweets, where it plays on negative emotions and fearmongering of the reader (vast numbers of refugees (destabilization of society) tied to an extremely high economical cost).

The sixth tweet is worded the following:

(...) Refugees from Syria over 10k plus more coming. Lots of young males, poorly vetted.

The message is snippet (…) to remove some Twitter screen names. Its propagation pattern was the same as the others (minus the tweet at the fifth position) and had a shorter life cycle than the other tweets (only beat by third position tweet with its four weeks). Its properties are around fearmongering, which in this case, is the insufficient vetting process of refugees. As seen in the literature, tieing refugees (at a general level) to disruptive events of society is common (Aigner et al., 2017, p. 297; Haug, 2019, p. 1), which makes our findings supporting this connection.

The seventh tweet is worded the following:

(...) We want a president that puts Americans first - not illegal aliens, not refugees. https://t.co/xNzraMVx6D

The first part of the tweet is snipped to remove the screen name tagged in the tweet. Of the tweet studied in-detail, this one had a unique propagation pattern. Its life cycle was only two weeks, and it was almost entirely even distributed (51.95% the first week). The author of the tweet is unknown, while the URL at the end of the tweet refers to a Fox News tweet (FoxNews, 2016). Its properties were at playing on the fear of illegal refugees (here referred to as aliens), and the negative effect that would impose on the American society.
The tweet was previously in this study classified as biased or one sided, since there were no arguments (balance) presented by Fox News in this particular tweet. While Fox News is seen as a traditional media outlet, it has previously been classified as a right-wing (named *right* in the original study) news outlet (Bovet & Makse, 2019, p. 6).
An interesting note here is that studies show that Fox News actively uses the hashtag *fakenews* on Twitter (Al-Rawi et al., 2019). This helps illustrate the issue the group shares with the literature in terms of the term; that it is used by various actors to fit their purpose and narrative (Vosoughi et al., 2018), and in need of further classification if it is to be continued to be used in academia.

5.4 Fake News detection and classification of URLs

In terms of manual fake news detection and classification of webpages, there is a novelty in the findings and approach. By finding the most shared URLs, validating them against the most extensive list to-date of fake news webpages/actors (1387 unique entries in total) and then further classifying them accordingly to the framework by Zannettou et al. (2019), we have been able to create a snapshot of some of the fake news actors present in the public discussion on Twitter related to the refugee crisis of 2016. Previous work at detection and classification has utilized smaller lists of fake news webpages (1001 entries), in combination with more strict inclusion criteria when deciding whether a webpage is listed as containing fake news (Bovet & Makse, 2019, p. 11). It is worth noting here (as further elaborated on under limitations), that the literature review had to be limited due to scope constraints, and even with a foundation of 42 articles, the possibility of missing out on relevant research in terms of fake news detection was present.

Judging the results, most of the URLs focused on the American presidential election, with a tendency towards conservative and right-leaning. Several of the pages are previously covered by the literature, which shares the same opinion about them (Bovet & Makse, 2019; Jang et al., 2018; Murungi et al., 2018). For more comparison towards the literature, see section 4.3

While network mapping was not feasible due to DPIA regulations (which meant removable of usernames), identifying and classification of webpages that creates and shares fake news helps build up the understanding of different web communities - as suggested by the literature to be a necessary step in advancing the research field (Zannettou, Sirivianos, Blackburn, et al., 2019, p. 30). Arguments can be made that the list of fake news webpages contains entries that previously are classified as biased and/or politically focused, which at the time of creation was not thought of or viewed as fake news. It is in the framework utilized in our thesis classified as fake news regardless of severity and original classification. In the tweets that have been manually checked, the previous classifications (in the webpage list checked against in the dataset) were not taken into consideration, as these were checked against Politifact, Snopes, Factcheck.org, and general literature (not limited to that collected in the literature review) to form a new classification accordingly to the framework used. This is a bit different from similar work conducted before, where the inclusion criteria for being classified as fake were either classified as fake news in general or conspiracy theories (Bovet & Makse, 2019, p. 11). Nevertheless, we argue that due to the extended classification framework by Zannettou et al. (2019), couplet with the manual process of checking available information from recognized fact-checking services and academia to the, most shared URLs, this approach holds some merits and act as a valid deviation from the previous literature it builds on. In any case, this thesis, in combination with the work of Zannettou et al. (2019) could be viewed as a evolution of the terminologies - even when the primary goal was not to further define the terminologies.

By conducting manual fact-checking (albeit small number due to scope constraints), we have attempted to address one of the primary challenges of fake news detection on news sites: that using automated processes that is dependable on linguistical patterns is extremely difficult and maybe not suitable for the purpose (Kostakos et al., 2018, p. 1083). While properties surrounding fake news and its spread is starting to get a foothold in the literature (Michela Del Vicario et al., 2016; Pal & Chua, 2019; Vosoughi et al., 2018), it stands to believe that there is a need of understanding fake news and its actors better before we can form more effective automated processes fit for this purpose. As a secondary result, doing this work could be a start at creating better automated processes for effective measurements aimed towards public warning messages related to fake news (B. Ross et al., 2018).

5.4.1 Fake news URLs timeline

As seen in figure 39, and in contrast to figure 35, the propagation patterns were a bit different when comparing most liked tweets with most shared URLs. Where the spikes in model 35 are apparent, they were not so apparent in model 39. Spikes in weeks 26, 36, and 43 (4.58%, 4.49%, 4.16%), although several of the other URLs have weeks with 3% or more of the diffusion, meaning where to set the bar of determining whether it was a spike or not is questionable. There was a visible, although small, increase in propagation over time leading up to the presidential election. As the fake news URLs were mostly focused at the presidential election, this likely affected the results, which also are in line with previous research relating to the spread of fake news throughout the presidential election (Budak, 2019; Pal & Chua, 2019).

5.4.2 Fake News prevalence and concentration

In the analysis of the dataset, there was found 435.099 tweets with URLs classified as fake news of the total 8.636.696 tweets containing URLs. This gives a percentage of 5,03%. The paper by Grinberg et al., where they look at fake news on Twitter related to the 2016 U.S presidential election, found that 5% of the political URLs were from fake news sources (Grinberg et al., 2019). Grinberg's findings were very close to the findings of our study. Bovet and Makse, on the other hand, found that 25% of tweets with links to news outlets linked to either fake or extremely

biased news (Bovet & Makse, 2019). Based on our findings and previous literature, it seems like every study and case is different when it comes to the percentage of tweets that are detected as fake news. This comes as no surprise as the different studies look at different cases and utilize different methods to detect fake news. Different cases and topics attract different actors, both regular users and bots, that may or may not spread fake news. When we look at the spread of fake news over time, the spread of fake news classified URLs, and we can see some periods with high spikes, where the number of tweets containing fake news URLs is higher. This is periods where the share of tweets with fake news classified URLs is more than usual. In figure 41, we can see several of these spikes in the dataset. The highest spike is 4,5 times higher compared to an average day. Since we were unable to get all the tweet-metadata to perform bot detection, it is unclear if the users sharing fake news are normal users or bots.
Previous research has looked at bots in different groups and topics to detect and analyze bot behavior (Abokhodair et al., 2015; Al-Rawi et al., 2019; Bastos & Mercea, 2019; B. Ross et al., 2019). Bastos et al. detected and analyzed a Brexit botnet. The Brexit botnet consisted of 13.493 bots. The bots in this network tweeted a total of 63.797 tweets (Bastos & Mercea, 2019). Right after the Brexit election, the botnet was removed. It is unsure if the botnet was deactivated or if Twitter identified and removed the botnet. Bastos et al. found the botnet shared/retweeted tweets with fake news made by real users. Based on previous research, it is quite possible that a large portion of the users disseminating fake news in the dataset were bots. Bots can share a large number of tweets in a short amount of time (Bastos & Mercea, 2019). In our dataset, the top 10% of users sharing URLs classified as fake news, shared 56% of all accounted for fake news URLs. Similar results where a large portion of the fake news on Twitter are shared by a small group, have been identified before (Bovet & Makse, 2019; Jang et al., 2018). These findings suggest that small groups/ few users can share large amounts of fake news that reach many Twitter users.

5.4.3 Top fake news stories

As seen in section 5.4, through the usage of the most extensive list to-date with fake news actors, we were able to find traces of fake news stories (URLs pointing to one specific news item). We saw that breitbart.com was very much present in these stories, and judging the top 20, eight (40%) relates to this webpage. It was established in section 4.3 that the webpage was very prominent in the dataset, which aligns with previous research (Bovet & Makse, 2019; Jang et al., 2018; Murungi et al., 2018).

Seeing table 18, four stories related to the webpages waterfordwhispernews.com, theamericanmirror.com, wikileaks.org, and toprightnews.com were present. While 3 of them were not previously covered by the literature or in this study, wikileaks.org has previously been labeled as right (leaning) news and used to spread fake news (Bovet & Makse, 2019, p. 4; Pal & Chua, 2019, p. 271).

The top 20 results also saw diversity in its focus, based on the URLs' name (which corresponds to the header of the news item). Europe, Russia, and Great Britain were present outside the American content. The stories also revolve around refugees, where 18 out of the 20 explicitly mentioned refugees or immigrants. From the literature used in this study, neither has investigated these URLs and the claims within. Nevertheless, the wording suggests a negative narrative towards refugees. Negative narratives related to refugees are nothing new, and support for this finding is not uncommon in the literature (Aigner et al., 2017; Bevensee & Ross, 2019, p. 4397; Haug, 2019, p. 1).

5.5 Dataset

With a dataset containing 14.3 million tweets, this study used one of the more extensive datasets to date within the field of research (Bevensee & Ross, 2019, p. 4396; Budak, 2019, p. 141; Rajdev & Lee, 2016, p. 17; Zannettou, Caulfield, et al., 2019, p. 353). While it might not be the largest (Zannettou et al., 2017, p. 406), it stands to believe it is of sufficient size and reach to fulfill the research objectives. The dataset was cleaned in advance of the analyzes (see section 3.8), alas a discussion on further cleaning could be beneficial. Even if keywords primarily focused on the refugee crisis were used to collect the data, the data that was collected was tainted by the American presidential election. Now, that does not render the results of this study invalid; however, the exclusion of tweets whose primary purpose was to focus its message around the American presidential election would be of interest to help identify the underlying themes and identifying fake news within these. 2016 had several overlapping geopolitical situations such as Brexit, the refugee crisis, the American presidential election, amongst others (Lindsay, 2016), and it stands to believe that several of these geopolitical situations was caught in the dataset.

As was seen in section 3.7 and further discussed in section 5.3.1, the data from November and December is lacking. This is troublesome. While we did reach out and discussed this both with our supervisors (which further discussed the matter with the University of Duisburg-Essen), the reason for this being the case remains unclear. In the end, with ten out of twelve months of sufficient data, this should not dismiss the credibility of this research, yet worth taking into consideration and not read too much into the data for November and December.

The structure of the dataset (through format given by the CSV file format standard), seemingly, was a good option for this type of research. While it is not explicitly stated which data format previous research has used, indications given through cues in the papers hints about CSV files being the used option in similar research (Bovet & Makse, 2019, p. 12). Additionally, other formats such as Excel would not work in this context, as its limit on 1 million rows would mean to either split and merge the data (with the possible errors that could cause) or removal of data.

5.6 Method and software

The choice of utilizing an exploratory case study as the research approach is something of value to discuss. We argued that the immaturity of the literature and lack of prior knowledge - both at the dataset as well as the refugee crisis in this context, meant that an exploratory approach would be the most feasible and beneficial to this study. Exploratory approaches within the field of research are nothing new (Al-Rawi et al., 2019; Nied et al., 2017; Pal & Chua, 2019), and thus an exploratory approach was in line with suggestions from the literature (Pal & Chua, 2019, p. 272). If one looks inward (the study group) while keeping the literature excluded, the feasibility of the study surfaces. Often traditional case studies with positivistic paradigms center around hypothesis testing, yet how could one effectively form and test precise hypotheses when one neither knows what the dataset contains and prior knowledge is limited? From the perspective of the group, it was not feasible to shape this study in such a way, and rather this more exploratory approach was crucial to address and answer the research questions.

The method and software were intertwined in this study. For most parts, the analyzes were conducted through KNIME workflows. Latter is interesting, as it turns out that neither of the articles identified in the literature review utilized this software. A valid concern here is whether this means that 1) the software is deemed unfit for purpose, 2) preferences from researches or 3) coincidence. The reason why remains unknown, yet this also creates and shows some novelty to the research field. We have shown that KNIME can be used in academics or otherwise, data analytics through its versatility and community verification as the software with its nodes and workflows is open source. In many of the instances where we utilized custom workflows, these have been documented to increase credibility and transparency in our work. KNIME had a steep learning curve, but once we were used to it, it was highly effective in quickly creating workflows to extract the data needed. The high amounts of documentation through the KNIME Website & forums, Youtube, and in combination with Google search was beneficial to understand both how to structure the workflows but also to understand why to analyze what one does analyze. Several times it turned out that the workflow was not the problem; instead, if you want certain data (such as occurrences of tweets containing the words looking for) - what are the questions (workflow) you need to ask to be able to find the answers you need? Doing so also helped to reflect critically on the approach throughout the study as vague questions would lead to shortcomings or wrong answers.

6 Conclusion and implications

This study aimed to investigate an extensive dataset, with the focus being at identifying if some of its content related to the public discussion was fake, and to such a degree, how much and what type. The context this dataset was collected in was the refugee crisis of 2016, and the group speculated in advance that this would be a natural topic to attract fake news in one of its forms. With this in mind, we shaped the research questions: *What were the most liked and shared tweets in a dataset related to the 2016 refugee crisis, and what were the characteristics of tweets?* and *Using open-source services and a mixture of manual and automatic fake news detection, is it possible to find the presence of fake news in the tweets, and if so, what type of fake news is present?*

These were kept at a high level (broad) by intention, as neither the group nor the originator of the data (the University of Duisburg-Essen) knew what the dataset contained. Based on the data collecting method, it was known that the dataset centered around the refugee crisis, yet to which degree the data collected reflected this remained to be seen. With such uncertainty, the research approach was set to an exploratory case study, with a mixed data analysis.

The starting point was to understand and agree upon the definitions within the study, as the literature utilizes ambiguous terms and definitions. We used Gelferts definition of fake news as a base to enter the literature review and ended up refining this further by following Zannettou et al. (2019) framework for fake news definition and classification (Gelfert, 2018; Zannettou, Sirivianos, Blackburn, et al., 2019). We then, through preliminary literature collection and an extensive literature review, bridged social media with fake news and set the context to the refugee crisis of 2016.

We found through six types of contextual findings that much of the tweets in the dataset focused on the American presidential election. Such a finding in itself was not entirely surprising, as the literature review highlighted the massive spam (extensive traffic) related to this event, yet, to see such the degree visible in the results was not something we expected in advance of the analyzes. The contextual findings also showed that the discussions picked up by the dataset reached globally, with participants from all across the world and through many different geopolitical events. Events such as the geopolitical situation in Syria, Brexit, European politics in general, Australian politics, amongst others, were present. Especially Syria/Syrian as a topic of discussion was present in several of the topic groups, which suggests that this particular geopolitical situation was often found in the center of discussions captured in the dataset. From the 20 most shared individual stories, we saw that content focused on refugees held negative narratives and not solely linked to American politics. Many of these stories originate from webpages known to disseminate fake news, such as Breitbart.com.

Through the broader classification of fake news (broader as more extensive compared to research prior to Zannettou), we found extensive amounts of fake news both in the most shared URLs and most liked messages. Of the most liked messages that were manually checked, more than one third contained fake news. In URLs, similar results.

Similar results are seen in previous research, yet, the novelty here is that we **1)** Performing both manual and automatic processes to detect fake news, **2)** Through the broader classification provided by Zannettou as well as the most extensive list of fake news actors, captured more of what was fake, **3)** Investigate the propagation pattern of some of the various fake news items (not exclusively to specific tweets and, **4)** Explain through fact-checking and literature why something classified as fake news is fake.

Doing so, we know more about the public discussion captured in the dataset, and we know some of the properties surrounding the fake news. It confirms the previous literature that states that fake news plays on the negative emotions of the readers.

As the privacy regulations and scope constraints prevented the group from performing practical network analysis of the actors caught in the dataset, some actors were visible regardless of when usernames were removed. Trump and his supporting actors on Twitter were highly present both in the most liked messages, as well as most shared URLs. The findings yet again aid painting the picture of the extensive amounts of fake news that surrounded the election seen in research prior to this study.

Although terminologies were not one of the primary focus points of this thesis, we saw the challenges of the term *fake news* in both academia and the content studied through the analyzes. As Vosoughi et al. pointed out in their research, the usage of the term gives little sense in academia due to the many ways it is used and the actors using it (Vosoughi et al., 2018). Our study shows and suggests that it is more beneficial to academia (as well as practitioners) to utilize frameworks such as that provided by Zannettou et al. to explain why something is fake, as it increases the granularity and helps better explain why something is fake or “false by design” (Gelfert, 2018).

Lastly, of the 42 articles collected and utilized in the literature review, none had attempted to utilize KNIME in their work. In this regard, this study brings two novelties that could be utilized in academics or practitioners.
1) The software is capable of analyzing large datasets for academic purposes related to the research field, and **2**) Suggests and highlights workflows suitably to analyze Twitter tweets. More details surrounding both KNIME and workflows are visible in appendix A.

6.1 Limitations and possible future research

This study is a result of an attempt to understand more about fake news in social media in the context of the refugee crisis of 2016. By using a large dataset, mixed analyses, and an exploratory approach, the aim was no less than to gain a new understanding of what various actors spoke about on social media. The belief of the group was, and is, that people act differently than what they say they do (and studies on these aspects indicated there might be some truth to this (Coulombe, 2014; Hair et al., 2011)). This meant the group wanted to look at unmoderated raw statements to get a more correct image of the as-was situation.

Due to the lengthy application process (final confirmation in March), coupled with the COVID-19 situation means the study held some additional external challenges. While the group is satisfied with the result, these two factors undoubtedly have affected the study in terms of scope. Even if the group were able to coordinate the efforts through electronic communication, the human interaction is not to be underestimated in academic work where one is required to utilize the knowledge of others in coordinated efforts. The following paragraph relates to the findings and approach of the study. It is structured in such a way that each limitation also provides future research options related to this limitation. It is kept extensive in an attempt to provide a range of possible future directions. It should be read not as highlighting all possible weaknesses, but to reflect the nature of this study; preliminary analyzing of a new dataset.

In terms of literature, this study is more an understanding of what people might write or act on a social medium in a crisis. It does, to some degree, further specify the terminologies and concepts identified in the study, yet - for most parts, it explains what currently “is/was” and attempting to classify it accordingly. Furthermore, from the literature that is brought in, one sees the ambiguousness of *fake news*, with indications that academia itself starts to question whether such a term should even be used in academia at all (Vosoughi et al., 2018). Also, there is a large amount of literature dedicated to bot detection and automated fake news detection. However, the scope of this study meant that literature that did not cover the three overlapping themes (fake news in social media set in the context of crisis) had to be excluded. Lastly, understanding of intricating human actions is too broad for the IS-literature and field alone to explain (Lerbæk & Olsen, 2019b; Talwar et al., 2019). Thus, the findings should be looked at as a step in the direction of building knowledge on this matter. Follow-up studies on both further defining the terminologies (expanding on the framework by Zannettou et al. (2019) to classify further on severity/impact) to decrease the ambiguousness with, particularly, fake news, would be beneficial. Also, further understanding of the dataset at hand to possible linking/isolating of the various geopolitical situations of 2016 would be of interest.

Social media is a collective term for many services, and even if Twitter is a popular platform – both to users and for academia to study (Al-Rawi et al., 2019; Bastos & Mercea, 2019; Buntain & Golbeck, 2017), Twitter alone likely cannot explain all users behavior on social media. There might be dynamics to the platform that is not current at the other platforms and vice versa, yet if and what these remain to be seen. Future research could aim to try distinguishing user behavior related to fake news across the platforms, to see whether it is possible to generalize these platforms or if these should be kept divided.

As for the research approach, which is set to an exploratory case study on research documents with mixed analyzes, some limitations arise in this regard. While it intends to help to understand the problem at hand is suitable in situations where literature is lacking or immature, and to use depth rather than breadth (Oates, 2006, pp. 142 - 143), there is just “so much” that can be said with a dataset that has been collected based on keywords, especially with the analyze limitation

the timeframe this study operated with. Furthermore, while manual evaluation of the material at hand gives the ability to focus on the depth of the data, it also is prone to human errors.
The goal of this study was never to generalize the findings in a population, and a natural question is whether the focus was the most beneficial to the literature. We argue that quantifiable identifying what is written in the public debate and couple with a manual qualitative evaluation to understand the characteristics of this content is a natural step in understanding some of the human actions and dynamics on social media platforms (Twitter). Future research could benefit from dwelling down into smaller selections of this dataset (or other) and conduct deeper qualitative analyzes to gain new insights into the semantics of the tweets. Doing so could both help understanding more of human actions on social media (or Twitter in isolation, see paragraph above) and/or divide the actions based on the several different overlapping geopolitical situations of 2016. Furthermore, by doing this, it would be possible to understand more of the dynamics that lead to people sharing and consummating fake news on social media.

Manually checking content shared on Twitter, such as most liked tweets and occurrences of known fake news actors (through URLs), is a great way to discuss in-details and classify content more precisely than relying solely on automated approaches. However, a study group of two, with the main work being conducted from March to the end of May meant the scope had to be narrowed to make the deadlines. Future studies could expand on both the selection as well as the depth of the selection. By depth in this context means more analyzes such as propagation patterns, statistical analyzes, possible track changes to both webpages, and tweet message alteration over time. The literature points out that fake news sees more alteration than traditional news over time (Jang et al., 2018); thus, tracking these changes could be beneficial to study the effects of fact-checking and “flagging” (as warning messages, debunked in discussions and so on). Secondary benefit by this would mean semantics would be more accessible when one understands more of the tweets (messages) at hand.

As the results showed, a majority of the data collected focused on the American presidential election. Point could be made that this affected some of the novelty of this research, and it would be a more beneficial approach to study the data with the removal of American presidential election content. This concern has been further discussed and addressed within the discussion and conclusion (6), yet, future studies could focus the task on such an endeavor to see the effect this holds on the results of the analyzes. It would also make it more accessible to analyze the different geopolitical situations caught in the dataset.

Lastly, several of the datapoints (longitude & latitude data, usernames, screennames, etc.) had to be removed from the dataset due to privacy regulations and compliance to the NSD/UiA agreement in the study, which meant network mapping, bot detection and similar was not feasible. As a result, the understanding of the various actors (and their impact) is limited in this study.
Future studies could apply for smaller portions of the dataset (ideally linked to specific geopolitical events in 2016) and new applications to see whether the information that was removed initially could be kept. That way, an in-depth study of the actors and their network would help to build up the understanding more of who the specific actors are that spreads fake news on social media. The literature identified and used in the thesis indicated an immaturity on the field, thus after trailing previous studies - as well as this one, in a new context and/or with new data could prove to be beneficial research.

6.2 Endnote: Lessons learned, meta reflections

From the pre-study phase in the autumn of 2019 to the finished product in spring 2020, this process has brought with its valuable knowledge and experience to the group outside just the material and work that makes out the thesis derived from the study. From initially believing that the application to NSD would be extensive but "simple" with enough documentation, to the end approvement, which involved a lengthy Data Protection Impact Assessment (DPIA) document followed up by even more argumentation between the study group and the University of Agder. While frustrating, it also was educational, and thought-provoking having to argue and reflect on matters such as 1) why the data was needed and what type of data it was, 2) what identifiers was within the data, especially when connecting and making "sense" of the data, 3) extreme political utterance and marginalization of groups of people contained in the document - how to mitigate this in the presented results to prevent conveying hate speech, 4) what measurements were taken to address privacy concerns on both technical level and also thesis, 5) what was the focus of the study and lastly 6) what security measures were put in place to ensure that the data was being kept out from the public in its raw form. While these concerns should be addressed in any master thesis, the details had to be a lot clearer before the analyzes could even start. This was a big challenge as this had to be done early before we knew much about both the topic (other than the previous year's assignments) and the actual data. Eventually, the approval was given and left is a densely argued and self-reflective thesis aiming to contribute to progressing the topic of fake news, or in other wording: false information on social media in a crisis event. Note that the group prefers to use false instead of fake, as throughout the last two semesters has shown the ambiguousness of the term. There were lengthy discussions within the group on this matter, but ultimately it was decided to keep the wording "fake" as it was used in great extent within the literature collected and due to the public perception of the word (everyone has an opinion about fake news)

Through structural and analytical work with the usage of KNIME, couplet with literature to argue and discuss what the findings of these analyzes might indicate, it has been good practice in arguing your case - academic and at general levels. It truly has been beneficial to the group members, and hopefully, these preliminary findings will find their way into future work conducted by people with a broader scope and more resources available to them (time, effort, knowledge, computing power)

On an ending note, this also marks the end of our education (for now….) and left standing are two thankful, tired, and hopefully a bit wiser individual (in terms of analytical skills, ethical reflections amongst others). We also believe that what presented here contains shallow levels of false information - hopefully none. Nevertheless, that is up to the future to decide, as experience, new research, and knowledge surfaces.

7 References

Abokhodair, N., Yoo, D., & McDonald, D. W. (2015). *Dissecting a Social Botnet: Growth, Content and Influence in Twitter*. Paper presented at the Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, Vancouver, BC, Canada.

Agrawal, R., de Alfaro, L., Ballarin, G., Moret, S., Di Pierro, M., Tacchini, E., & Della Vedova, M. L. (2019). Identifying fake news from twitter sharing data: a large-scale study. *arXiv preprint arXiv:1902.07207*.

Aigner, J., Durchardt, A., Kersting, H., Kattenbeck, M., & Elsweiler, D. (2017). *Manipulating the perception of credibility in refugee related social media posts*.

AIS. (2020, 2020). Senior Scholars' Basket of Journals. Retrieved from https://aisnet.org/page/SeniorScholarBasket

Akhtar, N., & Ahamad, M. V. (2017). Graph tools for social network analysis. In *Graph Theoretic Approaches for Analyzing Large-Scale Social Networks* (pp. 18-33).

Al-Dahash, H., Thayaparan, M., & Kulatunga, U. (2016). UNDERSTANDING THE TERMINOLOGIES: DISASTER, CRISIS AND EMERGENCY. *Management*, *2*, 1191-1200.

Al-Rawi, A., Groshek, J., & Zhang, L. (2019). What the fake? Assessing the extent of networked political spamming and bots in the propagation of #fakenews on Twitter. *Online Information Review, 43*(1), 53-71. doi:10.1108/OIR-02-2018-0065

Alam, F., Ofli, F., Imran, M., & Aupetit, M. (2018). A twitter tale of three hurricanes: Harvey, irma, and maria. *arXiv preprint arXiv:1805.05144*.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives, 31*(2), 211-236.

Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics, 6*(2), 2053168019848554.

Andrews, C., Fichet, E., Ding, Y., Spiro, E. S., & Starbird, K. (2016). *Keeping up with the tweet-dashians: The impact of 'official' accounts on online rumoring*.

Awan, I. (2014). Islamophobia and Twitter: A typology of online hate against Muslims on social media. *Policy & Internet, 6*(2), 133-150.

Awrahman, B., & Alatas, B. (2017). Sentiment analysis and opinion mining within social networks using Konstanz information miner. *Journal of Telecommunication, Electronic and Computer Engineering, 9*(1), 15-22. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85013135267&partnerID=40&md5=4cd0620eba8e6b28a0e8539938e13083

Ayers, J. W., Hofstetter, C. R., Schnakenberg, K., & Kolody, B. (2009). Is immigration a racial issue? Anglo attitudes on immigration policies in a border county. *Social Science Quarterly, 90*(3), 593-610.

Babcock, M., Cox, R. A. V., & Kumar, S. (2019). Diffusion of pro- and anti-false information tweets: the Black Panther movie case. *Computational and Mathematical Organization Theory, 25*(1), 72-84. doi:10.1007/s10588-018-09286-x

Bastos, M. T., & Mercea, D. (2019). The Brexit Botnet and User-Generated Hyperpartisan News. *Social Science Computer Review, 37*(1), 38-54. doi:10.1177/0894439317734157

Baydogan, C., & Alatas, B. (2018). *Sentiment analysis using Konstanz Information Miner in social networks.* Paper presented at the 6th International Symposium on Digital Forensic and Security, ISDFS 2018 - Proceeding.

BBC. (2015a, September 18, 2015). Migrant crisis: Croatia closes border crossings with Serbia. Retrieved from https://www.bbc.com/news/world-europe-34286432

BBC. (2015b, September 15, 2015). Migrant crisis: Hungary declares emergency at Serbia border. Retrieved from https://www.bbc.com/news/world-europe-34252812

BBC. (2016a, 2016). Donald Trump Jr compares Syrian refugees to Skittles. Retrieved from https://www.bbc.com/news/election-us-2016-37416457

BBC. (2016b, March 4, 2016). Migrant crisis: Migration to Europe explained in seven charts. Retrieved from https://www.bbc.com/news/world-europe-34131911

Berente, N., Hansen, S., Pike, J. C., & Bateman, P. J. (2011). Arguing the value of virtual worlds: patterns of discursive sensemaking of an innovative technology. *Mis Quarterly*, 685-709.

Berry, M., Garcia-Blanco, I., & Moore, K. (2016). Press coverage of the refugee and migrant crisis in the EU: A content analysis of five European countries.

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., . . . Wiswedel, B. (2009). KNIME-the Konstanz information miner: version 2.0 and beyond. *AcM SIGKDD explorations Newsletter, 11*(1), 26-31.

Bevensee, E., & Ross, A. R. (2019). *The Alt-Right and Global Information Warfare*.

Bhatia, A., & Jenks, C. J. (2018). Fabricating the American Dream in US media portrayals of Syrian refugees: A discourse analytical study. *Discourse & Communication, 12*(3), 221-239.

Binkowski, B. (2016, March 29, 2016). Are Refugees Overwhelmingly Young and Male? Retrieved from https://www.snopes.com/fact-check/refugee-invaders-meme/

Bolsonaro, J. M. (2020, February 4). Jair M. Bolsonaro. Retrieved from https://twitter.com/jairbolsonaro

Bond, G. D., Holman, R. D., Eggert, J. A. L., Speller, L. F., Garcia, O. N., Mejia, S. C., . . . Rustige, R. (2017). 'Lyin'Ted','Crooked Hillary', and 'Deceptive Donald': Language of Lies in the 2016 US Presidential Debates. *Applied Cognitive Psychology, 31*(6), 668-677.

Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature communications, 10*(1), 1-14.

Brazys, S., & Panke, D. (2017). Why do states change positions in the United Nations General Assembly? *International Political Science Review, 38*(1), 70-84.

Broadbandsearch.net. (2020). Avarage Time Spend Daily on Social Media (Latest 2020 data) Retrieved from https://www.broadbandsearch.net/blog/average-daily-time-on-social-media

Budak, C. (2019). *What happened? The spread of fake news publisher content during the 2016 U.S. Presidential election.* Paper presented at the The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019.

Buntain, C., & Golbeck, J. (2017). *Automatically Identifying Fake News in Popular Twitter Threads.* Paper presented at the Proceedings - 2nd IEEE International Conference on Smart Cloud, SmartCloud 2017.

Burbach, L., Halbach, P., Ziefle, M., & Valdez, A. C. (2019). *Who shares fake news in online social networks? An agent-based model of different personality models and behaviors in social networks.* Paper presented at the ACM UMAP 2019 - Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization.

Campan, A., Cuzzocrea, A., & Truta, T. M. (2018). *Fighting fake news spread in online social networks: Actual trends and future research directions.* Paper presented at the Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017.

Chouliaraki, L., & Zaborowski, R. (2017). Voice and community in the 2015 refugee crisis: A content analysis of news coverage in eight European countries. *International Communication Gazette, 79*(6-7), 613-635.

CIEM, U. (2019). Social Media Analytics for Society and crisis communication. Retrieved from http://ciem.uia.no/project/social-media-analytics

Clement, J. (2020, April 24, 2020). Countries with the most Twitter users 2020. Retrieved from https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/

Commission, E. (2020, February 13, 2020). EU data protection rules. *GDPR*. Retrieved from https://ec.europa.eu/info/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules/eu-data-protection-rules_en
Contorno, S. (2016, September 26, 2016). Donald Trump says Hillary Clinton lauded controversial Trans-Pacific trade deal. Retrieved from https://www.politifact.com/factchecks/2016/sep/27/donald-trump/donald-trump-says-hillary-clinton-lauded-controver/
Cooke, M., & Buckley, N. (2008). Web 2.0, social networks and the future of market research. *International Journal of Market Research, 50*(2), 267-292.
Coulombe, P. (2014). Low R2 value in social science or education research? Retrieved from https://stats.stackexchange.com/questions/114919/low-r2-value-in-social-science-or-education-research
Cui, A., Zhang, M., Liu, Y., & Ma, S. (2011). *Are the urls really popular in microblog messages?* Paper presented at the 2011 IEEE International Conference on Cloud Computing and Intelligence Systems.
Dahl. Ingvill Dybfest, & Wergeland. Atle Jørstad. (2019). Reagerer på «Radioresepsjonen»-tirade: «Nå er det på tide at de beklager». Retrieved from https://www.vg.no/rampelys/i/AdBv5z/reagerer-paa-radioresepsjonen-tirade-naa-er-det-paa-tide-at-de-beklager
Daily Caller. (2020, 2020). Retrieved from https://www.snopes.com/?s=dailycaller
Danielsen, F. (2019). Systematic Literature review. Retrieved from https://uia.instructure.com/courses/4822/files/?preview=834321
Darwish, K., Magdy, W., & Zanouda, T. (2017). *Trump vs. Hillary: What went viral during the 2016 US presidential election.* Paper presented at the International conference on social informatics.
Davidson, T., & Farquhar, L. (2020). Prejudice and Social Media: Attitudes Toward Illegal Immigrants, Refugees, and Transgender People. In *Gender, Sexuality and Race in the Digital Age* (pp. 151-167): Springer.
Del-Fresno-garcía, M., & Manfredi-Sánchez, J. L. (2018). Politics, hackers and partisan networking. Misinformation, national utility and free election in the catalan independence movement. *Profesional de la Informacion, 27*(6), 1225-1238. doi:10.3145/epi.2018.nov.06
Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., . . . Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences, 113*(3), 554-559.
Del Vicario, M., Gaito, S., Quattrociocchi, W., Zignani, M., & Zollo, F. (2018). *News consumption during the Italian referendum: A cross-platform analysis on facebook and twitter.* Paper presented at the Proceedings - 2017 International Conference on Data Science and Advanced Analytics, DSAA 2017.
Del Vicario, M., Quattrociocchi, W., Scala, A., & Zollo, F. (2019). Polarization and fake news: Early warning of potential misinformation targets. *ACM Transactions on the Web, 13*(2). doi:10.1145/3316809
Emery, D. (2016a, September 29, 2016). Did Hillary Clinton Call the Trans-Pacific Partnership the ‘Gold Standard’ in Trade Agreements? Retrieved from https://www.snopes.com/fact-check/hillary-clinton-called-trans-pacific-partnership-the-gold-standard/
Emery, D. (2016b, August 9, 2016). Did Hillary Clinton Say Christians in America Must Deny Their Faith? Retrieved from https://www.snopes.com/fact-check/clinton-christians-must-deny-faith/

Emery, D. (2017, May 12, 2017). Did Bill Clinton Fire His FBI Director One Day Before Vince Foster Died? Retrieved from https://www.snopes.com/fact-check/clinton-fbi-director-vince-foster/

Evon, D. (2016a, December 22, 2016). CNN and MSNBC in Ratings Freefall, Fox Soundly Beating Both Combined? Retrieved from https://www.snopes.com/fact-check/cnn-and-msnbc-ratings/

Evon, D. (2016b, July 26, 2016). No American Flags Present at the Democratic National Convention. Retrieved from https://www.snopes.com/fact-check/flags-banned-at-dnc/

Evon, D. (2020, March 4, 2020). Will an Asteroid Hit Earth in April 2020? Retrieved from https://www.snopes.com/fact-check/asteroid-warning-april-2020/

FactCheck.org. (2017a, February 16, 2017). FactChecking Trump's News Conference. Retrieved from https://www.factcheck.org/2017/02/factchecking-trumps-news-conference/

FactCheck.org. (2017b, June 6, 2017). Misinformation Directory. Retrieved from https://www.factcheck.org/2017/07/websites-post-fake-satirical-stories/

FactCheck.org. (2020a, 2020). Search pamelageller.com. Retrieved from https://www.factcheck.org/search/

FactCheck.org. (2020b, 2020). Search result "breitbart.com. Retrieved from https://www.factcheck.org/search/

FactCheck.org. (2020c, 2020). Search result "worldnetdaily". Retrieved from https://www.factcheck.org/search/

FactCheck.org. (2020d, 2020). Search results "infowars.com". Retrieved from https://www.factcheck.org/search/

Faris, R., Roberts, H., Etling, B., Bourassa, N., Zuckerman, E., & Benkler, Y. (2017). Partisanship, propaganda, and disinformation: Online media and the 2016 US presidential election. *Berkman Klein Center Research Publication, 6*.

Farley, R. (2016, June 28, 2016). Trump on Clinton's Tax Plans. Retrieved from https://www.factcheck.org/2016/06/trump-on-clintons-tax-plans/

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96-104.

Ferreira, R. R. (2018). Network of lies: The spread of fake news in the Brazilian presidential pre-campaign. *Observatorio, 12*, 139-162. Retrieved from https://www2.scopus.com/inward/record.uri?eid=2-s2.0-85058018400&partnerID=40&md5=d5f208c95c74f9ecc61f25802037a54f

Fichera, A. (2019, May 10, 2019). The 'Raccoon' Rant Not Written by Steve Harvey. Retrieved from https://www.factcheck.org/2019/05/the-raccoon-rant-not-written-by-steve-harvey/

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly, 80*(S1), 298-320. doi:10.1093/poq/nfw006

Fourney, A., Racz, M. Z., Ranade, G., Mobius, M., & Horvitz, E. (2017). *Geographic and temporal trends in fake news consumption during the 2016 US presidential election.* Paper presented at the International Conference on Information and Knowledge Management, Proceedings.

FoxNews. (2016, November 5, 2016). Tweet Retrieved from https://twitter.com/FoxNews/status/794948357718278146

Füller, J., Hutter, K., Hautz, J., & Matzler, K. (2014). User roles and contributions in innovation-contest communities. *Journal of Management Information Systems*. doi:10.2753/MIS0742-1222310111

Galán-García, M. (2017). The 2016 Republican primary campaign on Twitter: Issues and ideological positioning for the profiles of Ben Carson, Ted Cruz, Marco Rubio, and Donald Trump. *El profesional de la información, 26*(5), 850-858.

Gallego, M., Gualda, E., & Rebollo, C. (2017). Women and Refugees in Twitter: Rethorics of Abuse, Vulnerability and Violence from a Gender Perspective.

García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining*: Springer.

Garrett, L. (2020). COVID-19: the medium is the message. *The lancet, 395*(10228), 942-943.

Gelfert, A. (2018). Fake News: A Definition. *Informal Logic, 38*(1), 84-117. doi:https://doi.org/10.22329/il.v38i1.5068

Geller, P. (2020, 2020). Biography. Retrieved from https://gellerreport.com/about/

Georgiou, M., & Zaborowski, R. (2017). *Media coverage of the "refugee crisis": A cross-European perspective*: Council of Europe.

Gordon, L. (2016, November 8, 2016). Lindsay Lohan is using her weird accent to literally save the world. Retrieved from https://www.revelist.com/celebrity/lindsay-lohan-accent-shirt-charity/5673

Greenwood, S., Perrin, A., & Duggan, M. (2016). Social media update 2016. *Pew Research Center, 11*(2).

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science, 363*(6425), 374-378.

Grøtan, T. O., Fiskvik, J., Haro, P. H., Auran, P. G., Mathisen, B. M., Karlsen, G. H., . . . Brandtzæg, P. B. (2019). *På leting etter utenlandsk informasjonspåvirkning* (2). Retrieved from regjeringen.no: https://www.regjeringen.no/contentassets/4d850821991746ecbcd9477a475baf73/sintef-rapport_2019-01292_gradering_apen.pdf

Group, S. C. W. (2020, 2019). ANALYSIS TOOLS FOR SOCIAL MEDIA. Retrieved from https://sites.google.com/view/social-cybersec/tools?authuser=0

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139-151. doi:10.2753/MTP1069-6679190202

Haug, M. (2019). *Endorse the source – The impact of information assessment on news sharing behavior*.

Heidemann, J., Klier, M., & Probst, F. (2010). Identifying key users in online social networks: A pagerank based approach.

Heller, S. (2014, Oktober 16, 2014). Bat Boy, Hillary Clinton's Alien Baby, and a Tabloid's Glorious Legacy. Retrieved from https://www.theatlantic.com/entertainment/archive/2014/10/the-ingenious-sensationalism-of-the-weekly-world-new/381525/

Helsper, E. J., & Eynon, R. (2010). Digital natives: where is the evidence? *British educational research journal, 36*(3), 503-520.

Holmes, S. M., & Castañeda, H. (2016). Representing the "European refugee crisis" in Germany and beyond: Deservingness and difference, life and death. *American Ethnologist, 43*(1), 12-24.

Hughes, A. L., & Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. *International journal of emergency management, 6*(3-4), 248-260.

Inglehart, R. F., & Norris, P. (2016). Trump, Brexit, and the rise of populism: Economic have-nots and cultural backlash.

Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer, 51*(8), 56-59.

ITV. (2016, February 15, 2016). Bulgaria to close all external EU borders, PM announces. Retrieved from https://www.itv.com/news/update/2016-02-15/bulgaria-to-close-all-external-eu-borders-pm-announces/

Jacobsen, D. I. (2016). *Hvordan gjennomføre undersøkelser* (C. D. Akademisk Ed. Vol. 3): Cappelen Damm Akademisk.

Jacobson, L. (2016, December 14, 2016). "Before Leaving, Obama Enacts Rules to Take $3,000 From EVERY American.". Retrieved from https://www.politifact.com/factchecks/2016/dec/14/conservative-tribune/conservative-website-exaggerates-misleads-post-reg/
Jamieson, A. (2016, October 10, 2016). 'Big League Truth Team' pushes Trump's talking points on social media. Retrieved from https://www.theguardian.com/us-news/2016/oct/10/donald-trump-big-league-truth-team-social-media-debate
Jang, S. M., Geng, T., Queenie Li, J. Y., Xia, R., Huang, C. T., Kim, H., & Tang, J. (2018). A computational approach for examining the roots and spreading patterns of fake news: Evolution tree analysis. *Computers in Human Behavior, 84*, 103-113. doi:10.1016/j.chb.2018.02.032
Jones, M. O. (2019). The gulf information war| propaganda, fake news, and fake trends: The weaponization of twitter bots in the gulf crisis. *International Journal of Communication, 13*, 27.
Journal, T. W. (2020, 2020). About The Western Journal. Retrieved from https://www.westernjournal.com/about/
Jr., C. E. E. (2016, June 13, 2016). Donald Trump says Hillary Clinton wants to let 500 percent more Syrians into the U.S. Retrieved from https://www.politifact.com/factchecks/2016/jun/13/donald-trump/donald-trump-says-hillary-clinton-wants-let-500-pe/
Kaboub, F. (2008). Positivist paradigm. *Encyclopaedia of counselling, 2*(2), 343.
Kalajdzic. Pedja, & Vigsnæs. Maria Knoph. (2019). NRK fjerner «Radioresepsjonen»-innslag etter reaksjoner.
Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University, 33*(2004), 1-26.
Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology, 51*(1), 7-15.
KNIME. (2020a). KNIME Open Source Story. Retrieved from https://www.knime.com/knime-open-source-story
KNIME. (2020b, February 27, 2020). Palladian Nodes for KNIME (trusted extension). Retrieved from https://www.knime.com/community/palladian
Kogan, M., Palen, L., & Anderson, K. M. (2015). *Think local, retweet global: Retweeting by the geographically-vulnerable during hurricane sandy*.
Kostakos, P., Nykanen, M., Martinviita, M., Pandya, A., & Oussalah, M. (2018). *Meta-terrorism: Identifying linguistic patterns in public discourse after an attack.* Paper presented at the Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018.
Krogstad, J. M. (2019, Oktober 7, 2019). Key facts about refugees to the U.S. Retrieved from https://www.pewresearch.org/fact-tank/2019/10/07/key-facts-about-refugees-to-the-u-s/
Lacapria, K. (2015, 2015). False Flag. Retrieved from https://www.snopes.com/fact-check/isis-flag-refugees-germany/
Lacapria, K. (2018, 2020). Did Maine House Democrats Vote to Allow Female Genital Mutilation? Retrieved from https://www.snopes.com/fact-check/did-maine-democrats-vote-to-allow-fgm/
Lee, J.-S., & Nerghes, A. (2018). Refugee or migrant crisis? Labels, perceived agency, and sentiment polarity in online discussions. *Social Media+ Society, 4*(3), 2056305118785638.
Lerbæk, T. O., & Olsen, B. E. V. (2019a). *Falske Nyheter på Sosiale Medier og hvordan disse spres– En litteraturstudie*. Literature Review. Universitetet i Agder.

Lerbæk, T. O., & Olsen, B. E. V. (2019b). *Reasons to Why People Spread Fake News Items on Social Media* Survey. Universitetet i Agder.

Li, H., Dombrowski, L., & Brady, E. (2018). *Working toward empowering a community: How immigrant-focused nonprofit organizations use twitter during political conflicts*.

Lindsay, J. M. (2016, December 2016). Ten Most Significant World Events in 2016. Blog Page Retrieved from https://www.cfr.org/blog/ten-most-significant-world-events-2016

Loescher, G. (2001). The UNHCR and world politics: state interests vs. institutional autonomy. *International Migration Review, 35*(1), 33-56.

Maddock, J., Starbird, K., Al-Hassani, H., Sandoval, D. E., Orand, M., & Mason, R. M. (2015). *Characterizing online rumoring behavior using multi-dimensional signatures*.

Mallet. (2018, 2018). MAchine Learning for LanguagE Toolkit. Retrieved from http://mallet.cs.umass.edu/

Minanović, A., Gabelica, H., & Krstić, Ž. (2014). *Big data and sentiment analysis using KNIME: Online reviews vs. social media.* Paper presented at the 2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2014 - Proceedings.

Moore, C. (2016, October 10, 2016). Lindsay Lohan flashes a beaming smile as she cheers up Syrian refugee children in Turkish 'container town'. Retrieved from https://www.dailymail.co.uk/tvshowbiz/article-3828513/Lindsay-Lohan-flashes-beaming-smile-cheers-Syrian-refugee-children-Turkish-container-town.html

Murphy, M. (2018). Government bans phrase 'fake news' Retrieved from https://www.telegraph.co.uk/technology/2018/10/22/government-bans-phrase-fake-news/

Murungi, D., Purao, S., & Yates, D. J. (2018). *Beyond facts: A new spin on fake news in the age of social media.* Paper presented at the Americas Conference on Information Systems 2018: Digital Disruption, AMCIS 2018.

Narayanan, V., Barash, V., Kelly, J., Kollanyi, B., Neudert, L.-M., & Howard, P. N. (2018). Polarization, partisanship and junk news consumption over social media in the us. *arXiv preprint arXiv:1803.01845*.

Newman, D., Asuncion, A., Smyth, P., & Welling, M. (2009). Distributed algorithms for topic models. *Journal of Machine Learning Research, 10*(Aug), 1801-1828.

Nicols, C. (2017, February 3, 2017). America's refugee vetting robust, not 'haphazard' as Calif. congressman claims. Retrieved from https://www.politifact.com/factchecks/2017/feb/03/tom-mcclintock/americas-refugee-vetting-already-robust-not-haphaz/

Nied, A. C., Stewart, L., Spiro, E., & Starbird, K. (2017). *Alternative narratives of crisis events: Communities and social botnets engaged on social media*.

NSD. (2020a, February 13, 2020). Do you wonder if you have to conduct a Data Protection Impact Assessment (DPIA)? Retrieved from https://www.datatilsynet.no/en/

NSD. (2020b, June, 2018). GDPR: Nytt meldeskjema for behandling av personopplysninger klart når loven trer i kraft. Retrieved from https://nsd.no/article.html?a=/articles/article0065.html

Oates, B. J. (2006). *Researching Information Systems and Computing* (2006 ed.): Sage.

Oh, O., Kwon, K. H., & Rao, H. R. (2010). *An exploration of social media in extreme events: Rumor theory and twitter during the HAITI earthquake 2010*.

OpenSources. (2017). False, Misleading, Clickbait-y, and/or Satirical "News" Sources. Retrieved from https://docs.google.com/document/d/10eA5-mCZLSS4MQY5QGb5ewC3VAL6pLkT53V_81ZyitM/preview. Retrieved 07.05.20 https://docs.google.com/document/d/10eA5-mCZLSS4MQY5QGb5ewC3VAL6pLkT53V_81ZyitM/preview

Orso, A. (2016, October 26, 2016). "Recently though, the media has reported on thousands of illegal voters here in the commonwealth.". Retrieved from https://www.politifact.com/factchecks/2016/oct/26/john-rafferty/gop-attorney-general-candidate-cant-back-claim-wid/

Pal, A., & Chua, A. Y. K. (2019). *Propagation Pattern as a Telltale Sign of Fake News on Social Media*. Paper presented at the 5th International Conference on Information Management, ICIM 2019.

Palma, B. (2017, October 3, 2017). Did a Second Gunman Shoot From the Fourth Floor of the Mandalay Bay Hotel? Retrieved from https://www.snopes.com/fact-check/second-gunman-shoot-fourth-floor-mandalay-bay/

Palma, B. (2018, February 22, 2018). Did a Muslim Scholar Advocate 'Colonizing' U.S. and Canada with Sharia Law? Retrieved from https://www.snopes.com/fact-check/muslim-scholar-tariq-ramadan-sharia-law-colonize/

Politifact. (2017). PolitiFact's guide to fake news websites and what they peddle Retrieved from https://www.politifact.com/article/2017/apr/20/politifacts-guide-fake-news-websites-and-what-they/

PolitiFact. (2020a, 2020). Daily Caller. Retrieved from https://www.politifact.com/personalities/daily-caller/

PolitiFact. (2020b, 2020). Search result "worldnetdaily". Retrieved from https://www.politifact.com/search/?q=worldnetdaily

PolitiFact. (2020c, 2020). Search results "breitbart.com". Retrieved from https://www.politifact.com/search/?q=breitbart.com

PolitiFact. (2020d, 2020). Search results "infowars.com". Retrieved from https://www.politifact.com/search/?q=infowars.com

PolitiFact. (2020e, 2020). Search results for "pamelageller.com". Retrieved from https://www.politifact.com/search/?q=pamelageller.com

PolitiFact. (2020f, 2020). Search results for "sputniknews.com". Retrieved from https://www.politifact.com/search/?q=sputniknews.com

PolitiFact. (2020g, 2020). Search results for "truthfeed.com". Retrieved from https://www.politifact.com/search/?q=truthfeed.com

Poynter.org. (2018). Fakecheck content analysis. *2020*(16.05). Retrieved from https://docs.google.com/spreadsheets/d/11ytmr7-fpUNYWaIHoNhY1vmdftO_U5ew2JmLk2EnXIc/edit#gid=0

Prensky, M. (2001). Digital natives, digital immigrants. *On the horizon, 9*(5).

Putterman, S. (2018, December 7, 2018). Complex tale involving Hillary Clinton, uranium and Russia resurfaces. Retrieved from https://www.politifact.com/factchecks/2018/dec/07/blog-posting/complex-tale-involving-hillary-clinton-uranium-rus/

Quu, L. (2016, June 13, 2016). Wrong: Donald Trump says there's 'no system to vet' refugees. Retrieved from https://www.politifact.com/factchecks/2016/jun/13/donald-trump/wrong-donald-trump-says-theres-no-system-vet-refug/

Rajdev, M., & Lee, K. (2016). *Fake and spam messages: Detecting misinformation during natural disasters on social media*.

Rajput, Q., Haider, S., & Ghani, S. (2016). Lexicon-based sentiment analysis of teachers' evaluation. *Applied Computational Intelligence and Soft Computing, 2016*.

Reese, R. M., Reese, J. L., Kaluza, B., Kamath, U., & Choppella, K. (2017). *Machine Learning: End-to-End guide for Java developers: Data Analysis, Machine Learning, and Neural Networks simplified*: Packt Publishing Ltd.

RISE_SMA. (2020). What is RISE_SMA? Retrieved from https://social-media-analytics.org/

Robertson;, L., Farley;, R., Kiely;, E., Gore;, D. A., & Schaedel;, S. (2016, July 27, 2016). Trump's Press Conference. Retrieved from https://www.factcheck.org/2016/07/trumps-press-conference/
Rosa, K. D., Shah, R., Lin, B., Gershman, A., & Frederking, R. (2011). Topical clustering of tweets. *Proceedings of the ACM SIGIR: SWSM, 63*.
Ross, A. S., & Rivers, D. J. (2018). Discursive Deflection: Accusation of "Fake News" and the Spread of Mis- and Disinformation in the Tweets of President Trump. *Social Media and Society, 4*(2). doi:10.1177/2056305118776010
Ross, B., Heisel, J., Jung, A. K., & Stieglitz, S. (2018). *Fake news on social media: The (in)effectiveness of warning messages*. Paper presented at the International Conference on Information Systems 2018, ICIS 2018.
Ross, B., Pilz, L., Cabrera, B., Brachten, F., Neubaum, G., & Stieglitz, S. (2019). Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks. *European Journal of Information Systems, 28*(4), 394-412. doi:10.1080/0960085X.2018.1560920
Salvo, D., & De Leon, S. (2018). *Russia's Efforts to Destabilize Bosnia and Herzegovina*: German Marshall Fund of the United States.
Schell, C. (1992). The value of the case study as a research strategy. *Manchester Business School, 2*, 1-15.
Scopus. (2020, February 11, 2020). Analyze search results. Retrieved from https://www.scopus.com/term/analyzer.uri?sid=963e80b2f83d7fe7d67d4b4162a8ece7&origin=resultslist&src=s&sort=cp-f&sdt=a&sot=a&sessionSearchId=963e80b2f83d7fe7d67d4b4162a8ece7&count=177&analyzeResults=Analyze+results&cluster=scopubstage%2c%22final%22%2ct%2bscosubjabbr%2c%22COMP%22%2ct%2c%22SOCI%22%2ct%2c%22ARTS%22%2cf%2c%22PSYC%22%2cf%2c%22BUSI%22%2cf%2c%22ECON%22%2cf%2c%22AGRI%22%2cf%2c%22ENER%22%2cf%2c%22ENVI%22%2cf%2c%22MATH%22%2cf%2c%22SOCI%22%2cf%2c%22DECI%22%2cf%2c%22MEDI%22%2cf%2c%22EART%22%2cf%2c%22MATE%22%2cf%2c%22HEAL%22%2cf%2c%22NURS%22%2cf%2c%22PHYS%22%2cf%2bscolang%2c%22English%22%2ct&txGid=aca429dc591da35956276b0393122672
SCWG. (2020). ANALYSIS TOOLS FOR SOCIAL MEDIA. Retrieved from https://sites.google.com/view/social-cybersec/tools?authuser=0
Semmens, D., Moon, B., Bolliet, L., Amarasekara, I., & McKinnon, M. (2016). Science, Twitter and election campaigns: tracking# auspol in the Australian federal elections.
Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of fake news by social bots. *arXiv preprint arXiv:1707.07592, 96*, 104.
Sherman, A. (2016, October 9, 2016). Trump: Clinton won't use the term 'radical Islamic terrorists'. Retrieved from https://www.politifact.com/factchecks/2016/oct/09/donald-trump/trump-clinton-wont-use-term-radical-islamist-terro/
Sherman, A. (2018, July 13, 2018). Fake headline says New York Times said Trump would make the best president in 1984. Retrieved from https://www.politifact.com/factchecks/2018/jul/13/bling-news/fake-headline-says-new-york-times-said-trump-would/
Shoebat, T. (2020, April 16, 2020). The Nazis Are Coming Back: Covid-19 And The Return Of The 1930s. Retrieved from http://shoebat.com/
Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter, 19*(1), 22-36.

Skjeggestad;, S. R., Sandvik;, S., & Omland, E. (2017, February 1, 2017). Ny rapport mener økt innvandring kan utfordre velferdsstaten. Retrieved from https://www.nrk.no/norge/ny-rapport-mener-okt-innvandring-kan-utfordre-velferdsstaten-1.13353790
Snopes. (2018, 2020). Retrieved from https://www.snopes.com/fact-check/muslima-nurse-practitioner-son/
Snopes. (2020a, 2020). Search result "infowars.com". Retrieved from https://www.snopes.com/?s=infowars.com
Snopes. (2020b, 2020). Search result "worldnetdaily". Retrieved from https://www.snopes.com/?s=worldnetdaily
Snopes. (2020c, 2020). Search results "breitbart.com". Retrieved from https://www.snopes.com/?s=breitbart.com
Snopes. (2020d, 2020). Search results "express.co.uk". Retrieved from https://www.snopes.com/?s=express.co.uk
Snopes. (2020e, 2020). You searched for: pamelageller.com. Retrieved from https://www.snopes.com/?s=pamelageller.com
Spohr, D. (2017). Fake news and ideological polarization:Filter bubbles and selective exposure on social media. *Business Information Review, 34*(3), 150-160. doi:10.1177/0266382117722446
Starbird, K., Dailey, D., Mohamed, O., Lee, G., & Spiro, E. S. (2018). *Engage early, correct more: How Journalists participate in false rumors online during crisis events*.
Starbird, K., Maddock, J., Orand, M., Achterman, P., & Mason, R. M. (2014). Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing. *IConference 2014 Proceedings*.
Statista. (2020, 2020). Number of social network users worldwide from 2010 to 2021. *Social Media & User-Generated Content.* Retrieved from https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/
Stein, A., & Salime, Z. (2015). Manufacturing Islamophobia: rightwing pseudo-documentaries and the paranoid style. *Journal of communication inquiry, 39*(4), 378-396.
Swar, B., & Hameed, T. (2017). *Fear of Missing out, Social Media Engagement, Smartphone Addiction and Distraction: Moderating Role of Self-Help Mobile Apps-based Interventions in the Youth.* Paper presented at the HEALTHINF.
Talwar, S., Dhir, A., Kaur, P., Zafar, N., & Alrasheedy, M. (2019). Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior. *Journal of Retailing and Consumer Services, 51*, 72-82. doi:10.1016/j.jretconser.2019.05.026
Tandoc Jr, E. C., Lim, Z. W., & Ling, R. (2018). Defining “fake news” A typology of scholarly definitions. *Digital journalism, 6*(2), 137-153.
Thiel, K. (2014). Topic Detection LDA: Summarizing Romeo & Juliet or cataloging News. KNIME. Retrieved from https://hub.knime.com/knime/spaces/Examples/latest/08_Other_Analytics_Types/01_Text_Processing/25_Topic_Detection_LDA
TinEye. (2020). Reverse Image Search. Retrieved from https://tineye.com/
Tjernshaugen, K. (2016, January 8, 2016). Listhaug: Det er klart jeg er bekymret for mine barns fremtid. *Politikk.* Retrieved from https://www.aftenposten.no/norge/politikk/i/aP1M/listhaug-det-er-klart-jeg-er-bekymret-for-mine-barns-fremtid
Trump, D. J. (2016a, October 20, 2016). Tweet link 1. Retrieved from https://twitter.com/realDonaldTrump/status/788930678255517696
Trump, D. J. (2016b, May 22, 2016). Tweet link 2. Retrieved from https://twitter.com/realDonaldTrump/status/734468447829004288

Trump, D. J. (2016c, July 28, 2016). Tweet link 3. Retrieved from https://twitter.com/realDonaldTrump/status/758761945910669312
Trump, D. J. (2016d, August 15, 2016). Tweet link 4. Retrieved from https://twitter.com/realDonaldTrump/status/765283490887790592
Trump, D. J. (2019, 11 November 2019). Donald J. Trump. Retrieved from https://twitter.com/realDonaldTrump
Twitter. (2016, May 24, 2016). Doing more with 140 characters. Retrieved from https://blog.twitter.com/developer/en_us/a/2016/doing-more-with-140-characters.html
Twitter. (2020a, April 13). About Twitter's link service (http://t.co). Retrieved from https://help.twitter.com/en/using-twitter/url-shortener
Twitter. (2020b, March 23, 2020). How to use hashtags. Retrieved from https://help.twitter.com/en/using-twitter/how-to-use-hashtags
Twitter. (2020c, February 12, 2020). Overview. Retrieved from https://developer.twitter.com/en/docs/tutorials/consuming-streaming-data
Twitter. (2020d, May 12, 2020). #refugeesGR. Retrieved from https://twitter.com/hashtag/refugeesgr
UNDRR. (2020, February 7, 2020). Disaster. *Terminology.* Retrieved from https://www.undrr.org/terminology/disaster
Urquhart, C., & Vaast, E. (2012). *Building social media theory from case studies: A new frontier for is research*.
Valeeva, A. M. A. (2017, July 27, 2017). Is Russia Today a legitimate fact-checker? We did the math. Retrieved from https://www.poynter.org/fact-checking/2017/is-russia-today-a-legitimate-fact-checker-we-did-the-math/
Valverde, M. (2016, September 27, 2016). Trump says Clinton would bring in 620,000 refugees in her first term. Retrieved from https://www.politifact.com/factchecks/2016/sep/27/donald-trump/trump-says-clinton-would-bring-620000-refugees-her/
Valverde, M. (2017, February 2, 2017). Most Syrian refugees are women and children, as Keith Ellison said. Retrieved from https://www.politifact.com/factchecks/2017/feb/02/keith-ellison/rep-keith-ellison-correct-demographic-overview-syr/
Van Leuven, S., Deprez, A., Joye, S., & Ongenaert, D. (2018). How do Flemish politicians talk about migration? A study on the political framing of migration in Belgium 2016-2018.
Volkova, S., & Bell, E. (2017). *Identifying effective signals to predict deleted and suspended accounts on Twitter across languages*.
Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146-1151. doi:10.1126/science.aap9559
Walshe, C. E., Caress, A. L., Chew-Graham, C., & Todd, C. J. (2004). Case studies: A research strategy appropriate for palliative care? *Palliative Medicine, 18*(8), 677-684.
Watanabe, K. (2018). *Conspiracist propaganda: How Russia promotes anti-establishment sentiment online.* Paper presented at the ECPR General Conference, Hamburg.
Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *Mis Quarterly, 26*(2), xiii-xxiii. Retrieved from http://www.jstor.org/stable/4132319
whitehouse.gov. (2015). Infographic: The Screening Process for Refugee Entry into the United States. In.
Wilder, B., & Vorobeychik, Y. (2018). *Controlling elections through social influence.* Paper presented at the Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS.
Yablokov, I. (2015). Conspiracy theories as a Russian public diplomacy tool: The case of Russia Today (RT). *Politics, 35*(3-4), 301-315.

Yao, L., Mimno, D., & McCallum, A. (2009). *Efficient methods for topic model inference on streaming document collections.* Paper presented at the Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining.

Zakharchenko, A., Maksimtsova, Y., Iurchenko, V., Shevchenko, V., & Fedushko, S. (2019). *Under the conditions of non-agenda ownership: Social media users in the 2019 Ukrainian presidential elections campaign.* Paper presented at the CEUR Workshop Proceedings.

Zang, J., & Pottle, J. (2016). Does Fake News Work? Fake News and Ideological Bias in News Consumption. *Ethnicity, 45*(59), 39.

Zannettou, S., Caulfield, T., De Cristofaro, E., Kourtellis, N., Leontiadis, I., Sirivianos, M., . . . Blackburn, J. (2017). *The web centipede: Understanding how web communities influence each other through the lens of mainstream and alternative news sources.* Paper presented at the Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC.

Zannettou, S., Caulfield, T., Setzer, W., Sirivianos, M., Stringhini, G., & Blackburn, J. (2019). *Who let the trolls out? towards understanding state-sponsored trolls*.

Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2019). The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality, 11*(3). doi:10.1145/3309699

Zannettou, S., Sirivianos, M., Caulfield, T., Stringhini, G., De Cristofaro, E., & Blackburn, J. (2019). *Disinformation warfare: Understanding state-sponsored trolls on twitter and their influence on the web*.

Zhou, X., & Zafarani, R. (2018). Fake news: A survey of research, detection methods, and opportunities. *arXiv preprint arXiv:1812.00315*.

Zimmer, F., Scheibe, K., Stock, M., & Stock, W. G. (2019a). Fake news in social media: Bad algorithms or biased users? *Journal of Information Science Theory and Practice, 7*(2), 40-53. doi:10.1633/JISTaP.2019.7.2.4

Zimmer, F., Scheibe, K., Stock, M., & Stock, W. G. (2019b). Fake news in social media: Bad algorithms or biased users? *Journal of Information Science Theory and Practice*. doi:10.1633/JISTaP.2019.7.2.4

Appendix

In this chapter, the material of relevance to the thesis is attached. Some of the material contains little to no further explanation, as in those cases, it is solely meant to give context at previously placed in the text. One such example is appendix G, which acts as a reference of relevance to section 4.1.6.

Appendix A.1: KNIME Workflows

In the following sections, A.1 - A.17, we have highlighted some of the workflows utilized in the analyzes in the study.

Appendix A.2: Workflow 1: Data Preprocessing

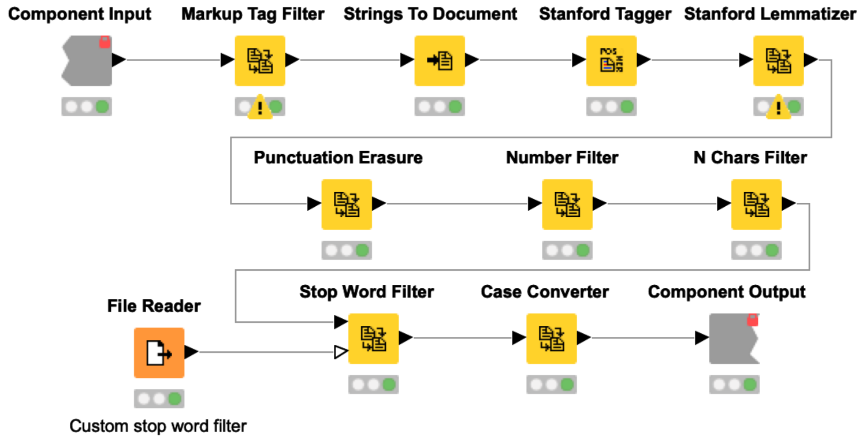


Figure 44 - Data Preprocessing

The data processing workflow starts by importing the dataset. The dataset is then filtered through a markup tag filter where markup language/programming language syntax is removed. The data is then converted to documents (as needed to ensure KNIME can process the text further, data is converted to a suitable format). The next node is the Stanford Tagger, where each term of the document is assigned a part of speech tag (e.g., verb, adjective, noun). After the tagger node is the Stanford Lemmatizer, which reduces each word to its original stem (e.g., tweets to tweet). The next node, Punctuation Erasure, remove punctuation characters form the documents. The Number Filter filters all numerical terms, numbers, and operators. The N Chars Filter removes all terms with less than specified n characters. The stop word filter removes specified words. Words that do not give much meaning to a text, e.g., the, a, in, an. The Case Converter converts all the documents to the same lower case.

Appendix A.3: Workflow 2: Locations

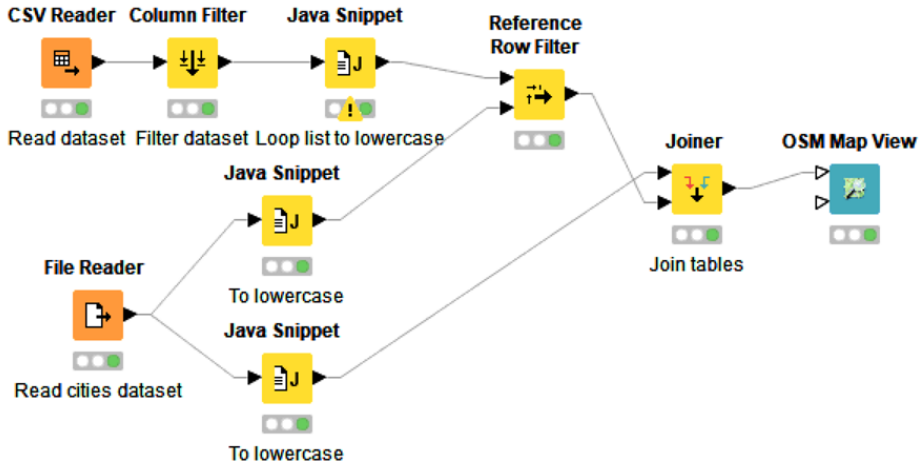


Figure 45 - Workflow locations

The locations workflow filters out all other columns but the locations column. The locations column uses a list of cities as a reference and removes all other strings that are not on the list. The list of cities has coordinates for longitude and latitude for each city (cities with more than 15 00 citizens). The locations column is then joined with the cities list so that each entry in has coordinates. The tweets are then mapped on a world map.

Appendix A.4: Workflow 3: Tweets timeline

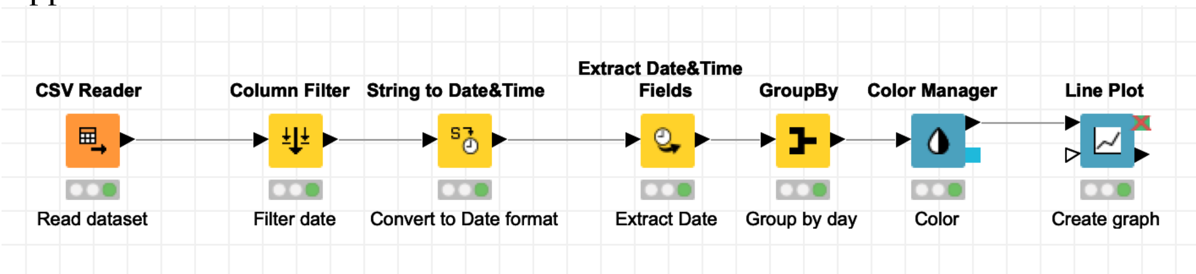
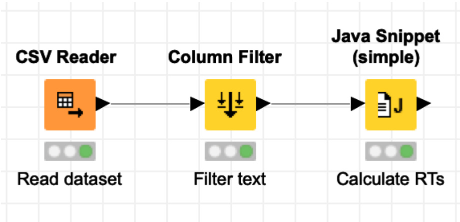


Figure 46 - Workflow Tweet timeline

The tweet timeline filters the dataset so that only the id and timestamp are left. The timestamp is converted to a valid date format. Then each date is extracted and grouped and put in a colored line graph. Note that there are various versions of this, where the focus was at keeping tweet text rather than ID (as ID was removed due to DPIA agreements)

Appendix A.5: Workflow 4: Calculate RTs




```
if (!Text.startsWith("RT"))
    out_Text = Text;
```

Figure 47 - Workflow for identifying Retweets

Retweets start with ‘RT.’ To filter out retweets, we used a java snippet that looks for tweets that start with ‘RT.’ If a tweet matches the criteria, we keep the tweets, if not, we remove it.

Appendix A.6: Workflow 5: Topic detection

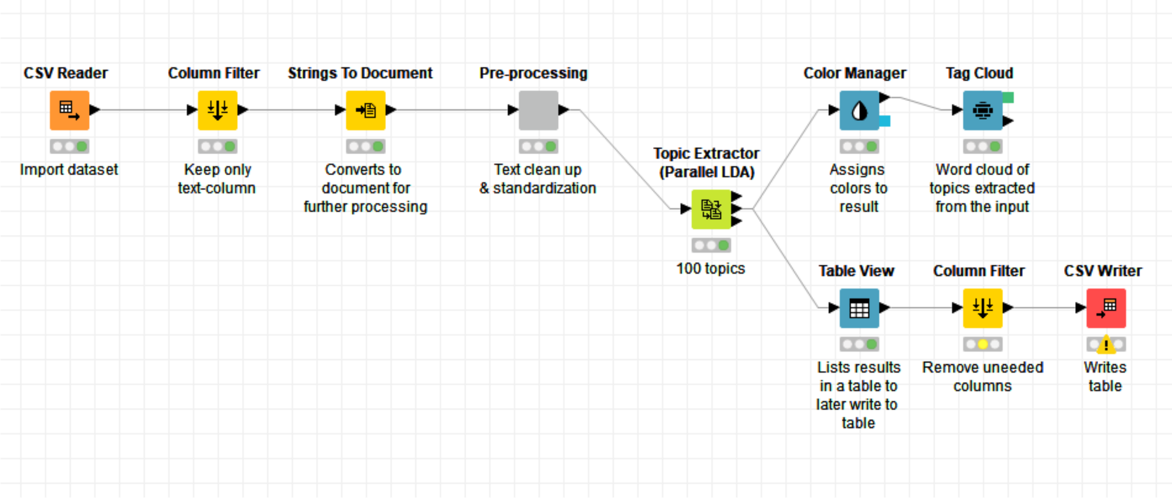


Figure 48 - Workflow for Topic detection

The topic detection workflow filters out all other columns but the text column. The strings are converted into documents, and then the documents are preprocessed (Appendix A.2). After the preprocessing is done, the documents are sent to the LDA (Latent Dirichlet Allocation) node, where the different topics are extracted. This workflow both illustrates the results in KNIME, as well as exporting them to a CSV file. A later version of this exported to excel files instead.

Appendix A.7: Workflow 6: Identify URLs

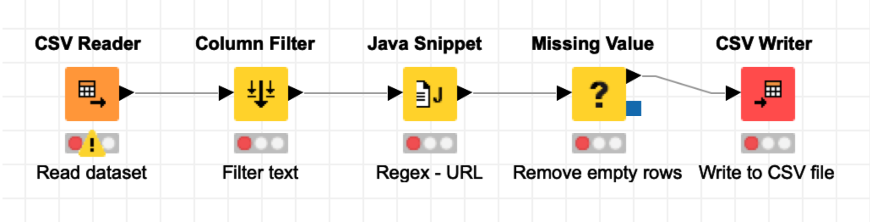


Figure 49 - Workflow for identifying URLs

The workflow starts by importing and reading the dataset. The dataset is filtered so only the text is left. Everything else is removed. The dataset is then sent to the java snippet node where regular expressions is used to identify URLs. If there are no URLs in the tweet the entry is left empty.

```java
Pattern pattern = Pattern.compile(
                "\\b(((ht|f)tp(s?)\\:\\/\\/|~\\/|\\/)|www.)" +
                "(\\w+:\\w+@)?(([-\\w]+\\.)+(com|org|net|gov" +
                "|mil|biz|info|mobi|name|aero|jobs|museum" +
                "|travel|[a-z]{2}))(:[\\d]{1,5})?" +
                "(((\\/([-\\w~!$+|.,=]|%[a-f\\d]{2})+)+|\\/)+|\\?|#)?" +
                "((\\?([-\\w~!$+|.,*:]|%[a-f\\d{2}])+=?" +
                "([-\\w~!$+|.,*:=]|%[a-f\\d]{2})*)" +
                "(&(?:[-\\w~!$+|.,*:]|%[a-f\\d{2}])+=?" +
                "([-\\w~!$+|.,*:=]|%[a-f\\d]{2})*)*)*" +
                "(#([-\\w~!$+|.,*:=]|%[a-f\\d]{2})*)?\\b");

// expression start


try{
    Matcher matcher = pattern.matcher(c_Text);

    while (matcher.find()) {
        out_Text = matcher.group();

        }
} catch (PatternSyntaxException e){
    System.out.println(e.getMessage());
}

```

Figure 50 - Regular expression node

The java snippet node uses regular expressions to identify the different URLs the tweets contain. The string regular expression for URL identification is compiled into an instance of the Pattern class. The pattern is then used to create a Matcher object that can match the regular expression against the dataset. We then loop through the dataset. If there is a match, the URL is stored in the out_Text variable and replace the original text. If there is no match, the original text will be left empty.

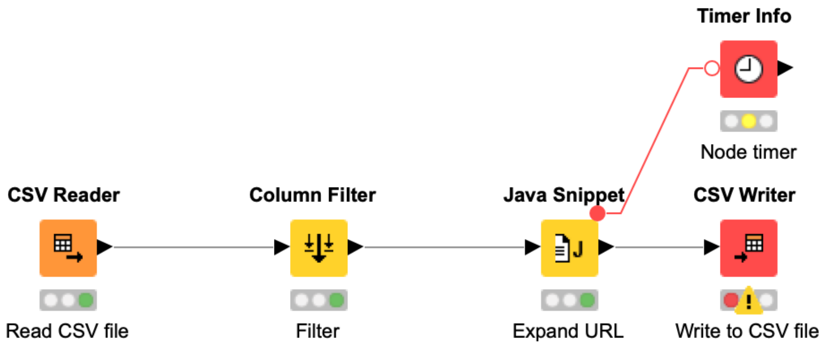


Figure 51 - Expand URL

Many of the URLs identified are short URLs, which means that The URLs have been though a URL shortener service. When analyzing URLs, we need a complete URL. To combat this, we used the workflow in figure 51 to expand and get the original URL. The list of URLs in tweets is sent to the expand URL node.

```
URL url = null;
    try {
        url = new URL(c_Text);
    } catch (MalformedURLException e) {
        e.getMessage();
    }
    // open http connection
    HttpURLConnection httpURLConnection = null;
    try {
        httpURLConnection = (HttpURLConnection) url.openConnection(Proxy.NO_PROXY);
    } catch (IOException e) {
        e.getMessage();
    }

    // stop following browser redirect
    httpURLConnection.setInstanceFollowRedirects(false);

    // extract location header containing the actual destination URL
    String expandedURL = httpURLConnection.getHeaderField("Location");
    //replace the short URL with expanded URL
    out_Text = expandedURL;
    //close connection
    httpURLConnection.disconnect();
```

Figure 52 - Expand URL node

The expand URL node creates a new instance of the URL class and passes it to the URL in the tweet. After we open an HTTP connection, we call the openConnection method that returns a URLConnection instance that represents a connection with the given URL. We then create a String variable and assign it to getHeaderField. If there are no getHeaderfield it returns null. The expanded URL is then passed to the dataset.

Appendix B: Hardware, Software, and data

Throughout this study, we used a dataset of Twitter data, with over 14.3 million entries. The dataset cannot be shared due to Twitter Developer Terms, privacy regulations, and our contract/application with NSD (Norsk Senter for Forskningsdata – Norwegian center for research data).

The main software used for data preprocessing and data analysis was KNIME. KNIME, the open-source data analytic tool, can be download from the KNIMEs website here: https://www.knime.com/downloads.

To analyze the dataset, we used two computers equipped with the security standards set by NSD. The large dataset required computers with enough RAM and a fast CPU. The first computer is equipped with 24GB DDR4 RAM, Intel Core i7-7700 2.8GHz CPU, and Nvidia GeForce GTX 1050 GB GDDR5 VRAM. The second computer is equipped with 24GB DDR3 RAM, Intel Core i5-2500k CPU, and Nvidia GeForce GTX 750 Ti.

Appendix C: Structural model and hypothesis testing from the previous study by the group

The group conducted a survey last semester related to why people share fake news on social media. Its primary purpose was to after trail and expand on a previous study (Talwar et al., 2019), but ended up with inconclusive results. While there was some novelty found in the expanded theory brought in and tested within the study (technical operations and creation of content), it ultimately ended up illustrating some weaknesses to both the method and type of study when trying to answer these questions. It is in this context used as argumentation that one has to study what people do, as what people say they do, and the action they perform often misaligns (Hair et al., 2011). Please ignore figure number here (10 and 11), as these are brought in from the previous report (Lerbæk & Olsen, 2019b, p. 14).

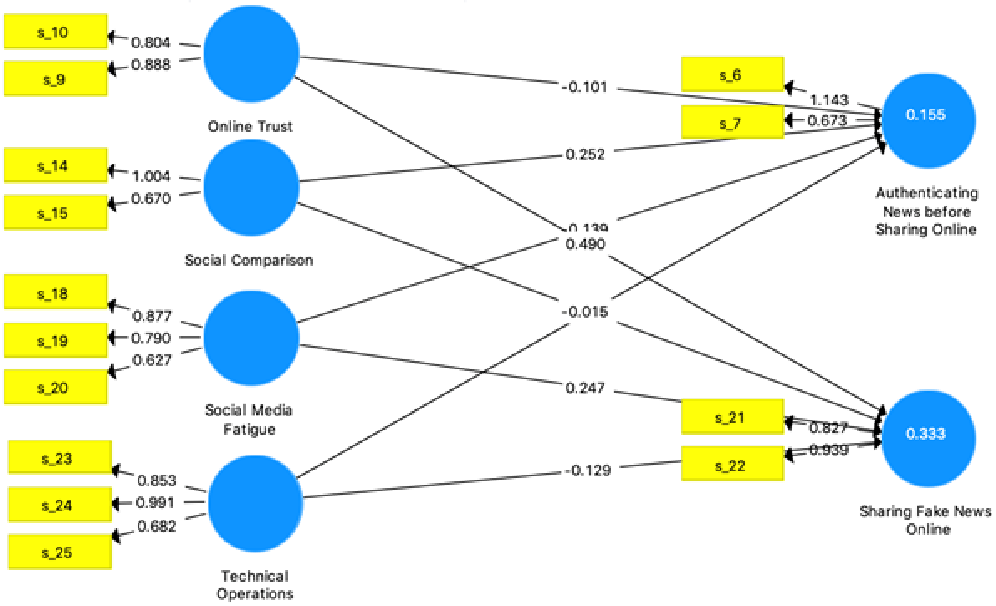


Figure 10 - Structural model edited

As seen in figure 8, the model displayed both low r-squared (R^2) values in combination of high cross loading values, which led us to test running a new consistent PLS algorithm test with model illustrated in figure 9. Compared to the model used in rest of the results this yielded little changes. *Authenticating news before sharing online* gets R^2 results of 15,5% (16,3% otherwise) while *Sharing fake news online* gets R^2 results 33,3% (33,7% otherwise).

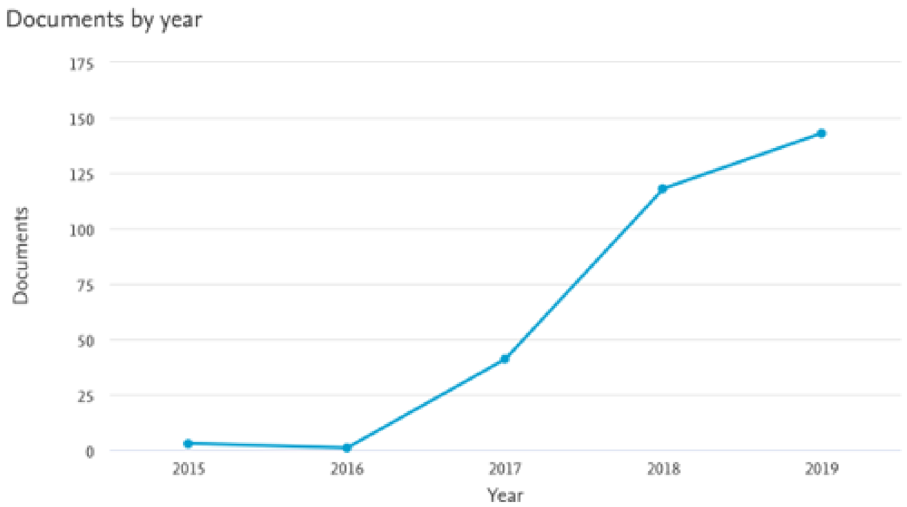
	Original Sample (C	Sample Mean (M)	Standard Deviatior	T Statistics (|O/ST	P Values
Creation of Content and Knowledge -> Authenticating News before Sharing Online	0.036	0.050	0.120	0.297	0.767
Creation of Content and Knowledge -> Sharing Fake News Online	-0.020	-0.022	0.090	0.222	0.824
Fear of Missing Out -> Authenticating News before Sharing Online	0.067	0.077	0.117	0.575	0.565
Fear of Missing Out -> Sharing Fake News Online	0.054	0.041	0.095	0.568	0.570
Online Trust -> Authenticating News before Sharing Online	-0.084	-0.079	0.095	0.879	0.379
Online Trust -> Sharing Fake News Online	0.405	0.401	0.097	4.189	0.000
Self-disclosure -> Authenticating News before Sharing Online	-0.127	-0.135	0.108	1.173	0.241
Self-disclosure -> Sharing Fake News Online	0.081	0.094	0.107	0.758	0.449
Social Comparison -> Authenticating News before Sharing Online	0.228	0.215	0.124	1.839	0.066
Social Comparison -> Sharing Fake News Online	-0.010	-0.011	0.115	0.092	0.927
Social Media Fatigue -> Authenticating News before Sharing Online	-0.130	-0.121	0.119	1.087	0.277
Social Media Fatigue -> Sharing Fake News Online	0.194	0.204	0.099	1.957	0.051
Technical Operations -> Authenticating News before Sharing Online	-0.239	-0.234	0.100	2.400	0.016
Technical Operations -> Sharing Fake News Online	-0.124	-0.130	0.102	1.211	0.226

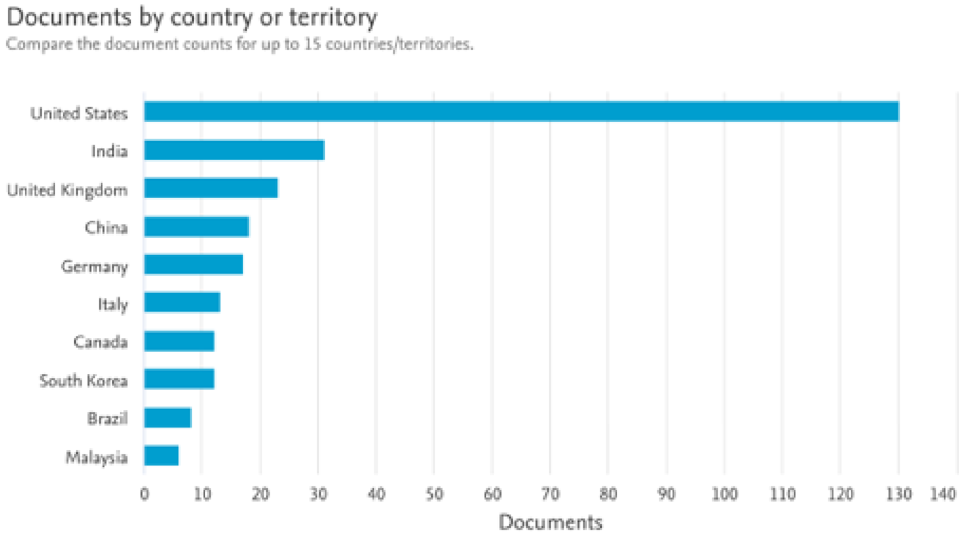
Figure 11 - P-values

Appendix D: Original literature review (autumn 2019) metadata

From the previous year's literature review, the following metadata about the literature review can be illustrated. It shows the rapid increase of the research on *fake news* and the overweight of American research on the subject.

4.2 Metadata





TITLE-ABS-KEY ("fake news") AND ("social media" OR "social network" OR "social networking") AND ("spread" OR "scope" OR "reach" OR "grow" OR "increase" OR "expand" OR "advance" OR "disseminate" OR "propagation") AND (LIMIT-TO (SRCTYPE , "j") OR LIMIT- TO(SRCTYPE , "p")) AND (LIMIT-TO(PUBSTAGE , "final")) AND (LIMIT-TO(DOCTYPE , "ar") OR LIMIT-TO(DOCTYPE , "cp")) AND (LIMIT-TO(SUBJAREA , "COMP") OR LIMIT-TO(SUBJAREA , "ENGI") OR LIMIT-TO(SUBJAREA , "DECI")) AND (LIMIT- TO(PUBYEAR , 2019) OR LIMIT-TO(PUBYEAR , 2018) OR LIMIT-TO(PUBYEAR , 2017) OR LIMIT-TO(PUBYEAR , 2016) OR LIMIT-TO(PUBYEAR , 2015)) AND (LIMIT-TO(LANGUAGE , "English"))

Appendix E: Concept matrix, further details

Within the concept matrix, work was performed to identify and easier visualize the occurrences of overlapping concepts and concepts in relation to each other. Below there are four snips of these overlaps. Its idea is to illustrate the different main focuses of the literature brought in in the literature review. The way these are structured is to make statements/queries in excel to count occurrences where other concepts are present (pseudocode: count X where Y and Z are present). As seen here, we see a clear dominance of Twitter being the platform of study, fake / false news being the main concept studied, and crisis is the context (also a concept) just a bit ahead of disaster.

Summary no. fake news		Summary no. Social Media Platforms		Summary no. disruption type	
Fake news / fake stories	25	Twitter	39	Crisis	11
Conspiracy theories	5	Facebook	13	Emergency	1
Misinformation	19	Linkedin	1	Disaster	9
False news / information / claims	15	Reddit	3		
Disinformation	14				
Rumors / rumor cascades	15				
Tweet / retweet cascades	3				
Hoax	5				
Concept in context of	Crisis	Emerency	Diaster		
Fake news / fake stories	3	0	2		
Conspiracy theories	0	0	0		
Misinformation	3	0	3		
False news / false information / Claims	2	0	1		
Disinformation	1	0	1		
Rumors / rumor cascade	6	0	4		
Tweet / Retweet cascades	0	0	2		
Hoax	0	0	1		
Summarized context	15	0	14		
Concept in context of	Twitter	Facebook	LinkedIn	Reddit	
Fake news / fake stories	23	13	1	3	
Conspiracy theories	4	5	0	3	
Misinformation	17	9	1	3	
False news / false information / Claims	14	8	0	3	
Disinformation	9	6	0	3	
Rumors / rumor cascade	14	5	0	3	
Tweet / Retweet cascades	3	1	0	0	
Hoax	4	5	0	3	

Figure 53 - Concept distribution

Appendix F: Tag cloud results, 20 topics 10 terms

While the thesis ultimately ended up using 10 topics, 10 terms result from the topic detection LDA workflow, tests were conducted with the sample size of 20 topics, 10 terms as well. The top results remained somewhat the same as with 10 topics 10 terms, although the increased topic doubles the term. The results of these tests are seen in figure 48 below. The size of the term is dependable on the frequency of the term (bigger - more frequent), and the color corresponds to the various topic.

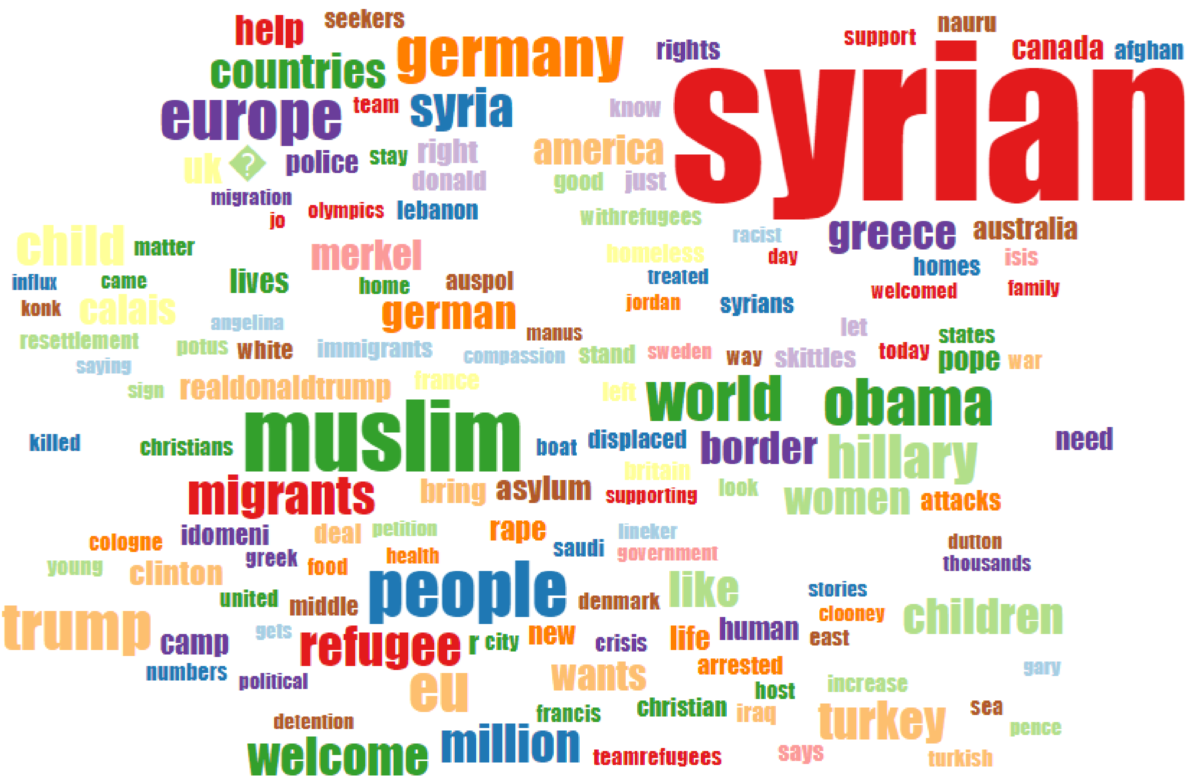


Figure 54 - Tag Cloud based on LDA analysis, 20 topics 10 terms

Appendix G: Topic and term weighting

Topic	Term	Frequency	Topic	Term	Frequency	Topic	Term	Frequency
topic_1	syrian	1.88%	topic_5	syrian	2.68%	topic_8	syrian	1.08%
topic_1	help	1.64%	topic_5	obama	2.11%	topic_8	australia	0.94%
topic_1	refugee	0.75%	topic_5	muslim	1.74%	topic_8	trump	0.71%
topic_1	world	0.72%	topic_5	christian	0.67%	topic_8	canada	0.67%
topic_1	new	0.71%	topic_5	amy	0.57%	topic_8	skittles	0.66%
topic_1	support	0.62%	topic_5	mek	0.57%	topic_8	auspol	0.56%
topic_1	work	0.51%	topic_5	america	0.55%	topic_8	nauru	0.53%
topic_1	stand	0.48%	topic_5	r	0.45%	topic_8	people	0.48%
topic_1	withrefugees	0.47%	topic_5	state	0.43%	topic_8	just	0.43%
topic_1	today	0.45%	topic_5	christians	0.42%	topic_8	refugee	0.42%
topic_2	welcome	2.29%	topic_6	syria	1.75%	topic_9	eu	2.04%
topic_2	syrian	0.90%	topic_6	million	1.71%	topic_9	child	1.83%
topic_2	people	0.58%	topic_6	syrian	1.53%	topic_9	uk	1.53%
topic_2	?	0.49%	topic_6	world	1.23%	topic_9	calais	1.31%
topic_2	love	0.39%	topic_6	war	0.87%	topic_9	europe	1.16%
topic_2	today	0.35%	topic_6	displaced	0.80%	topic_9	turkey	0.99%
topic_2	just	0.35%	topic_6	turkey	0.76%	topic_9	migrants	0.93%
topic_2	say	0.33%	topic_6	people	0.76%	topic_9	refugee	0.76%
topic_2	like	0.30%	topic_6	countries	0.68%	topic_9	asylum	0.76%
topic_2	brussels	0.29%	topic_6	lebanon	0.63%	topic_9	britain	0.61%
topic_3	syrian	2.77%	topic_7	germany	2.66%	topic_10	greece	2.47%
topic_3	hillary	2.72%	topic_7	muslim	1.70%	topic_10	border	1.60%
topic_3	trump	2.31%	topic_7	german	1.61%	topic_10	migrants	1.24%
topic_3	wants	1.64%	topic_7	merkel	1.27%	topic_10	europe	1.08%
topic_3	america	1.32%	topic_7	rape	0.96%	topic_10	syrian	0.97%
topic_3	muslim	1.08%	topic_7	europe	0.83%	topic_10	pope	0.93%
topic_3	obama	1.08%	topic_7	women	0.74%	topic_10	greek	0.80%
topic_3	realdonaldtrump	0.99%	topic_7	isis	0.72%	topic_10	refugee	0.72%
topic_3	clinton	0.94%	topic_7	police	0.67%	topic_10	sea	0.66%
topic_3	increase	0.91%	topic_7	attacks	0.67%	topic_10	idomeni	0.63%
topic_4	people	1.15%						
topic_4	europe	1.06%						
topic_4	like	0.82%						
topic_4	want	0.81%						
topic_4	country	0.65%						
topic_4	countries	0.63%						
topic_4	women	0.61%						
topic_4	need	0.60%						
topic_4	isis	0.59%						
topic_4	syrian	0.56%					Total =	30 373 196

Figure 55 - Overview of all the term weighting