Vrije Universiteit Amsterdam   Universiteit van Amsterdam

Master Thesis

The Islamic Elephant in the News Room

Using Natural Language Processing to Uncover Stereotypes of Muslims and Ethnic Minorities in Norwegian News in 2017-2019

Author: Laréb Fatima Ahmad   VU:2737338 | UvA:14118114

1st supervisor: Anna Bon

A thesis submitted in fulfillment of the requirements for
the joint UvA-VU Master of Science degree in Computer Science

August 29, 2024

" I am the Islamic elephant in your room. I am a suspicion, I am a mascot, I am politics"

from *Den Islamske Elefanten*, by Karpe

Abstract

Previous research has revealed that the population in Norway broadly has a negative and stereotypical impression of Muslims and ethnic minorities. Surveys conducted on Muslims and non-Muslims living in Norway reveal that the population sees news media as a significant source for the negative impression. There is little to no technical research conducted on how mainstream Norwegian news media writes about these minority groups. Existing research focuses on alternative news media sources, or takes a manual approach to analyzing news articles using frameworks from social sciences.

The aim of the thesis is to investigate if stereotyping of Muslims and ethnic minorities is detectable in Norwegian news. To this extent, I investigate what roles, labels and properties are often prescribed to the two. The collected descriptions are compared to those of Norwegians, Christians and other theoretical opposites.

Natural language processing was used to investigate how Muslims and ethnic minorities are described at an in-depth level. The data used in this study is comprised of articles from various newspapers, from the years 2017 to 2019.

The results revealed that the majority of the descriptions revolved around terrorism, warfare and the terrorist organization ISIS. Notably, Muslim women are also spotlighted when related to controversy. Articles about Muslims tend to default to stereotype-congruent language, for example describing women as "brides", and insinuating Islam hinders women's ability to communicate. Descriptions of their "opposites" are on the other hand positive or neutral, revolving around sports and trends.

From the results and the subsequent analysis, one can detect that there is a disproportionate reportage practice. Articles about Muslims and ethnic minorities largely revolve around war and extremism, and individuals are described in ways that align with biases.

Acknowledgements

Contents

CONTENTS

CONTENTS

List of Figures

LIST OF FIGURES

List of Tables

LIST OF TABLES

1

Introduction

Growing up in Norway as a child of Muslim immigrants from Pakistan, the study of this thesis is one that is close to my heart. My Muslim peers and I were always staunchly different than our classmates in the early 2000s and 2010s. Feeling like Norwegian society held a close eye on us, we did our best to not only fit in, but to oppose the stereotypes of Muslims who refuse to integrate and with a potentially dangerous set of beliefs. The news landscape in the aftermath of the 9/11 attacks on the World Trade Center, and during the rise of ISIS, Islamist extremism and far-right extremism in Western Europe, was a landscape that kept us on our toes constantly. Having to explain why our beliefs "encouraged" terrorism in class, or be the recipient of an individuals anger following ISIS-claimed attacks in Europe, and to see your own ethnic background explicitly written out in news cases of crimes and violence. These are experiences which contributed to consistently feeling like not being "Norwegian enough", opposed to the norm. As I may be perceived to speak on behalf of peers I share beliefs and backgrounds with, it feels as if we are not individuals, as if we do not have agency but rather that we are all motivated by what we share.

Can the root causes of these collective feelings of otherness of Muslims be excavated? Does Norwegian news media play a role in introducing bias and stereotyping? And if so, is it possible to objectively analyze and reveal this stereotyping in Norwegian news articles? The objective of this research is to analyze news texts from online news sites,. The paper attempts to answer the following research question:

Is stereotyping of Muslims and ethnic minorities detectable in Norwegian news?

1. INTRODUCTION

To answer this question I collected a large corpus of text. Using Natural Language Processing (NLP), this text was cleaned and processed, and syntactical data was extracted to collect descriptions of Muslims in the news. I developed this pipeline in Python code, which output a list of the most frequent descriptions.

After analyzing the descriptions, I composed meaningful insights that represent the image of Muslims in the news. With this, I will also answer the following sub-questions:

- *Can this be objectively answered?*

- *What method is the most suitable?*

This paper is organized into a sequence of chapters that provide a comprehensive overview of the study. The paper begins by establishing the context in which this research takes place, followed by a review of relevant literature that informs the study. The next section provides an explanation of the data used in the study, as well as the technical background that underpins the research. A detailed methodology chapter outlines the specific approaches and techniques employed in the study. The results of the study are then presented, followed by a discussion of their implications and significance within the context of news reporting. This discussion also addresses the challenges, limitations, and restrictions encountered during the study. Finally, the paper concludes with a summary of the key findings, and suggestions for future research.

I want to make the usage of certain terms throughout this thesis clear. In this thesis, I will use the word "news sites" and "newspaper" interchangeably for online news media sites, unless explicitly stated to be physical newspapers. Additionally, the news media that are discussed in this study is exclusively Norwegian news media, unless stated differently. Lastly, this thesis assumes that the perceptions of immigrants in Norway largely reflect the same sentiments as those towards Muslims.

Mainstream news media refers to established and well-known newspapers, that tend to be a part of larger media organizations. They cover a wide variety of topics, and follow a regular publishing schedule. Mainstream news media publish regional or national news articles, essays, and debate posts (also called a short essay). Essays and debate posts are written by individuals who do not work for or represent the newspaper, but people from the outside who are usually knowledgeable about the given topic.

These newspapers adhere to a neutral reporting of events in terms of the language in the articles, and usually upholds an expected degree of political correctness.

2

Context Analysis

This chapter firstly looks over the history of Muslims arriving in Norway and the news media landscape at that time. Section 2.1.1 describes the statistics of Muslims in Norway in 2017, 2018 and 2019, aligned with the years of the articles analyzed in this research. Section 2.2 presents results from studies conducted in Norway about stereotypical beliefs that people in Norway may hold about Muslims, along with the perceived reasons of these beliefs. The last section will describe how Muslims are depicted in Norwegian news media, and the experiences of Muslims in Norway with the news media. This section ends with the experiences of respondents interviewed and surveyed by me.

2.1 Brief History of Muslims in Norway

Muslim immigrants arrived in Norway first in the 1970s. They came from countries such as Pakistan, Morocco, India and Turkey. It was around this time that Norway had discovered large quantities of oil, and thus had opened its borders to migrant workers from the Global South. They were hired to help rebuild the nation following the war. Nearby nations such as Germany and France had already closed their borders to migrant workers at this time (37, 147).

The news media landscape at this time covered world events which concerned the rise of conservative Islam, and Islam as a state religion in many countries, particularly in the Middle East (21, 147). Thus, the rise of Islam was often referred to as a key part of world politics. In the 1970s, news articles about unrest in countries such as Iran, Libya and Egypt would often represent Islam as traditional and Muslims as fanatics, accompanying such statements with descriptions of atrocities executed in the name of the religion, by

dictators such as Gadafi in Libya and Sadat in Egypt. In one article, the rise of Islam in world politics was described as *"Islam på fremmarsj"* ("Islam on the march"), a comparison to soldiers preparing for an attack or invasion. In 1979, an article titled *"Islamfrykt i vesten"* ("Islam-fear in the West") was published, detailing the Western world's fears on growing Islam with dictators as leading figures (21).

In Bollingberg's thesis "Muslimer i Aftenposten på 1970-tallet", he analyzed Muslims in one Norwegian news outlet in the 70s. In many articles, Muslims and Islam are compared to Christianity and Christians, describing Islamic prayer as similar to Christian mass, and imams as priests. As much as these comparisons helped to familiarize the Norwegian public with the Muslims newly introduced to the country, they also emphasized that Muslims are separate from Norwegians. There was a focus on all the things which Muslims do not do, as compared to Norwegians, such as the Islamic abstaining traditions and women exposing their hair. Aspects of Sharia law was also zoomed into, such as violent punishments and death sentences (21).

The media surrounding Muslims and Islam when the first Muslims immigrants arrived in Norway was describing them as similar in some aspects, but yet starkly different, particularly in aspects such as views of women, abstinences, and Islam as somewhat violent in its dogma. Taking all this into consideration, one must consider that all in all, Muslims and Islam were discussed sparsely in the media in the 70s. However, when discussed, the articles often spotlighted dictators in Muslim majority nations such as Egypt and Iraq, and insinuated that all Muslims have the same beliefs and follow the same practices.

2.1.1 Statistics of Immigrants in Norway

The Muslim population in Norway is a minority that has existed in the country for roughly 40-50 years. The majority of Muslims in Norway have immigrated from or have parents who immigrated from Pakistan or Somalia, the rest have background from other Muslim-majority countries such as Iraq, Iran, Turkey, Syria and Afghanistan, and a minority are converts to Islam. According to Statistics Norway *(Statistisk sentralbyrå)* there were between 148 000 and 250 000 Muslims in Norway at the start of 2016 (103). However, the total number is difficult to establish for a number of reasons, such as many with backgrounds from Muslim origin countries may not identify as Muslim and many may not register officially as a member of a mosque for a variety of reasons. The top 10 countries of origin of immigrants in Norway and Norwegians with immigrant parents is listed in tables 9.1, along with the number of people from each country, for the years 2017 to 2019. Out

of the top countries of origin, four of them have a majority Muslim population, namely Pakistan, Somalia and Iraq along with Syria in 2019 (104). Syria enters the top 10 as refugees from the country come to Norway.

2.2 Stereotypes and Perceptions of Muslims and ethnic minorities in Norway

Participants in a street study conducted by Bye et al. in 2014, 244 people were asked to rate how some given social groups in Norway are perceived by the Norwegian general population (24). Groups were judged on warmth and competence, status and competition. People with perceived high status and competitiveness are deemed cold, but competent. With low status instead, they are deemed cold and incompetent. On the flip side, with perceived high status and non-competitiveness, people are deemed warm and competent. With low status, they are deemed warm, but incompetent. To measure warmth and competence, respondents rated the groups' perceived friendliness, warmth, kindnatured-ness, sincerity, competence, confidence, capability and skillfulness, on scale of 1 to 5, where 5 indicated "to a very large extent". Measuring the groups' status consisted of participants being asked to think of how the groups are perceived on a broad level to have qualities such as prestigious jobs, economic success and good education. Lastly, competition was measured by subjects comparing which groups were perceived to receive more or less resources than others, which groups loses power when another gains power, and which groups when getting preference over others, makes it harder for the other groups. On average, the social group of Muslims and immigrants were judged by the respondents to be perceived by the Norwegian public as moderate in warmth and to have low competence. Within the group of immigrants, those with background from Muslim countries were disproportionately negatively perceived and were ranked moderate to low on competence and warmth.

In 2017 and 2022, The Norwegian Center for Holocaust and Minority Studies (HL-center) (42) conducted a broader survey mapping and analyzing the perceptions and attitudes of Muslims, Jews and the population as a whole against Muslims and Jewish people (55, 85). For the relevance of this thesis, only the parts focused on Muslims will be highlighted. The Muslims included in the survey were people who had either immigrated themselves or their parents had, and who had lived in Norway for longer than 5 years. A majority of the respondents had backgrounds from Pakistan, Kosovo, Morocco, Iraq, Iran, Bosnia-Herzegovina, Afghanistan, Palestine, Turkey and Somalia. While 23150 Muslims were

asked, only 1305 responded, and this small group of respondents has to be taken into account when discussing the answers. 135 Jewish people responded, and 1653 people from the rest of the population responded. One must also take into account that these are simply answers to statements on various given scales, and one cannot conclude that the answers to these statements correlate to how the broader population treats Muslims in their day to day lives.

In the survey, respondents were to answer how much they agree or disagree with given statements and answer certain questions, which represent stereotypical beliefs about Muslims. Respondents answered on a scale of 1 to 5 or 6, from "totally disagree" to "totally agree" or "dislike it strongly" to "like it strongly". Table 9.2 presents the statements and questions used in the survey, and what percentage of the participants believed in the negative stereotype.

Respondents from 2017 typically cited "fear of terror" and "fear of extremism", "negative representations in news media", "xenophobia" along with "Muslim people's lacking will to integrate", "violence", "Islam" and "oppression of women" as reasons for why the population at large may have negative attitudes against Muslims in Norway. Another thing of note is that in interviews conducted, Muslim respondents frequently referred to themselves and Norwegian people as separate, and as not something they are a part of (55).

The 2022 report detailed the respondents answered more extensively. The majority of reasons given for negative attitudes towards Muslims included the words "prejudice", "lack of knowledge/insight" and "xenophobia", according to the broader population (including Jewish and Muslim respondents). Many of them referred to a lacking ability to differentiate between extremists and "normal" people, and little desire to understand other cultures and beliefs among Norwegian society. From the Muslim respondents, words such as "misinformation" and "false information" were often used in explanations and referred to scarce or wrong knowledge about the religion. Other explanations include generalizations of Muslims, for instance using wording which refers to "Muslims" as a whole instead of the individual or group in question. Moreover, many explanations include a general negative attitude towards religion in Norwegian society, as well as a tendency to contradict religion to "modern" and "Western" societies (85).

When Muslims are visibly religious, for example when they are seen wearing a headscarf or hijab, dressing modestly (something which stands out in non-Muslim majority societies particularly during summers) or grow a beard, they stand in "contrast" to modernity, and thus become the "they"-group to the Norwegian notion of "us" (61, 140). In addition

to outwardly different, Muslim immigrants are religiously different, and often come from a background which is ethnically and culturally unlike than those found in Norway and Western Europe (16).

Another major reason given for the negative attitude towards Muslims included the word "terror", referring to terrorist attacks outside of Norway. Negative attitudes towards Muslims in Norway correlate to incidents outside of Norway, usually conflicts in the Middle-East (55, 67, 85). Respondents referred to specific terrorist organizations such as ISIS, Al-Qaida and Taliban and their violent actions, and reports on their actions used as reasons to generalize all Muslims, as the terrorists themselves claim to represent the religion. Notably, the respondents also mentioned the 9/11 world trade center attacks, something many studies point to as a leading reason for Islamophobia in the West (45, 69, 73, 140). Many refer to terrorists as "ruining the reputation of Muslims". Furthermore, corrupt regimes in Muslim-majority countries are also pointed out in the answers, such as Iran and Saudi-Arabia.

Throughout all the respondent groups, most pointed to the media as a reason for prejudice towards Muslims. One common response correlated to the lack of knowledge of other religions, in addition to the media mostly representing and talking about Muslims and Islam in the context of religious extremism and Middle Eastern conflict. This leads to people filling the empty gaps of their knowledge of Islam with what they see on the news. The issue is the same with attitudes towards immigrants, where the news media mostly highlight crime, as news media does, and crime is most prominent in the biggest cities in Norway, the same cities with the highest populations of immigrants. Many respondents in 2022 pointed to the political far-right in Norway, such as organizations like the anti-Muslim Stopp Islamiseringen av Norge (SIAN) (127), the political party FrP ("People's party") (47) and the alternative, far-right and conservative news media site Resett.no, which was shut down in 2022 (62). A majority in 2022 believed that Muslims in the media are over-represented by extremists. Around 60% of the answers to this question refer to the news media being problematic and contributing to prejudice. When asked why the respondents believe "normal" or average Muslims are not a larger part of the representation of Muslims in the media, the majority answered that this is related to news outlets' focus on publishing articles which "sell" the best. Stories about Muslims everyday life, not being much different from the average Norwegian, may not be something to write news about (5, 61, 69).

Overall, the majority of the respondents were not in favor of the (negative) statements in both studies. Furthermore, there was a decrease from 2017 to 2022 in the percentage

who supported the negative statements. The exception of this was the statement "Muslims oppress women", which the majority of the respondents agreed with. Still, a significant amount of the respondent believe in and supported negative statements against Muslims. The report concluded that almost one third of the population supported widespread prejudiced attitudes towards Muslims.

2.3 Muslims in Norwegian News Media

The Fritt Ord Foundation (2017) (46) published an analysis of Norwegian news media from 2016 (116). It showed that news about Islam and about Muslims were the most common themes that year. Of the articles covering Islam and Muslims, 11% were overwhelmingly negative. More than half of the articles about Islam and Muslims cover news cases in other countries, in which more than 60% of these cover terror attacks, radical Islamism and ISIS. 16% of the articles about Islam and Muslims cover criticism to Islam and are usually sent in by citizens rather than written by journalists. Most of these articles depicted Islam as a threat to Norwegian culture. The analysis revealed that ISIS and Islamist terror attacks were the focus of most of the news coverage about Muslims and Islam in Norwegian news media in 2016. However, the media itself does not equate Islam and Muslims to terrorism.

On the other hand, Muslims are depicted as a single organism or group particularly when talking about religious traditions or practices (113, 116). 85% of the coverage is revealed to be neutral in and of itself, but that does not mean that it cannot be regarded as wholly negative by individuals. For instance, individuals may make a connection between Muslims and terrorism even after consuming language-wise neutral media (113). The report concluded that Islam is depicted in a nuanced way, and Islam criticism is met with people challenging the criticism. Both sides have equal and neutral coverage. However, past studies have revealed that most news coverage about Muslims tends to surround either extremists and terrorism, or very liberal Muslims, both of which subvert expectations of Muslims in some way (45, 61).

2.3.1 Effects of News Media Coverage

Potential or perceived bias in news articles which cover topics about Muslims people can have real life and dangerous consequences. As reported by the HL-center and The Centre for Equality and Diversity (KUN) study in 2022(41), the comment section on such news articles tend to include severely negative and often blatantly discriminatory statements directed towards Muslims. Some informants emphasized the effect hateful comment sections

can have on young Muslims and children. Studies have shown that those who engage in internet debates are less than 6% of the adult population in Norway (88). However, the impact of the negative comments have a large span and can be perceived as particularly penetrative and seem to be guaranteed under a news article covering a crime conducted by someone who happen to be Muslim (84). Respondents made connections between what is written in news media and how this affected their daily lives. For instance, some feel unsafe in public and like they need to make themselves invincible, to "be on their best behavior", particularly in the aftermaths of larger news cases where the perpetrator may have been Muslim. Many feel as if it is their responsibility to prove to society that the perpetrator is an outlier and to make active efforts to distance themselves from the representations of Muslims in news media (39, 69, 73, 140). This ties back into the generalization effect the news media has on the larger population perception of Muslims, but also the perception of Muslims by others (84, 113). The Equality and Anti-Discrimination Ombud (LDO) conducted in 2018 a study analzing hate speech in the editorial-controlled comment sections of the Facebook and Twitter pages of two Norwegian news outlets; NRK and TV2. The study revealed that 7% of the comments were hate speech, mostly under news of refugees and migrants, in which ethnic minorities, Muslims and politically active people were mostly targeted (146).

2.4 Norwegian Muslims' Experiences with News Media

The HL-center's reports from 2017 and 2022 highlights that Norwegian news media, to a large extent, stopped the practice of listing a perpetrators ethnicity or religious beliefs, following the terrorist attack in Norway, 22nd July 2011, where the attacker was a Norwegian far-right extremist (55, 85). Contrary to this, several respondents from the 2022 study referred to their experience about the tendency of news media to highlight a persons ethnic background or religious affiliation, when the news are negative, such as referring to a perpetrator as "The Muslim man". Others referred to the tendency to explicitly mention religion even when the perpetrator may be severely mentally ill. On the contrary, white perpetrators often have their psychological medical history highlighted instead (85). This is evidenced by Jarmer's findings in his analysis of articles from established news site Verdens Gang (VG), from 2017 to 2019. In his research, he noticed a tendency to *not* use the word "terror" when describing mass shootings conducted by white, non-Muslims. However, the word "terror" appears frequently in news about Islam (61).

In another study conducted by the HL-center and KUN (41) in 2022, interviews followed analyses were conducted (84). The reference group for this study were representatives of Muslim organizations in Norway, academics and resource people relevant to the quality of life of Muslims in Norway. These include but are not limited to the Muslim organizations Islamic Culture Centre Norway (ICC) (101) and Islamsk Råd Norge (IRN) (90) and representatives from religious communities. The interviewees were Muslims born and raised in Norway and Muslims who had immigrated to Norway and lived there for 5 or more years. Subjects were geographically from all over Norway. Out of the 23150 people asked, 1305 people replied.

Several respondents reported having experienced an overwhelmingly negative representation of Muslims in the media, often related to terrorism and the political situations in other countries. Several studies argue that Islam is indeed portrayed more negatively in media than other religions (5, 113, 140). Islam frequently is written about in such a way that the norm is oppression of women, violence and terror, and anything else is outside of the status quo, traits that are "as opposed to the norm", or need to be proven wrong, but are still the point of reference (69, 140). This aligns with previous research done on linguistic bias and particularly stereotype congruent and incongruent behavior, which is typically worded in such a way that reinforces stereotypes by emphasizing the stereotype-incongruent behavior as atypical (19, 53, 149).

In the same study, respondents described the experience of news media particularly highlighting a persons "foreign" background when they have done something negative. However, people with the same background are described as Norwegian only when the news are positive. Statements such as these were common in this study, both among Muslim and non-Muslim respondents (69, 84). Moreover, news media coverage of far-right politicians who propagate xenophobia and Islamophobia were mentioned. Lastly, misinformation and generalizations made by the wording news outlets choose to use in news about Muslim people. When asked "Do you experience that the way Islam is portrayed in the public is prejudiced?", 57.4% of Muslim respondents answered "often" and 25.2% answered "sometimes" (84).

2.4.1 Interviewing Norwegian Muslims About Their Experiences

In the context of gathering data for this thesis, I conducted an anonymous survey and some interviews of Muslim people in Norway. As there were very few respondents in the interviews and the survey (8 in total, 5 in the survey and 2 in the interviews), and the answers from the respondents were often irrelevant (not about experienced with news

media) or included little details, the data I collected was not used in the methodology. The answers were irrelevant as they were collected early on in the process of conducting this research, when I had yet to formulate the research questions. Additionally, some of the questions do not relate to this study. I started off the early stages of this research broadly, not focusing in on Norwegian news until later, and the interviews and surveys were conducted early on.

However, I believe it is meaningful to include some of their answers as an insight into real experiences with discrimination from Muslim people in Norway, albeit just a handful of people. All the questions from the interview and survey are not included due to their irrelevancy, but the questions to the answers discussed below are stated.

To the question *"Have you experienced discrimination in text, reading e.g. in news or TV programs?"*, one respondent described that they felt the media misuses their power and perpetuates prejudicial believes especially about minority groups and Muslims. The respondent emphasizes the comment section on news articles being overrun by hate speech, and the experience of a disproportionate amount of hate speech in the comment sections of articles about someone with a minority background, versus comments under articles of someone with a majority background. As this respondent is politically active on Facebook, they also point to the excessive hate speech in the Facebook comments of news articles too. These descriptions particularly align with the findings of the study on comments under posts of a large Norwegian news broadcaster's Facebook page, as described in section 2.3.1 (146).

I asked the respondent about their experiences with discrimination, and all of them described it as negative and disheartening. One respondent experienced that the events covered in the news directly had an effect on how they were treated by their colleagues. They describe more negative profiling, being gossiped about, unfair treatment, and being met with prejudiced and stereotypical remarks about Muslims. Such a news event was for instance the coverage of an terrorist attack in the name of Islam. Another respondent described feeling the need to be on your toes after the events of the 22nd July terrorist attacks in Oslo in 2011, before it was known that the attacker was a white, Norwegian far-right extremist. During this time, people were quick to blame Muslims. The same respondent also mentions seeing a lot of hate speech against Muslims on the internet, such as on Facebook, before the perpetrator's identity was know.

3

Literature Review

Using NLP to analyze text is not a novel thing. In fact, many studies have used various NLP and machine learning procedures to attain meaningful conclusions from texts on various subjects, including on racial biases, in multiple languages.

This chapter of the paper attempts to summarize a collection of articles and papers conducted on related topics in order to showcase existing research. More specifically, this section first describes stereotyping and bias by looking into previous research on the topic, describes other works done using NLP techniques to detect racial biases in texts, other research using microportraits extraction, and lastly analyses done on Muslims' and Islam's depiction in mainstream Norwegian news media. The chapter ends with my explanation for choosing microportraits, and the results I expect to see based on the literature review. Finally, I denote my contribution to the research and beyond existing studies.

3.1 Stereotypes and Bias

Delgado-Rodríguez and Llorca (2004) describe bias as a wrongful assessment of one's experiences and exposures as a fact or truth applied to the broader population, or as a lack of internal validity, and as something that is separate from a random error (33). Beukeboom (2017), a co-founder of the Micrportraits framework (40), denoted linguistic bias as systematically specific ways to describe social groups with consistent word or semantic choices (17). Social categories or groups and stereotypes help to simply the processes of perceiving people, and people within the same social groups tend to hold the same beliefs about other social groups (17, 81). Negative stereotypes tend to materialize as a consequence of prejudice and discriminatory associations about social groups (17, 122).

The Social Categories and Stereotypes Communication (SCSC) framework was proposed by Beukeboom and Burgers (2019) for analyzing how stereotypes that are applied to social categories, such as "immigrant" or "Muslim", flow and grow in language (18). The authors studied how stereotypes turn into knowledge: through repetitive exposure to the the same labels directed at the same social groups, or particular characteristics of those who belong to a specific social group. Especially when these characteristics are seen as fundamental to a social group, for instance a veil on a woman "guarantees" a Muslim, then the labels associated with a veiled Muslim woman are likely to be conserved. People are labeled according to the social group they are put in, then are connected to the labels that the given social group are associated with, rather than with labels "non-typical" to the group (122). Then, people from the social group are usually denoted with certain descriptions, for instance descriptions of being oppressed with a veiled Muslim woman (17, 18, 40).

3.2 Bias Detection in Text with NLP

Bosco et al. (2022) detected stereotypes from a collection of Italian Facebook messages, specifically the comments under a sympathetic post from an Italian singer over a span of a few days, about the 2015 shipwreck where 700 migrants died on their way to Italy. The dataset was annotated giving each comment a hate speech ranking from 0 to -3 among other categorizations, such as two binary columns: if stereotypes are present and if prejudice is present. Lastly, a variable to describe the negative standpoint of the comment was added: discredit. The authors conducted a lexical analysis where they discussed and compared the contents of the comments which appeared in each of their categories. NLP techniques used for this were tokenization followed by Term Frequency-Inverse Document Frequency (TF-IDF) scores. Their results revealed that where hate speech occurs, stereotypes often appear alongside it (22).

The dataset was then used to train a binary classification model to detect stereotypes from a hate speech dataset consisting of data from Twitter. Two baseline models were used, one using GilBERTo, an Italian BERT model, and an Support Vector Machine (SVM) model. They then trained the two models on their TF-IDF scores of lexical information. The trained models performed slightly better than the baseline models for detecting racial bias on the Twitter data. After investigating which words are more and less relevant for the classification and adjusting the data used to train the models, the scores improve more. Interestingly, the investigation revealed words such as "muslim" in the top 4 and "terror-

ist" in the top 20 most relevant words in the training to detect racial stereotypes.

In their study, Shimi et al. (2024) analyzed detection of gender stereotypes (126). The dataset used consisted of Twitter data in English and Spanish which included sexist and non-sexist language about women, and was annotated with respect to the authors' research goals. The aims of the study were two-fold. Firstly, binary classification of sexism in comments, and secondly to classify the intention of the commenter. The authors utilized 3 different classification models for the tasks: Multi-layer Perceptron (MLP), XGBoost and Convolution Neural Network (CNN). The MLP classifier model was using alongside language agnostic embeddings to accommodate the English and the Spanish texts. The XGBoost model used TF-IDF scores as input. For the CNN model, the data esd padded and the model esd trained with hyperparameter optimizers. The results revealed that the MLP model performed the best on the first task, while the XGBoost model outperformed the other models for task 2.

The 2019 study *Racial Bias in Hate Speech and Abusive Language Detection Datasets* by Davidson et al. analyzed racial bias in datasets for detecting hate speech, similar to Bosco et. al's (2024) (32). The authors used texts from tweets written in both standard English and African-American English.

In their approach for classification, they used regularized logistic regression models with Bag-of-Words (BoW) features as input. Tweets were categorized to be likely written by a white person or a black person, as this would affect the intentions behind certain word usages. Across the datasets used on the classifiers, all revealed that black-aligned tweets were more likely to be classified as negative, except in the case where the white-aligned tweet was likely racist in intention. Their conclusions suggest that people should be vary when using these datasets in their models, as the racial bias present in them can lead the models to methodically discriminate against African-Americans.

Terman (2017) conducted a research on the description of Muslim women in American news media coverage (140). The data she analyzed came from the *New York Times* and the *Washington Post*, with articles that contained the term "woman" from the 1980s and up to 2014. Terman studied two main subjects in this research, with a goal of finding confirmation bias of stories about Muslim women discussing oppression of women. Firstly, she located what countries most articles that concern women's rights were discussing. Secondly, it was established what most articles concerning Muslim women discuss. For the

first goal, statistical models and negative binomial analysis was used and for the second goal, the author used Structural Topic Modelling, which is built on top of Latent Dirichlet Analysis (LDA). Terman's study revealed that US media did indeed report more on Muslim women when they were associated with countries which do not respect women's rights, and non-Muslim women in countries that do respect them. Secondly, stories about women in Muslim countries more often were about gender-rights concerning topics, than when they were about non-Muslim women.

3.3 Microportrait Extraction in News Media

Fokkens et al (2018) created and applied their microportraits pipeline (described in section 5) to a collection of news articles and blogs related to politics in the period January to March 2017. The microportraits, or descriptions, were extracted from the words "Dutch" or "Muslims", the results of which were then compared to outputs of LDA. The most typical descriptions for both groups were analyzed and given overall sentiment scores.

The microportraits were extracted using syntactic patterns and entity coreference resolution. On a sentence level, nouns were extracted as labels in the sentence, and properties were extracted from modifier and attribute words. Roles were identified through syntactic dependencies with the basic rule that subjects are agents in a sentence, objects are patients and indirect objects are recipients. For example, in the sentence "she gave her friend a gift", the subject is "she", the object is "gift" and the indirect object is "her friend". This thesis takes heavy inspiration from the article and research written and conducted by Fokkens et al. in 2018.

The results of the study show that Dutch people in the news during this time were mostly regarded when they won something, for instance in sports. On the other hand, in articles about Muslims, the most frequent words were "assault", "president" and "American", referring to Trump's travel ban of people from certain countries. Articles discussing both "Dutch" and "Muslims" usually regarded "elections", "integration" and "democracy". Other typical descriptions of Muslims included "rape", "insult", "murder", "radical", "conservative", "extremist" and "moderate", while descriptions of Dutch included "famous", "free", "Dutch origin", "beautiful", "win", "score" and "break".

The collected microportraits were annotated by four students as negative or positive, and the results revealed that microportraits about Muslims were all negative. Meanwhile, about Dutch people, they were either neutral or positive. In order to validate the results,

around a thousand descriptions were selected randomly and manually checked, where most of them were correctly extracted from the text.

Kluska (2023) conducted a similar analysis by utilizing the framework described by Fokkens et al. (2018) in her thesis on exploring queer stereotypes in Polish online news media (70). A comparison is done between two Polish online news papers and how each of them describe queer people, where one is expected to be more conservative than the other. For her analysis, Kluska defined three sets of neutral terms. One set with words expected to be neutrally described by both news sites called control terms, words to refer to queer people called target terms, and words to refer to social groups usually contrasted to queer people in Polish news, called counter terms. Kluska created a pipeline for microportrait extraction from Polish text from scratch. The steps in the pipeline included morphological and syntactic tagging of the data, followed by common NLP pre-processing tasks such as tokenization and lemmatization. The author extracted the same information using *spaCy* as is described in the NLP pre-processing subsection in this thesis (see section 6.1.2). The extractions of properties, labels and roles formed lists of descriptions, in which a list of descriptions became one microportrait. Kluska compared their findings against results of LDA topic modelling and of word clouds, on a subset of articles. The subsets included 10 articles per news paper, which included at least one term from each group of terms, in total 60 articles. In addition to this, the author applied sentiment analysis on the extracted descriptions. The results revealed that while one news site did indeed display a stereotyped depiction of queer people, surprisingly the other news site showed some tendencies to also do so.

3.4   Analyses of Muslims in Norwegian News Media

Pold wrote in 2022 his thesis on the perceptions of Muslims and Islam based on the analysis of Norwegian texts from alternative and traditional mediums (113). The study used exclusively articles which in some way expressed the views of the newspaper itself. Articles from the following mainstream newspapers: Vårt Land (76), Aftenposten (AP) (9), Dagen (29) and Dagbladet (DB) (28) and the alternative newspapers Resett (62), Document.no (35) and Human Rights Service (HRS) (123) were used. In his analysis, Pold chose one article from each news paper published between 23rd and 28th September in 2020, a weeks span from when the Norwegian government at the time publicly released the "Action Plan to Combat Discrimination and Hatred Towards Muslims (2020-2023)" (107). The articles

collected were all reactions to this plan. They were individually analyzed in depth on sentence level and document level, in addition to analyzing the usage of the words "muslims" and "islam". The utilized Discourse-Historical Approach (DHA) took into consideration linguistics as well as social structures and understandings of language. Pold concluded from his findings that only in alternative media were Muslims and Islam written about as one homogeneous entity. Furthermore, alternative media wrote of Islam and Muslims are creating problems, cannot co-exist with Norwegian society, and that they have a perceived elevated status in Norwegian society. The author pointed out that due to the nature of this analysis, one should not make general conclusion from it.

Jarmer (2020) wrote his thesis on the depiction of Islam in the established newspaper VG and the alternative newspaper Resett (61, 62). Using quantitative and qualitative content analysis on 153 article from 2017 to 2019, Jarmer researched how criticism and stereotypes are perpetuated in media, but also how these stereotypes and criticisms are challenged. In his study, Jarmer concluded that VG actively sought to challenge stereotypical narratives of Islam, for instance by amplifying voices of Muslims in the newspaper, showing for instance Muslim men as victims of violence instead of the violent one, and Muslim women as ambitious and proactive instead of passive and oppressed. However, the analysis also revealed that the Muslims which were represented in VG were a few given people, whom generally adhere to the same beliefs and politics, rather than showing a diversity of beliefs or for instance conservative Muslims' point of views. This may contribute in creating an image of Muslims where they are mostly alike, which in this context were "moderate" and more befitting the Norwegian society, or at least do not challenge the status-quo's.

A comparison of articles from the established mainstream newspaper AP and the large alternative newspaper HRS was conducted in Fossheim's thesis from 2018 (45). Articles from the period 2001 up until 2017 were analyzed. Fossheim conducted a comparative qualitative and quantitative analysis. From his findings, he concluded that AP's coverage and representations of Islam are to a large extent diverse and nuanced as the newspaper actively avoided equating Islam and Islamism. Instead it largely focused on problems related to integration, immigration, or those which Muslims in Norway face, without focusing on Islam. However, similar to Jarmer's findings from VG, Aftenposten showcased a certain type of Muslim, namely a "liberal" or "moderate" one, a Muslim that is integrated into the Norwegian society. This newspaper also allowed room for Islam criticism, for instance by publishing comments and debate content from people like Hege Storhaug, a well-known

critic of Islam and founder of HRS. Hege Storhaug is widely criticised to be Islamopho-
bic and far-right. Similar to VG, Aftenposten publishes polarizing content by publishing
a staunch Islam critic along with mostly liberal, "just-like-regular-Norwegians"-Muslims.
However, the newspaper notably published people's positive opinions of Islam more than
the negative. On the other hand, while HRS bring up valid criticisms of Islam, such as
the role and view of women, they pited the West against Islam and Muslim immigrants,
and viewed them as a threat to Western society. Contrary to Aftenposten, HRS actively
seeked to point out connections between Islam and terrorism, and increased immigration
in Norway to decreased safety in the society.

On behalf of the Fritt Ord foundation (46), the data analysis and news media monitoring
company Retriever conducted an analysis on Norwegian news media from 2016 to research
how Islam and Muslims are represented in it (116). Retriever completed a qualitative and
quantitative analysis of the news texts. The study revealed that almost 55,000 articles
collected from across newspapers, magazines and journals, published in 2016, mentioned
or discuss Muslims and Islam. Of these, almost half were about events outside of Norway,
and more than half of these discussed terrorism, ISIS and extremism. From the articles
which discussed local events inside of Norway, the focus was on Muslim individuals, or
was a debate post. 16% of the articles about Muslims/Islam inside the country focused
on criticism of Islam, and 16% were debate posts or essays which challenged Islam-criticism.

3.4.1 Choosing Microportraits Framework

The reviewed literature described a myriad of machine learning models and computational
text analysis methods one can utilize to detect racial bias in texts and other interesting
insights from said data. Picking microportraits as the main framework for this study
comprised of many reasons. The data available to me was in the form of HTML files, and
the time it would take to convert this raw data into labeled datasets and then meaningful
embeddings to train and test multiple machine learning models with, would be out of my
scope. Additionally, there were limitations met due to the lack of well-established and
documented resources for machine learning and language modeling for Norwegian. I go
more into depths of the challenges and limitations of the data and methodology of this
research in chapter 8.4. Lastly, Fokkens et al.'s (2018) study using microportraits was
deeply interesting and a novel approach and alternative to topic modeling. Microportraits
are a straightforward technique that can yield more fine-grained and meaningful results

compared to general topic modelling, revealing biases that are subtle and thus harder for a topic model to classify (40). Current research into the topic of ethnic and religious minorities in Norway uses non-computational techniques. The framework serves as an insightful first step to analyzing Norwegian news texts in the context of stereotypes and bias, from which one can build upon using more complex machine learning and artificial intelligence methods.

3.4.2 Expected Results

Based on previous research conducted on analysis of representations of Muslims in news, and on similar topics of racial bias detection, my expectations for the results of this thesis are of a similar vein. I predict the outcomes will show unbalance of negative and stereotype-congruent news reports on articles concerning Muslims as individuals or groups, or referring to Islam or Muslim-majority nationalities. I particularly expect the results to show a tendency of associating Islam with terror by nature of disproportionate reporting.

3.4.3 Contributions to the Research

This thesis contributes to the research by adapting microportraits to the Norwegian language and creating a pipeline to be used on Norwegian text, which has not been done before in my knowledge. Additionally, there is little research done using NLP methods, let alone microportraits, on mainstream Norwegian news texts which analyzes the descriptions of Muslims and Islam, as evidenced by the literature reviewed. The code used for this study can be found on GitHub [1].

3.4.4 Contributions Beyond Existing Studies

Outside of the research area, this thesis is a part of putting a spotlight on Islamophobia and the role that subtle knowledge sharing through news articles, and the language and the rhetoric used in them can contribute in creating a prejudiced landscape. I am completing this thesis in 2024, in a time where there is news of violent anti-Muslim riots in England (23, 36, 87, 150), and lack of coverage from major news sources (2, 20, 68, 83) on the deaths of tens of thousands of civilian Palestinians (26, 109, 137, 141). The importance of the this topic continues to be contemporary.

---

[1]Link to GitHub repository: `https://github.com/fatimaaaaaaa/The-Islamic-Elephant-In-The-Room/tree/main`

4

Data

This chapter details the data that is used for the experiment. Firstly, the data source is described briefly, followed by the contents of the data itself and introductions to the different newspaper used in the corpus. Lastly, the context in the dataset are briefly presented in the form of exploratory data analysis (EDA). The EDA presents statistics of the articles, as well as common words and phrases that appeared in the texts.

4.1   Norwegian Newspaper Corpus

The Norwegian Newspaper Corpus (Norsk aviskorpus) is a collection of news articles scraped from various online Norwegian news sites. As such, all the articles in the corpus are in raw HTML format. The corpus consists of news articles from 1998 to 2022, with all articles up to 2019 publicly available for download from The Norwegian National Library's pages (108). Data from 2020 and onwards is available in University of Bergen's corpus management and analysis tool Corpuscle, where it is possible to search through it but it is not available for download (106). For the purposes of this thesis, data from 2017, 2018 and 2019 were used.

4.2   Norwegian Newspapers

The corpus was compiled from a diverse range of 10 news sources, providing a comprehensive representation of the Norwegian media landscape. Table 9.3 shows an overview of the newspapers used in the corpus, in addition to their abbreviations in the data and whether it is a national or regional paper.

The sources include Adresseavisen (AA), a regional newspaper that serves the central Norway region, with a focus on local news and issues (7). Aftenposten (AP) is one of

Norway's largest and most widely read newspapers, covering national and international news (9). Bergens Tidende (BT) is a regional newspaper that focuses on news from Bergen and its surrounding areas, and is the largest newspaper published outside of Oslo (142).

Dagsavisen (DA) is a newspaper with a long history of affiliation with the Labour Party in Norway (30), while Dagbladet (DB) is another major Norwegian newspaper known for its investigative journalism and exposé articles (28). Dagens Næringsliv (DN) specializes in national and international economic and financial news (105).

The corpus also includes regional newspapers such as Fædrelandsvennen (FV), which serves the Kristiansand area in southern Norway and its surrounding regions (48), Nordlys (NL), which covers news from northern Norway (89), and Stavanger Aftenblad (SA), which serves the Stavanger area in southwestern Norway and its surrounding regions (8). Lastly, the corpurs includes Verdens Gang (VG), also one of Norway's most widely read newspapers (49).

4.3 Exploratory Data Analysis

The characteristics and metadata of the dataset are explored in this section. Basic dataframe querying built into the Polars framework was used to gather an overview over the following:

- Number of articles were published each year

- Number of articles published per newspaper per year

- Which month of each year had the most articles published

- Number of articles published about Muslims or Islam

In order to collect articles about Islam/Muslims, a regEx pattern search was conducted to check if an article included at least one instance of a string containing a term word. For instance, for the term "Islam", the regEx pattern was "\w*(islam)\w*". Thus, the search would match with any words which include "islam", such as *"islamofobi"* ("Islamophobia") or *"snikislamisering"* ("Stealth Islamization"), as well as the word itself. These articles are referred to as relevant articles in this section.

4.3.1 Articles Statistics

This section discusses how many articles were published each year, in which month that year most articles were published, and from which news paper the most articles were

collected from. This is both for all articles and for relevant articles. For a visual overview of the data discussed in this section and detailed statistics for all newspapers and all months, see figures 4.1 and 4.2, and 4.3 and figures 4.4, respectively.

In 2019, a total of 135814 articles were published, 2587 of which were relevant articles. In the dataset, AP published the most articles in 2019, both the most articles in total (27200) and the most relevant articles (567). In 2019, most articles were published in January for all articles (12749). The largest amount of relevant relevant articles were published in August (322).

From 2018, a total of 154433 articles were published, and 12321 of those were relevant articles. AP published most of the articles this year (35148), but DB published the the most amount of relevant articles (2449). Most articles in 2018 were published in January (15344), however most relevant articles were published in August (1343).

The most articles and also most relevant articles were collected from 2017. 179872 articles were published in total, where 18188 articles were relevant ones. Also in 2017, most articles were collected from AP (32422). Most relevant articles were collected from AA (4501). January was when most articles were published (18334), while March was when most relevant articles were (2005).

4.3.2   Common Words and Phrases

In order to get a quick overview of the contents and topics within the relevant articles, I created a word cloud. As seen in figure 4.5, it is made up of the 50 most frequently occurring words. Most of the words are neutral, such as *"muslimer"*, *"landet"*, *"norge"*, *"tyrkia"*, *"iran"* and *"islam"*. These translate to "Muslims", "the country", "Norway", "Turkey", "Iran" and "Islam". However, other words put the neutral words into context: *"angrep"*, *"drept"*, *"taliban"* and *"politiet"*, which translate to "attack", "killed", "Taliban" and "police". Other words of note mention governments (*"regjeringen"*) and authorities (*"tyrkiske_myndigheter"*), as well as presidents like "trump" and "erdogan".

Bigrams and trigrams are two and three words which frequently appear in sequence. The top 10 bigrams and trigrams found in the relevant articles are displayed in table 9.6. Most of these are names of people. Additionally, there are some organizations present such as the terrorist group "Boko Haram", the Norwegian anti-Islam group Stopp Islamiseringen av Norge (SIAN) mentioned by "Stopp islamiseringen", and the umbrella organization for Muslims in Norway Islamsk Råd Norge (IRN). Other word combinations of note are "Swore allegiance", commonly written when the topic is about "swearing allegiance" to an Islamic

terrorist group. Lastly, "Alnoor Islamic Center" likely refers to the mosque in Norway that was attacked in 2019.
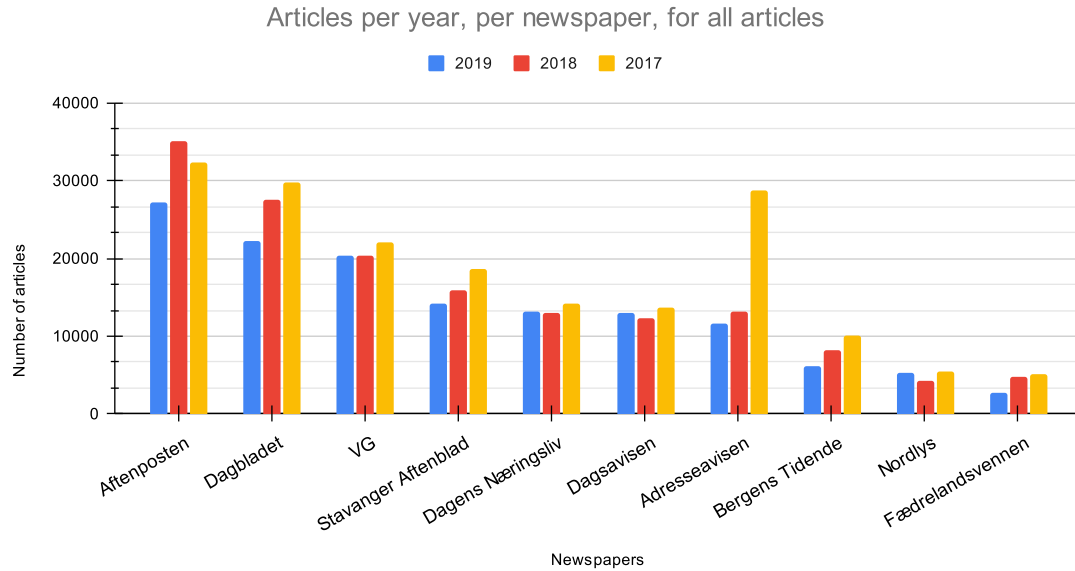
Figure 4.1: Articles per year, per newspaper, for all articles
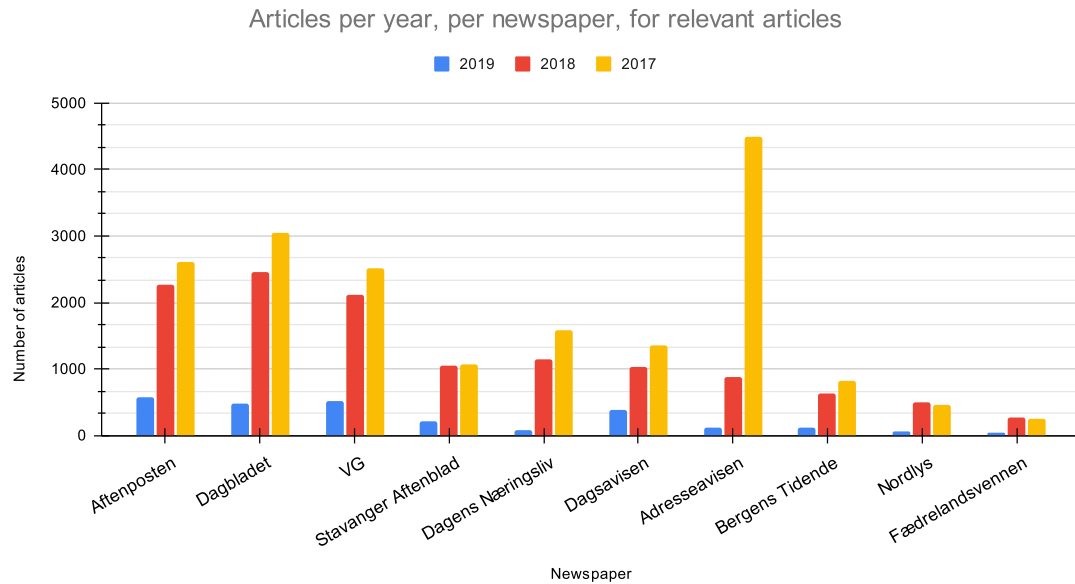
Figure 4.2: Articles per year, per newspaper, for relevant articles
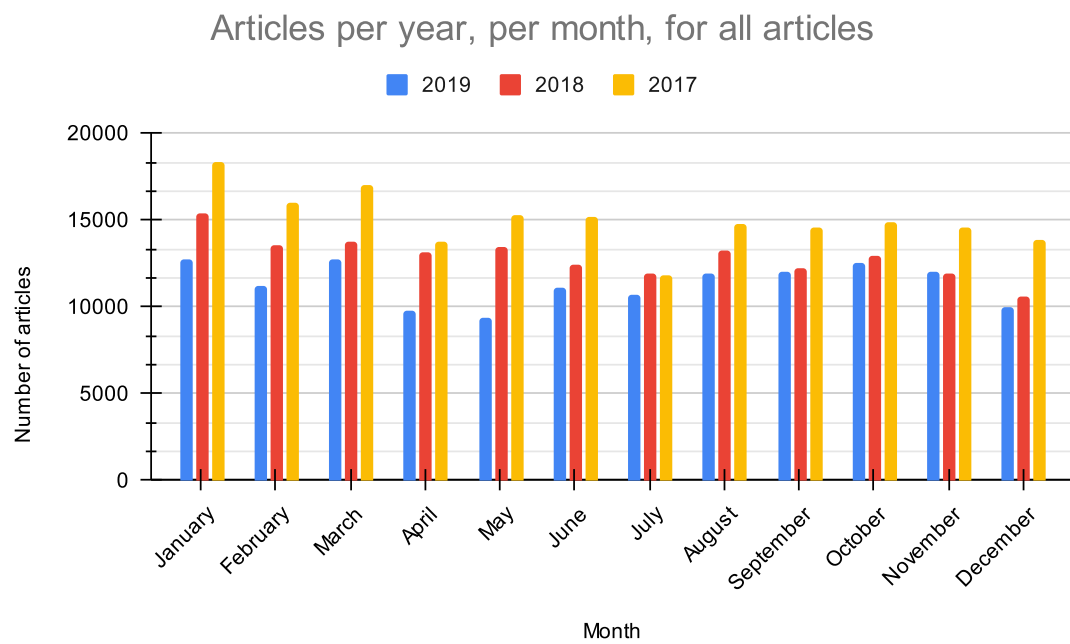
Figure 4.3: Articles per year, per month, for all articles
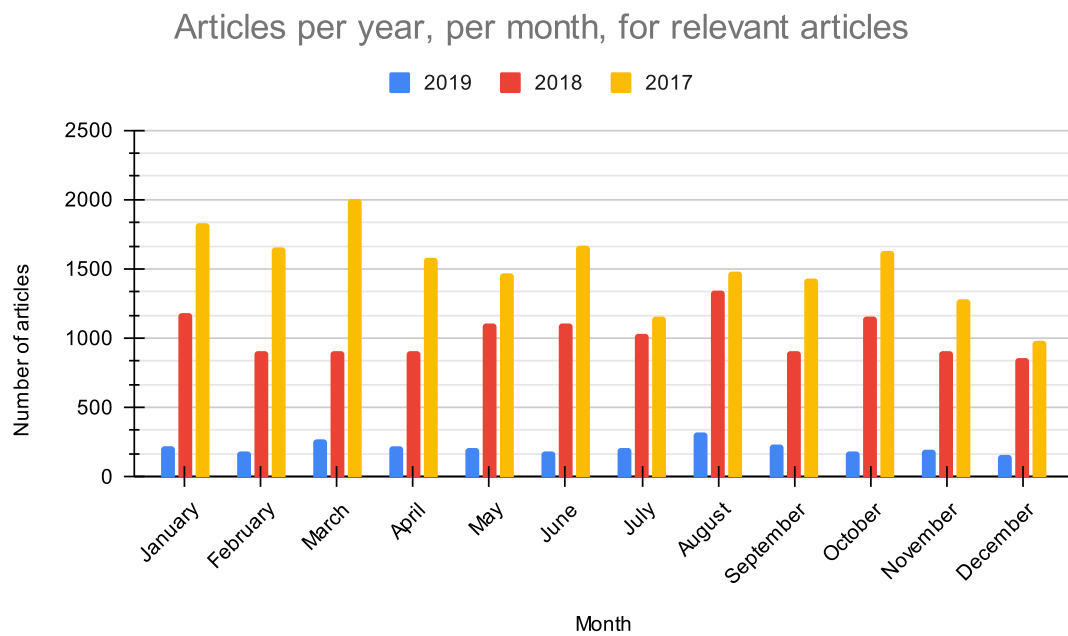
Figure 4.4: Articles per year, per month, for relevant articles

Figure 4.5: Most frequent terms which appeared in relevant articles

5

Theoretical background

This chapter goes in depth into the theoretical and technical background of which this paper relies on. This includes describing what microportraits are and the framework which they are built upon.

5.1   Microportraits

A microportrait is a collection of descriptions which represent the framing and stereotyping of an entity. Microportrait extraction is a method of gathering these descriptions from a given text. Descriptions include characteristics of an entity, such as the label or property assigned to it, or the role that the entity plays in an event. Take for example the given text "the fierce athlete won when the shocked opponent slacked". Descriptions include the label "athlete", the property "fierce" and the role *agent* in the event of "won". The label "opponent" has the property "shocked", along with the role *agent* in the event "slacked".

Microportraits as a concept rely on Beukeboom and Burgers (2017) SCSC framework. In their article *Linguistic Bias*, the authors described that it is a human tendency to put people they meet into social categories, and social categories usually have a set of traits and characteristics which we apply to all of those whom we put in a given category. This can have the negative effect of taking away people's individuality, the diversity within a group, and ignore an individual's context. At the same time, it emphasizes other people's differences compared to oneself and one's own social category. This process is also referred to as stereotyping, as was further described in the literature review in section 3.1.

In other words, microportraits rely on the assumption that stereotypes can be interpreted through analysis of content or frame in text.

5.2 Natural Language Processing

In order to conduct the research, a stack of different technologies were used. This section describes the various tools in depth. All the tools were utilized using the Python programming language.

5.2.1 Libraries

spaCy spaCy is a library which exists for performing NLP tasks in Python (133). The library includes methods for automatically performing computations such as tokenization, lemmatization and part-of-speech tagging. spaCy comes with pre-trained pipelines for some languages that are optimized to perform certain NLP tasks. The large Norwegian Bokmål pipeline *nb_core_news_lg* was used in this research (132). This pipeline is trained on data from the Norwegian Bokmål Universal Dependencies (UD) (34) and the NorNe list of Norwegian named entities (64, 65).

Pandas and Polars Pandas (2008) is well-known open-source Python library for data science tasks and for manipulating and working with dataframe objects. It includes functionality such as reading and writing CSV files and dataframe handling such as selection and filtering (111).

Polars was released in 2020 and offers the same functionalities, but with faster performance than Pandas (112). Polars lacks the same extensive documentation as Pandas due to its young age, but does regardless perform consistently. Pandas was only used where it was more convenient for this thesis, but Polars was largely used on its own.

BeautifulSoup4 Beautiful Soup 4 (bs4) is used for extracting and handling text from HTML and XML files, used in this project to handle raw HTML data from the Norwegian News Corpus (15). It offers functionality such as automatic parsing of HTML pages, and navigation of them using HTML tags and attributes.

NLTK Natural Language Toolkit (NLTK) is collection of libraries for performing various NLP tasks (144). Tasks conducted using NLTK tools include collocation tasks like finding words which commonly occur together in pairs or triads, and topic modelling for the purpose of investigating themes in the data.

5.2.2 Other tools

Dataframes The data used in this project was largely stored in the form of dataframes, which is a data structure that stores data in a way similar to spreadsheets (31). Is it a 2-dimensional data structure which is commonly used in data science.

Google Colab Colab is Google's browser-based online notebook for writing and running Python code (51). The code for this project was written and executed in Google Colab.

32

6

Methodology

This chapter describes the pipeline that was used to extract microportraits from raw data of news articles. A graphical representation of the pipeline can be seen in figure 6. Firstly, the raw data needed to be downloaded and cleaned so that it could be used in the pipelines. The cleaned data was used as input data, consisting of tables with cells that held the article text, its published date and its related news site. Afterwards, standard NLP procedures were conducted on the texts, such as lemmatization and tokenization. This processing allowed for gathering of meaningful insight from the data, as described in section 4.3. From the processed data, microportraits were extracted.

Finally, the pipeline was evaluated by selecting a number of articles randomly and putting them through the pipeline. The same articles were put through a topic modelling pipeline, and the results of the two were compared.

The procedure of gathering microportraits is detailed in section 6.2. The methodology used in this paper was adapted from the methodologies of Fokkens et al. (2018) and Kluska (2023) (40, 70). None of the code from either of the papers were reused or referenced. The pipeline used in this project was entirely created from scratch [1].

---

[1] Link to GitHub repository: `https://github.com/fatimaaaaaaa/` `The-Islamic-Elephant-In-The-Room/tree/main`
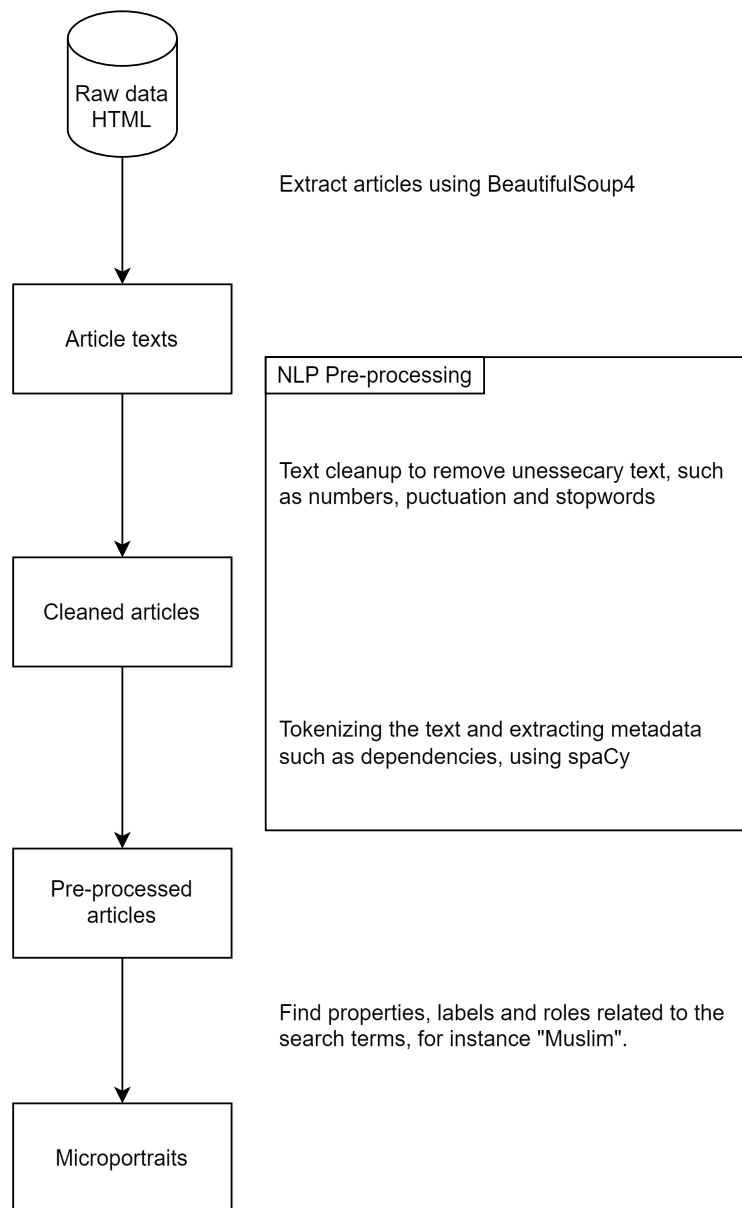
Figure 6.1: Flowchart of the microportrait extration pipeline

6.1 Pre-processing

The first step was to gather the data by downloading the files for each year, directly from the Norwegian National Library's website (108). Each file was a folder, which consisted of two folders per newspaper, one for each official written Norwegian: Nynorsk and Bokmål (72). Each raw article is an extraction of the whole HTML page.

For the purposes of this research, only articles in Bokmål are considered, as this is the most used standard.

6.1.1 Handling Raw Data

Using the bs4 Python library (15), the article body texts and dates were extracted from the HTML. Firstly, each article for each newspaper was looped through, and the newspaper name was collected. The body of the articles were found by using the *.document()* method of bs4, which automatically locates the main body of text on a page. Then, all the *div* elements of *type="text"* (*<div type="text">*) were collected and joined together as one piece of text. This avoided collecting for instance text under pictures and tables. Similarly, the date an article was last updated was found by locating *attribute* elements with *name="date"* (*<attribute name="date">*). The dates were restructured so that all extracted dates had the same *%d.%m.%Y* pattern. For example: 02.04.2019. The text, date and newspaper for each article were saved in Dataframe structures. Thus, one row in a dataframe consisted of data of one article.

There was a large amount of articles for each newspaper. In order to mitigate Google Colab's computation limits and simplify the process, a new file was saved each time the dataframe reached 2000 rows, and the data was split into several files.

6.1.2 NLP Pre-processing

The first step of the pipeline included removing rows which had any empty cells, and deleting duplicate rows. Following this, numbers and punctuation was removed. As it was noticed that long hyphens are commonly used in Norwegian text, often in titles and to indicate speech, these were removed and replaced with space. Other punctuation and special characters (such as currency signs and quotation marks) were also removed.

The next step was to remove stop words and short words. Stop words include words such as "and", "or", "not", "also", which do not carry the meaning of a sentence or without which the meaning can still be interpreted. The list of stopwords that were used in this pipeline were defined by Github user *Lotfi-AL*, who published a list combined from

multiple sources (80). Short words in this context were words shorter than 3 characters. The processed articles, along with their date and newspaper, were also saved in Dataframes. To keep the file sized small, these were saved in new files instead of in the same Dataframes as the original articles.

For the next steps, the NLP library *spaCy* was used (133). The texts were converted into *spaCy* documents and split into words, or tokenized. For each token, the following attributes was gathered (131):

- The given token

- The lemmatized token, or the root of the token word

- The part-of-speech (POS) tag of the token, which denotes what type of word it is

- The token's morphological features such as its tense

- The head word, which the token is directly syntactically related to

- The lemmatized head word

- The dependency of the token, meaning its syntactic relation to the head word

An example of this process is given in table 9.8, with the sentence "The smiling woman", where "smiling" is the given token. spaCy's own terminology for the attributes are also present in the table. The collection of this syntactic information was used for extracting descriptions later in the pipeline.

The extracted *spaCy* data was saved per article in a json format, as depicted in figure 6.2. The json format includes the attributes of each token in an article.

```
{
    "article": {
        "sentence": [
            {
                "text": " lett veier", // the first sentence in the article
                "data": [
                    {
                        "token": "lett",
                        "lemma": "lett",
                        "head": "veier",
                        "head_lemma": "vei",
                        "pos": "ADJ",
                        "dep": "amod",
                        "morph": "Definite=Ind|Degree=Pos|Gender=Neut|Number=Sing"
                    },
                    {
                        "token": "veier",
                        "lemma": "vei",
                        "head": "veier",
                        "head_lemma": "vei",
                        "pos": "NOUN",
                        "dep": "ROOT",
                        "morph": "Definite=Ind|Gender=Masc|Number=Plur"
                    }
                ]
            },
            {
                "text": "another sentence",
                "data": [
                    {
                        "token":"another",
                        ...
                    }
                ]
            }
        ]
    }
}
```

Figure 6.2: Example of how token metadata were saved in json objects

6.2 Microportraits

This section goes into depth of the process of extracting the actual microportraits from the *spaCy*-processed articles. Firstly, the terms used to extract the portraits are presented, then the extraction pipeline is described in detail.

6.2.1 Definition of Terms

In order to investigate the differences in descriptions of the minority group Muslims in news media as compared to the majority group Norwegians, five lists of terms were defined. Table 9.7 defines all the terms in Norwegian and English. *Target terms* include neutral words relating to Muslim groups and individuals, as well as neutral words related to Islam. The *target terms nationalities* list was defined to include nationalities of minority groups in Norway which are largely perceived to be majority Muslim. The selection of nationalities were chosen from the top ten immigrant groups in Norway originating from Muslim-majority countries in 2019, and also influenced by my personal perception from living in Oslo (104).

Minorities in Norway are sometimes described with "Norwegian" as a prefix combined with their second nationality, usually this categorizes the individual as Norwegian-born or with a Norwegian citizenship. For example *norskpakistansk* ("Norwegian-Pakistani"). Furthermore, occurrences of nationalities includeing the *norsk* prefix in news articles denoted that the aritcle most likely involved domestic cases. To cover instances of nationalities occurring this way, another list was created called *target terms norsk*. I collectively denote *Target terms*, *target terms nationalities* and *target terms norsk* as "target terms" non-italicized.

Counter terms were chosen according to neutral words used to refer to the larger population and specifically Norwegians, and words frequently employed in contrast to Muslims or Islam.

6.2.2 Extracting Microportraits

As described in section 5.1, a microportrait is a collection of descriptions that include labels, properties, agent roles, patient roles and recipient roles of a given word in a document. Descriptions were extracted for each term, on a sentence-level, utilizing a rule-based method. For each sentence, each token was inspected. The sentences processed included at least one of the terms from the lists described in the previous section. The lists will also be referred to as term categories. Descriptions were found using dependency-relations

between the tokens and given terms. Which relations to look for were decided using UD for Norwegian, which correspond to relations used in *spaCy* (34, 110). Thus, the data used for the extraction pipeline were the spaCy-processed articles.

Properties were found by inspecting the head word in a sentence. First, I checked if the head word was in one of the term categories. Then, I checked each token in the sentence where the head appeared. If one token worked as a modifier or an attribute to the head, then that token was extracted as a property of the head. In short, if head equals the given term, and a token exists with relation modifier or attribute to the head, then that token is a property of the head.

Labels were found with a similar but opposite approach. If in a sentence, there exists a token which was same as the given term, and it had a modifier or attribute relation to the head, then the head word was extracted as a label.

Lastly, roles were extracted in three categories: agents, patients and recipients. If the given term exists in a sentence, its relation to the head word determined its role. If the term has a role of agent, patient or recipient in relation to the head, then the head was extracted as a description.

Similar to the case for Polish and Dutch (40, 70), Norwegian had no ready-to-use semantic role labeler. Thus, the following rules were used to extract the roles:

- Agents = Subjects of active constructions, objects of passive constructions (UD: 'nsubj','obl')

- Patients = Objects of active constructions, subjects of passive constructions (UD: 'obj','nsubjpass')

- Recipients = Indirect objects (UD: 'iobj')

The top 10 most common microportraits for each type of description were found using frequency counts. The built-in method in the polars library *polars.Series.value_ counts()* automatically counts the occurrences of a word in a dataframe. This was done for each term, for each description, for each year.

6.3 Evaluation

To evaluate the microportrait extraction pipeline, a random selection of articles were picked as a sample to represent all the data. Ten articles were chosen from each category under target terms, meaning ten articles where a target term appeared, ten where a term for

nationalities appeared, and ten for *target terms norsk*. As the focus of my research is the depictions of Muslims, I chose to not include counter in the evaluation set.

The sampled articles were put through the extraction pipeline. Moreover, topic modelling was performed on these same articles to compare the results of the two. In the section below, the topic modelling process is described further in detail. Topic modeling was chosen as it is described as an alternative to microportraits extractions by Fokkens et al (2018) (40).

6.3.1 Topic Modelling

Topic modelling was conducted using the gensim library. Two models were created, a parallellized LDA model and a TF-IDF model using the *gensim.models.ldamodel.LdaModel* and *gensim.models.TfidfModel* methods of the library, respectively. The method used here was the one described by Nicha Ruchirawat (119, 120).

The data used for topic modelling was the already pre-processed data as described in section 6.1.2. Articles that did not include at least one word from any of the term categories were discarded. In order to optimize the topic modeling results, some extra steps were taken to modify the data further. Firstly, bigrams and trigrams (collectively referred to as n-grams) were identified. The n-grams were then merged to act as one word instead of two or three, and added back into the dataset.

A Pointwise Mutual Information (PMI) method was applied for this task. PMI calculates the likelihood of two or more words co-occurring within a given document, or in this case, within all articles found by searching for given terms. Words that occur less than 50 times were filtered out, before NLTK's bigram and trigram finding tools were applied to the data. In order to find relevant n-grams, a filter was applied to only find n-grams of adjective-noun combinations. Finally, as a last filter words that consisted of a single character were removed. A second modification applied to the data was to filter the remaining words further to include only nouns and verbs.

6 models were created, two for each of the 3 target term categories. Both took numerical representations of the tokens (embeddings), as input. The LDA model used BoW embeddings, while the TF-IDF model used TF-IDF embeddings. The BoW embeddings output numerical vectors representations of the data, which in this case included the tokens and the number of its occurrences in a document. The TF-IDF embeddings included a weight for each token, that denoted the likelihood of that token occurring in a document and in the whole corpus (60). To determine the optimal number of topics (k) for the models, the Coherence Measure tools from NLTK were used. After considering the results of the

ten different topic models, I settled on three topics represented by the 10 most frequently occurring words within it.

42

7

Results

The resulting microportraits extracted from the pipeline are described in this section, divided into target terms and counter terms. Following this, the results of the evaluation of the pipeline are described, which as conducted on a sample set of articles. This section is divided into presenting the microportrait extraction, then the topic modelling.

7.1 Microportraits

The results of the microportrait extraction are presented in this chapter. They are separated by term category, namely target terms and the control terms. The target terms are split into those regarding Islam and Muslims, the *target terms*, those regarding nationalities, the *target terms nationalities*, and those which combine "Norwegian" with the nationalities, the *target terms norsk*. Tables displaying the results are referenced in the respective sections. For each section, the most frequent microportraits are explained and the relevant news context around them is briefly discussed. The contexts will be discussed in detail in section 8.

7.1.1 Target terms

target terms The top 10 microportraits extracted from articles which contain one or more of the Islam focused target terms were all labels, as seen in table 9.9. The most common label was *"stat"* or "State" to the query term *"Islamsk"* or "Islamic", referring to the terror organization of the same name. The 2010s were affected by the Syrian civil war that broke out in 2011 and the aftermaths of which involved the overtaking of parts of Syria and Iraq by the Islamic State of Iraq and Syria (ISIS). The topic was extensively covered

in Norwegian news media and seemed to be the majority representation of any Muslims or Islam shown in media at the time.

Furthermore, Norway experienced a handful of its citizen leaving the country to join the war on the side of the group (54). A quick search for *"islamsk stat"* on the online news site Dagbladet reveals countless articles on the topic from the years 2017 to 2019. Articles cover some of the actions of the organization, but also individuals from the group usually with residence in a Western country such as Norway and the United Kingdom. One can assume that knowing this, most of the other top 10 microportraits for this category likely occur in articles covering the terror organization too, such as *"kvinne"* ("woman") and *"krig"* ("war").

"Brorskap" ("Brotherhood") however likely refers to coverage of the political situation in Egypt at the time. The Muslim Brotherhood is a political organization in Egypt. It is also labeled as a terror organization in some countries such as Russia, Saudi Arabia and even Egypt itself (148).

Most of the representation of Islam and Muslims thus cover news and topics which occur outside of the country. One microportrait that could refer to domestic news was *"fellesskap"* ("fellowship/community"). In 2017 The Islamic Community of Bosnia and Hercegovina in Norway left the other Norwegian Muslim organization IRN over the hiring of the semi-public figure Leyla Hasic, who bore the face-covering niqab (10, 57?).

target terms nationalities All of the microportraits that occurred with nationalities terms were also labels. A majority referred to international news, covering mostly the Turkish and to a lesser degree Iranian politics and warfare (see table 9.10). For example, descriptions such as *"myndighet"* ("authority"), *"styrke"* ("power") and *"offensiv"* ("the offensive") appearing with the query term *"tyrkisk"* ("Turkish" (adjective)) and *"luftrom"* ("air space") with *"iransk"* ("Iranian" (adjective)). Given the time frame of 2017 to 2019, the articles covered news of disputes and conflicts arising between the governments of Turkey and Iran and those who represented the large Kurdish minorities in the two countries. Additionally, Turkish military was involved in the Syrian civil war at the time.

target terms norsk As seen from the results in table 9.11, the collection of these microportraits were in significantly fewer numbers than the rest. Some topics that can be seen covered relates for instance to the case of two Norwegian Somali sisters held captive by the terror organization ISIS. For instance *"norsksomalier"* ("Norwegian-Somali" (noun)) with the role *"redd"* ("afraid"), and *"norsksomalisk"* ("Norwegian-Somali" (adjective))

with the label *"fremmedkriger"* ("foreign fighter") referring to people who travel from their countries to join a war in another country.

There was a lot of news coverage on the sisters, who first left in 2013, alongside the news coverage of the Syrian civil war and the terror organization as whole. The majority of these stories discuss the life of the two women as and the speculated life of women in the organization. It is of interest to note the way the Muslim women dominating the Norwegian news landscape at the time were the ones associated with ISIS. Articles were usually complimented with pictures or graphics of completely covered women in desert landscapes and often with children.

Another woman covered extensively in Norwegian news at the time was Bahareh Letnes of Iranian descent. The 10th most frequent microportrait "bahareh" with the label of *"norskiransk"* ("Norwegian-Iranian" (adjective)) refers to her. Most likely, the most frequent microportrait *"kvinne"* ("Woman") with the label of *"norskiransk"* also refers to Letnes. Letnes owns a company which deals with import and export of fish and seafood between Norway and Iran, and was revealed to the dating the then Minister of Fisheries of Norway. Controversy arose when the minister of Fisheries traveled to Iran with Letnes, and did not heed to security protocols related to traveling to Iran as a politician.

To summarize, in all three categories there were similarities in the topics covered. in that they were mostly related to war. Specifically, the conflicts in Syria and nearby, and related to the terror organization ISIS. Additionally, Muslim women tend to be emphasized in the news coverage in controversial topics such as related to terror or to the discussion of other people they might be related to (girlfriend, employee, victims of terror). While microportraits found by the *target terms* and the *target terms norsk* discussed both domestic and international themes, the ones found by the *target terms nationalities* exclusively focused on international news related to war in the Middle East. Some domestic topics in the two other term categories were also related to the same wars in the Middle East.

7.1.2 Counter Terms

The top 10 microportraits for counter terms reveal the association of terms such as *"nordmann"* ("Norwegian" (noun)) with roles such as *"vinne"* ("to win"), "liberal" with the label *"demokrati"* ("Democracy"). Associating "Norwegian" as a noun with the role "to win" was a common occurrence, as these often surround discussion of Norwegian athletes competing in a sports competition. The term "Western" with the label "Country" specifically denotes countries in North America, Australia and Western Europe, and is usually

used in contrasts other countries or where the mentioned countries behave similarly. The term "Liberal" tied to the label "(political) party" refers to the Liberal party of Norway.

There is a stark difference that is detected in the microportrait extracted from the target terms and the counter terms. The ones discovered from the target terms tend to be more specific and connected to specific news cases. Many of them are related to the same topics of ISIS and war in the middle east. If not explicitly, the other descriptions could implicitly be connected to the same context, such as the labels "Environment" and "Society".

7.2 Evaluation

To evaluate the microportrait extraction pipeline, a random selection of articles were picked as a sample to represent all the data. Ten articles were selected from each term category within the target terms.

The articles were put through the extraction pipeline, and used to perform topic modelling on. Firstly, the resulting microportraits are presented. Then, the results from the topic modelling by the LDA models and the TF-IDF models, respectively.

7.2.1 Microportraits

A total of 90 microportraits were extracted from the selected 30 articles by the pipeline. None of the which appeared more than three times. I will therefore not consider the quantity of microportraits found, but rather the content of them. All of the extractions for *target terms*, *target terms nationalities* and *target terms norsk* are listed in tables 9.13, 9.14 and 9.15, respectively. The descriptions are noted in Norwegian and English in the tables.

76 of the microportraits were associated with *target terms*, extracted by the terms *"nikab"*, *"islamsk"*, *"muslimsk"*, *"muslim"*, *"islam"*, *"hijab"*, *"burka"*, and one by *"moské"*. Descriptions of note include "Woman-oppressive" and "extreme" for *"islam"* and "forbid" for the terms *"burka"*, *"hijab"* and *"nikab"*. Moreover, *"islamsk"* was connected to *"stat"* as it was for all the articles, and *"muslim"* to "radicalize".

Another 32 microportraits were extracted by the *target terms nationalities*. These were related to the terms *"aghansk"*, *"irak*"*, *"iran*"*, *"pakistansk"*, *"somali*"* and *"tyrk*"*,

where * denotes varying suffixes such as *"-er"* or *"-sk"*. A single microportrait was extracted by *"marokkansk"*. Notable descriptions are "Military helicopter" to the term *"pakistansk"*, "Revolution army" to *"iranske"*, "Islamist group" to *"somaliske"*, "Terrorism legislation" to *"tyrkiske"*.

Finally, for the *target terms norsk* terms, only neutral descriptions such as "old", "Norwegian", "police" and "settled" were found in 6 microportraits. The microportraits were found for the terms *"norskiraker"*, *"norskpakistaner"* and *"norskiraner"*.

7.2.2 Topic Modelling

The results of the topic modelling are depicted in figures 9.1 to 9.6. Figures 9.1 and 9.2 show results extracted from the *target terms*, figures 9.3 and 9.4 for *target terms nationalities*, and figures 9.5 and 9.6 for *target terms norsk*.

The tables in the figures depict three topics each. The topics are comprised of the 10 words denoted as "value", and the tables show the frequency each word in the selected articles. The "frequency" represents percentage. For example, the word "hijab" in figure 9.1. It appeared less than 2% in the selected articles, in relation to topic 0. The word "nikab" appeared more that 2% in relation to topic 2.

LDA For the *target terms*, topic 0 includes word such as *"islamistgruppen"*, *"alshabaab"*, *"angrep"* and *"westgatekjøpesenteret"* the most. These translate to "The Islamist group", "Al-Shabaab", "attack" and "The Westgate Shopping Mall". All of these appeared more than 1% of the time related to this topic. Topic 1 includes words like "young", which appeared more than 1%, and "authorities", "attack", "investigation" along with a few of the given terms like "islam", "islamsk" and "muslimsk" which appeared less than 1%. Topic 2 involves "nikab" appearing more than 2.5% and "forbid" appearing more than 1.5% for the related topic. To a lesser degree the words "IS", "extreme", "the state" and "terror network". See figure 9.1.

For the *target terms nationalities*, topic 0 includes *"siktet"* ("charged") and *"danmark"* ("Denmark") appearing over 2% related to this topic, while *"Iranske"* ("Iranian") and *"myndigheter"* ("Authorities") appear less than 2%. Topic 1 regards *"gammel"* ("old"), *"bosatt"* ("settled") and *"iranske"* ("Iranian") appeared over 2% of the time. Except for *"gammel"*, no words appeared over 2% of the time, in topic 2. Words that appeared less were *"norskpakistaner"* ("Norwegian-Pakistani"), *"drept"* ("Killed") and *"kvinne"* ("Woman). See figure 9.3.

Lastly, topic 0 gathered from the *target terms norsk* include words such as "Norwegian-Pakistani", "the murder", "the police" and "charged" appearing less than 2%. Topic 1 includes the word "settled", appearing more than 8% of the time. Words "Drammen", "Fredrikstad" (two cities), "Iraqi" and "accused" all appear over 4% related to topic 1. For topic 2, the words "Iranian", "Norwegian-Iranian" and "Altostari" (a Norwegian-Iranian) appear 4% of the time.

The small number of articles resulted in the higher frequency of some words, as there were fewer articles to collect diverse words from. See figure 9.5.

TFIDF

For *target terms* as seen in figure 9.2, the topics 0 and 1 include most of the same words, such as "terror network", "IS", "Italian", "Krekar" and "staten". What differs between them is the frequency, where all except twi words for topic 0 appear in more than 30%, while for topic 1 they appear closer to 0% of the time. This is likely the TF-IDF model taking a lot of the same words and separating a few of them in another topic. For instance, there might be different words outside the 10 most frequently occurring for topic 1 that allude to a different topic. However, this is unlikely due to the low frequencies. Topic 2 involved words such as "Islamist group", "attack", "Al-Shabaab" and "Westgate shopping mall", appearing less than 1% of the time. From the words, it looks like topic 2 in the TF-IDF model include denotes the same theme as for topic 0 in the LDA model.

For the *target terms nationalities*, for topic 0 the word "settled" appeared over 4% of the time, with "Krekar" (a public figure), and "old" appearing over 2% of the time. Topic 1 includes "Denmark", "Iranian" appearing over 2%, and "charged" and "citizenship" appearing less than 2%. Topic 2 involves "Norwegian-Pakistani", "old" "the police" and "woman" occurring more than 2%. The results are detailed in figure 9.4.

Finally, for the *target terms norsk*, topic 0 denotes "settled" appearing 50% in related to this topic, with "Fredrikstad", ""accused", "Drammen" and "Norwegian-Iraqi" occurring more than 20%. Topic 1 involved words such as "settled", "travel", "trial" and "Krekar" (a public figure in Norway). All of these words appeared less than 10%. From looking at the words, it looks like topic 0 and 1 denote the same theme but differing in the frequencies of the words. Similar to the case for *target terms* described above. For topic 2, the words "Iranian", "Norwegian-Iranian" and "Altostari" appeared over 30%. See figure 9.6. The high frequency numbers for topic 1 might be due to few articles. The high frequencies for topic 0 and 2 are harder to explain, but it is likely due to a combination of the processing

the articles underwent before commencing topic modeling and the TF-IDF model itself grouping most occurrences of these words together to form a "topic".

7. RESULTS

8

Analysis and Discussion

For the discussion, I analyze the context around the findings. Results collected from the target terms are divided into three sections which encapsulate the main topics in the findings. These topics group the related microportraits together, since the analysis and discussion around them naturally also aligns. Terrorism, foreign fighters and women were chosen as themes to sufficiently represent the results for the target terms. Results from the counter terms are discussed as a whole, then compared to those from the target terms. Lastly, the results of microportraits from a selected sample of articles are discussed against the results of topic modelling on the same articles.

8.1 Target terms

8.1.1 Terrorism

A central topic discovered on the coverage of news related to Muslims and Islam was ISIS. News media outlet have a duty to report to the public the events of the world and enlighten them on relevant news topics. However, there is a fine line between reporting on terror and spreading the messages of the terrorists (130).

Many of the microportraits which fall under this category were those unveiled by the *target terms nationalities* and the *target terms* terms, as presented in section 7.1.1. During the years 2017 to 2019, international forces were well into Syria, including Iraq, Turkey and the USA, and a lot of ISIS-seized land was being retaken (11, 92, 94, 99).

Solum argues in her thesis on Norwegian news coverage of ISIS, that it was discussed in surplus and more than it needed to be (130). She makes the argument that terrorist actions have potentially little meaning if it does not get attention, particularly media attention (130). For instance, sharing videos and tweets made by ISIS is showcasing the groups

point of view, in a way allowing them a platform and an opportunity to control their own narrative in the public eye (13, 130). Particularly the influence of women in war, whom attract media attention several times more than men (13). More coverage and reportage of ISIS leads to more attention and even more coverage on them, greatly expanding their reaches (130). One can make an argument for the extensive coverage on the basis that this is the online news landscape today. It needs to be quick and many articles need to be churned out to stay up to date(25).

8.1.2 Foreign Fighters

According to the Norwegian Police Security Service (PST), around 30 foreign fighters with connections to Norway were in Syria in 2019 (114). A foreign fighter is defined as an individual that chooses to join a conflict in another country than their own, deeply sympathetic to a cause abroad they seek to aid in (54, 63). Hegghammer (2014) defines them as also ideologically motivated, different from a soldier, rebel or terrorists (54). Wars and conflicts have attracted fighters from all over the world, and Norwegians have joined foreign wars long before the conflict in Syria, such as in the Spanish Civil War and in the Finno-Russian Winter War (54). The Syrian conflict attracted foreign fighters on a larger scale than other conflicts for several reasons such as easier access to travel options, and engagement and recruitment through the Internet and social media (54, 63). As previously mentioned, ISIS had large reach in social media, but more in general, every aspect of the conflict could be reported on by individuals on and off-sight for anyone to access and engage with (52, 54).

One possible explanation of the media attention foreign fighters received was that it was far-right extremist Muslims that went, which is unexpected from a fairly Social Democratic and secular country. In contrast to for instance the Spanish Civil War, foreign fighters from Norway joined forces to fight *against* Fascism, while ISIS can be argued to perpetuate Fascistic rhetoric (54). Additionally, among Norwegians who travelled were well-known radical Islamists with social media followings, who upon travelling shared about it on Facebook (52). This was one way Norway and other Western countries were at all connected to the conflict in the Middle East. Another reason could be that foreign fighters were seen in what the terror group published, for instance the case of "Jihadi John". He was a British foreign fighter who attended many of their propaganda videos (50). Some Norwegian foreign fighters were also reported to be alongside prominent figures in ISIS and present in propaganda videos (63). Many foreign fighters died while in Syria, PST reports

about 30 Norwegians killed (114). Lastly, many foreign fighters returned home to their original countries. This would spark discussions on the level of threat the returners posed to the country, and how to treat them considering their situation (52, 54, 63).

8.1.3 Women

In 2013, two sisters from Norway left to Syria. In 2019, it was discovered that they were still alive in a refugee camp in Syria and that they wished to return home with their children (58, 86). This was extensively covered in the news, particularly in AP as they had journalists who managed to travel to the camp and interview the sisters (58). Questions of whether or not to bring them and their children home, or to only retrieve the children, were of debate (59). Several politicians stated that the now mothers would not be receiving any help, the argument being that they joined ISIS (4, 121). However, the sisters explained themselves by stating they left to Syria for "humanitarian aid", mostly cited in quotation marks, and not to join the terror group (58, 86). Articles referred to the sisters as ISIS-mothers, and one politician called them terror-women (4, 121). Moreover, the articles never fail to mention that the women were married to ISIS warriors who had since died in warfare. The main discussions in articles from 2019 after the sisters were found, were focused on how to bring children born to Norwegians in Syria home and the discourse surrounding the ethics of bringing children without their mothers (4, 59, 121). One article also alluded to there being public discourse on the "nature" of the children and whether they were destined to grow up to be terrorists too (59). However, instead of speculating on the outrageous concept of a child's nature, the article presented an interview with an expert of children of war. Åsne Seierstad released a book about the sisters in 2016, something which likely contributed to the media attention their case received (152). Stereotype-congruent behavior for Muslim women would be passive and submissive, to not decide for themselves, but rather have men do it. Despite the irony of leaving for ISIS, it is a rebellious and surprising choice to make, traveling to a war zone and joining a terrorist group. Certainly a decision that shows agency, as opposed to the "norm". Some of the women who did so from the West, have expressed feeling oppression in not being able to practice their religion without prejudice, and freedom in choosing to do what they believed would align with their beliefs. Narratives that describe these women as "terror-women" and "ISIS-mothers" can thus be to realign the women with stereotype-congruent behavior, as terrorists, as brides and as mothers (52, 82, 151).

8. ANALYSIS AND DISCUSSION

Religious and faith communities in Norway can receive government aid (6, 134). The Norwegian Muslim umbrella organization IRN received this funding until 2017, after a series of disagreements within the group. What ultimately triggered this decision was that in the same year, IRN hired Norwegian-Bosnian Leyla Hasic (38). This sparked widespread discussion in Norway as she wore niqab, a garment in which only your eyes are shown. The decision met objection within IRN too, with The Islamic Fellowship of Bosnia and Hercegovina leaving IRN (10), hence the microportrait "Islamic Fellowship". This situation was a particularly interesting one, as the discussion fully revolves around Hasic bearing niqab, and not about her as a person.

Critique to the organization's decision to hire her involved demands to revoke the aid IRN received, some of the reasoning being that they received a grant in order to hire more people full-time. As one article put it, for "building bridges", before stating that a niqab-bearing women was hired, clearly contrasting the two ((56, 139), figure 8.1). Moreover, it seemed like many had the impression that Hasic would be the "face" of the organization, instead of doing paper work (118, 138). An article questionably titled *Woman in Niqab is IRN's New Face* published by VG, also stated that Hasic was hired for administrative tasks (138). Others saw this hiring decision to be offensive particularly as face-covering garments were banned the year prior in schools and universities, even referring to it as "a type of declaration of war" (38, 118). It is hard to believe that Hasic's new position would attract this much attention and discussion without the widespread news coverage.

Part of the critique of niqab in Norway and globally regarded the inability to identify people with a face-covering (1, 14, 135). Post-pandemic, we know that it is indeed possible to do so, and many choose to regularly wear face masks (14). Even before the pandemic, there was and still is a precedence for face covering in Norway during the cold months (135). The narrative then perhaps ironically shifts to the idea of Muslims women needing "saving" from Islam and equating niqab to oppression. Simultaneously, to wanting to remove religious and particularly Islamic symbols from the public in Europe, fearing the "dangers" the niqab poses or represents (1, 14, 117, 135, 140). Lastly, an interesting observation was that very few articles cite any statements from Hasic herself about the situation, albeit referring to her appearing on TV for instance (56, 118, 139).

In 2018, Norway's then fisheries minister and deputy chairman in the right-wing conservative Progress Party Per Sandberg travelled to Iran with his then girlfriend Norwegian-Iranian Bahareh Letnes (44, 77). Sandberg and Letnes's relationship was already in the spotlight due to their large age gap, and Letnes being a immigrant with a conservative

politician (77, 124). Upon their travel to Iran, they were met with criticism for a multitude of reasons. One was due to Sandberg's position as a politician in parliament travelling to a controversial country. The second being that he brought his work phone along without reporting it, and which was revealed to be likely hacked in Iran (43, 44, 78). Furthermore, the couple received backlash for celebrating the Iranian national day, or as one article put it: "participating in celebration of Islamic revolution day" (125). The aftermath of this led to Sandberg stepping down from both his positions (71). Then prime minster Erna Solberg confirmed that PST conducted a search on Sandberg shortly after him and Letnes started seeing each other, due to a report of concern on Sandberg to the Prime Minster's Office (43).

During this time, a lot of Letnes's private life was revealed to the public such as when and why she applied for asylum in Norway. Moreover, speculations on what her relationship to the regime in Iran were in the spotlight. She expressed being distraught over the situation and even having received death threats (77). One can read from the articles surrounding the situation that Letnes received inflated media attention for choices that Sandberg made, to the point where Letnes was made out to be a possible threat to national security (44). Additionally, Letnes was simply mentioned in most articles about Sandberg during this time, hence the connection to the term "Norwegian-Iranian".

The findings from the result align with Muslim people's experiences and results from surveys conducted on Muslims in Norway, as extensively analyzed in chapter 2. The word choices in some of the articles were questionable and lead the narrative in a negative direction. For all the individuals discussed in this section, their non-Norwegian heritage was consistently mentioned. Most of all, there was extensive coverage of terrorism in the Middle East. This is particularly clear for the women I discussed, where there was consistent connections made to stereotypical behaviors.

The two sisters were described as "brides" and "mothers" and their relation to terrorism was emphasized. The discourse which followed Leyla Hasic's hiring largely failed to give Hasic a voice by not including interviews with her, at the same time as comparing her niqab to something that hinders connections and communication with other people. Niqab clearly signals the Muslim identity of a woman. By not including quotes from Hasic in articles about her, the idea of Islam oppressing women can be understood.

In articles where Bahareh Letnes is discussed, the fact that she was an asylum seeker from Iran were often stated alongside her opinions about her country. This is significant

because she was in a relationship with a politician, and the articles could allude to there being something to suspect about her - a potential threat to national security.

Islamsk Råd fikk statlig støtte til brobygging og dialog. Ansatte medarbeider med niqab.

Figure 8.1: From the news site Aftenposten, the title reads *Islamsk Råd received government aid for building bridges and dialogue. Hired a coworker with niqab.* (56)

8.2   Counter Terms

Shifting to look at the results of the microportraits acquired by the counter terms brings a significant change of tone. Topics of interest include Norwegian wins in sports such as in Ski World Championship in 2019, in chess, in football and other sports. Descriptions such as "Western Country" and "Western World" occur in articles often, usually regarding international politics. "Liberal democracy" also appears in regards to politics, particularly about the USA. The description "Norwegian (noun) million" refers to trends popular among Norwegians, and articles mentioning "Christian (adj) values" mostly regard the Christian Democratic Party of Norway.

Overall the microportraits found for this category of terms can be interpreted to be neutral or broadly describing something with ambiguity. The verb "To play" with the noun "Norwegian" does not invoke positive or negative associations, neither does "Young" or "A million". None of the microportraits were directly referring to specific events, but were rather vague and could apply to many news cases.

To summarize, based on the results, one can infer that there is indeed a large difference in the majority of depictions of Norwegians verse Muslims in the news. Most of the microportraits found by the target terms revolve around stereotype-congruent behaviors and expectations, if not all. It is obvious for descriptions surrounding terrorism and warfare. For instance, in the case of Leyla Hasic. The extensive news coverage from which it was emphasized that there was something wrong with her garment of choice, when working for an Islamic organization, created an atmosphere of danger and fear of the unknown. Baharah Letnes was frequently connected to the regime in Iran and speculated to be luring a politician currently in government to sympathize with and create relations to a controversial regime. The landscape at the time these news were discussed was one with frequent news of Islamic extremism, which I believe may have reinforced these fears. In comparison, microportraits found by the counter terms regarded mostly sports, and for the most part covered a variety of topics and cases.

8.3   Evaluation

The evaluation of this pipeline could have been developed and refined more than what it was. This is further discussed in section 8.4. Nevertheless, the results from the randomly selected 30 articles yielded some insights.

8.3.1 Microportraits

The themes from the extracted microportraits largely aligned with those found in the broader analysis conducted in the previous section.

ISIS and Islamic extremism are topics that can be alluded to by the microportraits extracted from the *target terms*, with descriptions such as "extreme" and "radicalize". Moreover, the association of *"hijab"*, *"nikab"* and *"burka"* with the word "forbid" likely revolves around discussions of banning the garments. This topic was mentioned briefly in the analysis of the discussions of Leyla Hasic, as described in section 8.1.3.

Warfare seems to be the main theme in the extractions from the *target terms nationalities*. Descriptions such as "military helicopter", and "revolution army" along with "terrorism legislation" for the term *"tyrkiske"* hint to the warfare revolving around ISIS. The context around this was discussed in the previous section. One topic discovered here which did not appear in the top 10 microportraits from the whole dataset, was that likely associated with *"somaliske"* ("Somalian") and "Islamist group". The radical Islamist group *Al-Shabaab* made the news due to several terror attacks in Somalia in the years 2017 to 2019 (27, 93, 95, 98).

The small size of microportraits extracted from the *target terms norsk* only returned neutral descriptions, likely due to the only being 6 extractions.

8.3.2 Topic Modelling

From the results of the topic modelling, the LDA model yielded better results. This was apparent in the cases for *target terms* and *target terms norsk*, where the TF-IDF model groups a lot of the same words for two of the three topics it detected. For this reason, only the results from the LDA model will be regarded in the comparison to the microportraits extraction pipeline.

For the *target terms*, as seen in figure 9.1, the topic models revealed topics such as the 2013 Al-Shabaab attack on the Westgate mall in Kenya which was straight-forward as words in the topic clearly alluded to this, denoted by topic 0 (66). The most frequent word in this topic is "hijab", which I could not understand why was related. From analyzing the selected articles, the two words were revealed to be related to a news case of a political party's campaign (91, 128).

Another topic revealed was that of the online terrorist network called Rawti Shax. In 2019, the Norewgian-Iraqi Mullah Krekar was in Italy for leading the terrorist network

(136). This network was allegedly connected to ISIS as well (96). This topic was denoted by words such as "Islamic", "state", "extreme", "IS", "rawti", "shax", "terror network" and "connected". However, the most frequent words in this topic (topic 2 in the figure) were "niqab" and "forbid", which are unrelated in reality, likely related to the same political campaign mentioned in topic 0. Topic 1 comprises of words such as "attack" and "investigation", however there is little to allude what is attacked or investigated.

The LDA results for the *target terms nationalities* likely involves the topics of a planned assassination of an Arabic Iranian separatist group, in Denmark in 2018 (12, 97). A Norwegian-Iranian man was accused to be involved in this plan (129). Words such as "assassination", "Denmark" and Iranian" allude to this. The other topic is alluded to by words such as "Norwegian-Pakistani", "Woman", "killed" and "accused". There was a case of a Norwegian-Pakistani man who was accused in 2018 for having ordered a contract murder of his wife (143, 145).

Topic 0 and 1 in figure 9.3 likely refer to the same news case, of the assassination, while topic 2 refers to that of the murder. Words such as "Fredrikstad" and "Drammen" which are cities in Norway, and "Iraqi" in topic 0 and 1 are more difficult to pin-point how they relate to the given news case. In fact, they seem to be related to Mullah Krekar, the Norwegian-Iraqi mentioned for *target terms*, where two others who were charged lived in the two cities mentioned (136). The word "Krekar" is in fact in the list for topic 1. The two topics seem to have been categorized together for the *target terms nationalities*. For topic 2, the word "Somali" and "Oromo's" (an ethnic group in Ethiopia) (3) seem unrelated to the news case. These word are likely related to the topic of "Al-Shabaab" from the *target terms*.

Lastly, for the *target terms norsk*, topic 1 likely denotes the same topic as for the *target terms*, about the terror network Rawti Shax and Mullah Krekar, as denoted by words like "Norwegian-Iraqi", "Italy" and "Krekar". This is depicted in figure 9.5. Topic 2 relates to the topic of the assassination in Denmark, as discussed for *target terms nationalities*, as denoted by words such as "Norwegian-Iranian", "Denmark", "assassination" and "ASMLA-members", which the Arabic Iranian separatist group in question (129). Topic 0 seem to combine the news case of the murder of the Pakistani women mentioned in *target terms nationalities*, and of another case: *"agjengen"* or *"A-gjengen"* is the name of a Norwegian gang, which were charged for dealing drugs and arms in 2017 (79).

8.3.3 Microportraits vs Topic Modelling

The topic modelling revealed several news cases which the microportraits extraction did not, such as about the Westgate Shopping Mall attack in Kenya, and of the planned assassination in Denmark. Although these results bring some insight into what kind of news are covered about Muslims and ethnic minorities in Norway, they are of little value with respects to the research question of this study. The LDA model depicts specific news cases, but often times groups different news cases together. Workarounds or adaptations to this are possible in terms of training the topic model to understand the language and the contexts of your dataset, for instance. A topic model thus requires optimizations in order to yield solid results. This is in contrast to the microportraits extraction pipeline. While the pipeline does have room to improve and could yield even better results, as discussed in chapter 8.4 and mentioned in section 9.1, the version I created from scratch served better for the case of this research. The microportraits allude to overall themes related to the specific words from the target terms. Moreover, the description types, the properties, labels and roles, give insights into the narratives in which these descriptions appeared. Based on these results, I would suggest the two frameworks to be used for different tasks.

8.4 Challenges

In the process of writing this thesis, I have met several challenges. In this section, I define the different challenges which I believe had a significant impact on the overall study. In short, these include limitations in domain knowledge, limitations with the Dataset, limitations of the microportraits framework and errors of the pipeline. Lastly, I consider restriction I met during the research such as with NLP tools.

8.4.1 Limitations in Domain Knowledge

Firstly, the topics of this study delve into Humanities and Linguistics, which are two areas of research I am not familiar with as a Computer Science student. The knowledge of Humanities which I possess may have led to lacking analyses, insights and conclusions that could have been described more in detail with respect to related studies and research in Humanities. Alas, as this is not a Humanities thesis, I did not go out of my way in order to research the Humanities aspects of the topic. While I was more familiar with the NLP part of Linguistics, I was however not familiar with most Linguistics terms. This resulted

in a substantial amount of time spent on familiarizing myself with relevant terms, which I assessed as vital to conduct this study.

8.4.2 Limitations with the Dataset

Another restriction to account for is that the publicly available Norwegian Newspaper Corpus only included news articles until 2019. Attempts to scrape news sites for data of newer articles directly from the news sites led to obstacles such as trackers that block attempts to scrape from their pages, one-page websites from which it was difficult to scrape from, and requiring users to login to access full articles. The corpus is available with articles up until 2020, however it has restricted access and only exists as an online database which can be queried and traversed, not downloaded, further discussed in section 4. A consequence of the dataset available only up until 2019 is that undoubtedly, most news cases covering Muslim and Muslim countries for instance will focus on ISIS. This was a time where the organization and their crimes were at the forefront of news media globally.

8.4.3 Limitations of Microportraits

While microportrait extraction yield results from which one can make different interpretations than those of topic modelling, microportraits only take into consideration the context which you infer from it and which the subset of articles surround. For instance, in Fokkens et al.'s (2018) study the article used for the analysis are picked from a period surrounding elections in The Netherlands. In the case of my results, without manual inspection and validation I cannot confirm that the microportraits which for example seemingly connect Islam and terror, do so in an objectively negative or subtly offensive manner, or if it is in a context which seeks to disperse the apparent connection.

Another limitation is that microportraits do not include context. It can be challenging to infer context without manual inspection, which sometimes is not possible. One example from the results is that of the top 10 microportraits for the counter terms (see table 9.12). Most of the microportraits are unspecific, and delving into the context of the news landscape at the time did little to change that in this case. Whereas the context surrounding the microportraits found for the target terms were easy to find. On the other side, this can be seen as a feature of microportraits framework. More specific results which give obvious context clues also might hint to then stronger ties to biased or stereotype-congruent language usage, while unspecific microportraits may imply little stereotyping.

Microportrait extraction is still a novel and little researched framework which has room to develop. It is a simple method that be applied along with others methods for better insights. Microportraits only look at words and extract the relationships between words. Any message, subtle meaning, and relations between entities in the text needs to be interpreted by the researcher.

The creators of the framework mention themselves that it is still a method that needs more and thorough evaluation (40), who's paper was reviewed in chapter3.

8.4.4   Limitation and Errors in the pipeline

The pipeline is lacking in the evaluations part. This part could be significantly improved given better capacity and technical knowledge. Additionally, the lack of linguistics knowledge adds to a lackluster evaluation of the results, as I would not be able to catch where they are not entirely as expected.

On a more technical note, the pre-processing steps skipped two characters that are used in Norwegian to denote quotes, namely the "« »", possibly because it is rare enough to not be included in automated special characters filtering tools. Also during the pre-processing, some words were erroneously lemmatized. For example, the word *"ramme"* ("affect") was lemmatized to "ram" which is not a word. However, I assume all words were lemmatized in a consistent manner, although possibly wrongly, and that this would not impact the results, but rather require an additional step of analysis. Moreover, the pipeline considered organization names that include a query term as a description, such as "Islamic *State*" and "*Islamsk* Råd Norge". These are not actually descriptions, but names.

8.4.5   Restrictions

There exists an active NLP milieu in Norway, with many resources and tools for NLP tasks open-source or in the works by research groups at universities. These are, however, often lacking descriptions on how to use them or with descriptions which I could not understand. Some tools were also many years old. Attempts to come in contact with the authors of the tools for help on how to get started or how to overcome bumps I met along the way, resulted in either no replies or at most two, in which there was little help to get. However, researchers are busy and I had to make due. I believe this study could have been stronger if I was able to use NLP tools specific to Norwegian (102), such as the Norwegian BERT model NorBert (74, 75). Many resources are also to be found in the corporate sphere,

albeit behind paywalls directed to be sold to mainly companies, such as by Retriever and NorskGPT (100, 115).

Lastly, several of the referenced papers and articles in this study are written entirely in Norwegian, as the topic takes place in a Norwegian context. Thus, much of the similar research conducted has been in Norwegian, such as surveys.

9

Conclusion

A large portion of Norway's immigrant population are from Muslim-majority countries. Muslims as a minority group immigrated to Norway first in the 1970s, mainly from countries such as Pakistan and Turkey. Several studies studies on Muslims in Norway, have revealed that stereotypical perceptions and negative attitudes are well present today. Norwegian Muslims have themselves expressed noticing these attitudes reflected in the news, and feeling the consequences. There exists few studies on this topic in the Norwegian context. I believe this thesis contributes to a niche research area and highlights the issues of subtly perpetuated racial biases against Muslims through news media in Norway.

This paper used a novel computational linguistics framework along with NLP techniques in order to analyze the depiction of Muslims and ethnic minorities from Muslim-majority countries, in the landscape of Norwegian news media. The framework was that of micro-portraits, which are collections of descriptions of a single entity within a given document. This was first introduced by Fokkens et al. in 2018 (40).

For the purpose of comparing the description of Muslims, ethnic minorities from majority-Muslim countries, and the descriptions of Norwegians as social groups, a microportrait extraction pipeline was created. These were extracted using terms related to the three groups, such as "Islamic", "Turkish", "Norwegian-Pakistani", and "Secular" and "Western".

The results were the ten most frequently occurring descriptions for each of the term groups. They gave insight into what type of news cases were most frequently covered about Islam and Muslims, and unveiled that the most prominent topics that were covered surrounded terrorism, foreign fighters and women.

The descriptions which fell under the terrorism category referred to articles discussing the terror group ISIS and warfare between them and countries surrounding Syria were involved in at the time. Articles about foreign fighters examined people from the West

travelling to join ISIS. Three cases of Muslim women in the news were discussed, as the descriptions highlighted them: Bahareh Letnes who was involved with a high-profile politicians as controversies arose, Leyla Hasic who was cast in the spotlight due to bearing niqab, and the case of two sisters that traveled Syria and were found alive many years later. This was in contrast to a majority of the microportraits found by the non-Muslim related terms, which were mostly vague and not about specific news cases.

In order to evaluate the pipeline, I selected a sample of articles. These were put through the pipeline, and conducted topic modelling on, separately. The comparison of the results from the two showed that microportraits revealed a more fine-grained insight into how Muslims are described in the news, than the topic model did. However, the two could be used to enhance the outcomes.

Conclusively, I want to refer back to the research question. *Is stereotyping of Muslims and ethnic minorities detectable in the news?* Given the resulting descriptions and by comparing those to the descriptions of Norwegians, stereotyping is indeed detectable. Most notably, when the news are about individuals. For instance, news about individual Muslim women often described them with stereotype-congruent roles such as "mothers and "brides", and emphasized their connection to something dangerous, such as to terrorism or to a foreign country. *Can this be objectively answered using NLP?* The microportrait framework misses the context of the words it finds, and the relations words have to other words in a document. However, this can contribute to make the found descriptions more objective. NLP serves to strip text from meaning and gives objective results to be interpreted by the researcher. Lastly, *What method is most suitable?* Between topic modelling and microportraits, the latter solved this task better as the former attempts to solve a classification problem instead. I believe the two can be used in conjunction to yield even more meaningful insights from news texts.

9.1 Future Work

This research has laid the ground work for further studies on racial bias in Norwegian news, particularly about ethnic and religious minorities. However, future work can and should be conducted to expand the conclusions made in this thesis. One thing that could likely make a difference is to use newer news articles. One could also use articles from a larger span of years and analyze the evolution of racial bias in the news, or even to articles

only surrounding specific news events and the coverage of that across a span of a short time. Moreover, creating a tagged dataset with meaningful lexical information would be useful for a variety of machine learning tasks. For instance, training the Norwegian language model based on BERT (NorBERT) to detect stereotypes of religious or ethnic minorities on Norwegian texts or Norwegian news articles. Additionally, at the time of writing this thesis in 2024, it is the age of wide-spread usage of Artificial Intelligence (AI) and particularly AI-powered language models. One can for instance train a model on a variety of articles, and attempt to generate new articles from it to detect racial bias in the news by analyzing the generated texts.

9. CONCLUSION

References

[1] **Muslim women wearing the niqab in Spain: Dialogues around discrimination, identity and freedom**. *International Journal of Intercultural Relations*, **75**:95–105, 2020. Available from: `https://www.sciencedirect.com/science/article/pii/S0147176719303335`. 54

[2] Mandour, Mohamed and Mandour, Doja and Mandour, Ignacio M. D. and Alsharif, Samir. **Attacks, arrests, threats, censorship: The high risks of reporting the Israel-Gaza war**. *Committee to Protect Journalists*, August 2024. Available from: `https://cpj.org/2024/08/attacks-arrests-threats-censorship-the-high-risks-of-reporting-the-israel-hamas-war` 20

[3] McKenna, Amy . **Oromo**. *Britannica*, July 2024. Available from: `https://www.britannica.com/topic/Oromo`. 60

[4] Helle Aarnes. **Åpner for å hente flere barn til Norge. Men IS-mødrene må bli igjen i Syria**. *Aftenposten*, April 2019. Available from: `https://www.aftenposten.no/norge/i/vQl9GX/aapner-for-aa-hente-flere-barn-til-norge-men-is-moedrene-maa-bli-igjen-i-syria`. 53

[5] Mona Abdel-Fadil and Liebmann Louise Lund. *Chapter 16 Gender, Diversity and Mediatized Conflicts of Religion: Lessons from Scandinavian Case Studies*, pages 281–298. De Gruyter, Berlin, Boston, 2018. Available from: `https://doi.org/10.1515/9783110502060-021`. 7, 10

[6] The Norwegian Tax Administration. **About government aid (state aid)**. Available from: `https://www.skatteetaten.no/en/business-and-organisation/reporting-and-industries/industries-special-regulations/state-aid/about-state-aid/`. 54

REFERENCES

[7] ADRESSEAVISEN. **Om Adresseavisen**, January 2024. Available from: `https://kundeservice.adressa.no/om-adresseavisen/`. 21

[8] STAVANGER AFTENBLAD. **Stavanger Aftenblad**, January 2024. Available from: `https://www.aftenbladet.no/`. 22

[9] AFTENPOSTEN. **Aftenposten**, January 2024. Available from: `https://www.aftenposten.no/`. 17, 22

[10] ØYSTEIN ALDRIDGE. **Norges største muslimske menighet trekker seg fra Islamsk Råd**. *Aftenposten*, March 2017. Available from: `https://www.aftenposten.no/norge/i/wooPM/norges-stoerste-muslimske-menighet-trekker-seg-fra-islamsk-raad`. 44, 54

[11] ANDREASSEN, TOR ARNE. **Det hvite hus: Militær uttrekning fra Syria har startet**. *Aftenposten*, December 2018. Available from: `https://www.aftenposten.no/verden/i/m6bavq/det-hvite-hus-militaer-uttrekning-fra-syria-har-startet`. 51

[12] ANDREASSEN, TOR ARNE. **Derfor jubler USA for arrestasjonen av en norskiraner i Danmark**. *Aftenposten*, October 20198. Available from: `https://www.aftenposten.no/verden/i/G1RAm6/derfor-jubler-usa-for-arrestasjonen-av-en-norskiraner-i-danmark`. 60

[13] ANSARI, SAMINA. **Kvinner er også i krig for og mot IS | Samina Ansari**. *Aftenposten*, Juli 2017. Available from: `https://www.aftenposten.no/meninger/debatt/i/3neWe/kvinner-er-ogsaa-i-krig-for-og-mot-is-samina-ansari`. 52

[14] SHAIMAA BALLOUT. **Revisiting the Veil Ban during the COVID-19 Epidemic: The Shift in the Social Context**. *Dirasat: Human and Social Sciences*, **50**(4):306–317, 2023. 54

[15] BEAUTIFULSOUP4. **Beautiful Soup Documentation — Beautiful Soup 4.12.0 documentation**, February 2024. Available from: `https://www.crummy.com/software/BeautifulSoup/bs4/doc/`. 30, 35

[16] LARS ERIK BERNTZEN. **Den eksistensielle trusselen, en sosiologisk studie av politisk motstand mot islam, muslimsk kultur og innvandring til Norge**. May 2011. Available from: `https://hdl.handle.net/1956/4905`. 7

[17] CAMIEL J. BEUKEBOOM AND CHRISTIAN BURGERS. **Linguistic Bias**, 07 2017. Available from: `https://oxfordre.com/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-439`. 13, 14

[18] CAMIEL J. BEUKEBOOM AND CHRISTIAN BURGERS. **How Stereotypes Become Shared Knowledge: An Integrative Review on the Role of Biased Language Use in Communication about Categorized Individuals.** *Review of Communication Research*, **7**:1–37, 2019. 14

[19] CAMIEL J. BEUKEBOOM, CATRIN FINKENAUER, AND DANIËL H. J. WIGBOLDUS. **The negation bias: When negations signal stereotypic expectancies.** *Journal of Personality and Social Psychology*, **99**(6):978–992, 2010. 10

[20] BOFFEY, DANIEL. **Israel accuses BBC of 'modern blood libel' over reporting of hospital strike**. *The Guardian*, October 2023. Available from: `https://www.theguardian.com/world/2023/oct/19/israel-accuses-bbc-of-modern-blood-libel-over-reporting-of-hospital-strike`. 20

[21] EIVIND BASTE BOLLINGBERG. **'Muslims' in Aftenposten in the 1970's: How muslims where reffered to in the context of muslim immigration to Norway and islamization in the muslim world.** June 2018. Available from: `https://hdl.handle.net/11250/2762971`. 3, 4

[22] CRISTINA BOSCO, VIVIANA PATTI, SIMONA FRENDA, ALESSANDRA TERESA CIGNARELLA, MARINELLA PACIELLO, AND FRANCESCA D'ERRICO. **Detecting racial stereotypes: An Italian social media corpus where psychology meets NLP**. *Information Processing & Management*, **60**(1):103118, 2023. 14

[23] BROWN, MARK. **Woman, 53, jailed over 'blow the mosque up' Facebook post after Southport riots**. *The Guardian*, August 2024. Available from: `https://www.theguardian.com/uk-news/article/2024/aug/14/woman-53-jailed-over-blow-the-mosque-up-facebook-post-after-southport-riots`. 20

[24] HEGE H. BYE, HENRIK HERREBRØDEN, GUNNHILD J. HJETLAND, GURO Ø. RØYSET, AND LINDA L. WESTBY. **Stereotypes of Norwegian social groups**. *Scandinavian Journal of Psychology*, **55**(5):469–476, 2014. Available from: `https://onlinelibrary.wiley.com/doi/abs/10.1111/sjop.12141`. 5

REFERENCES

[25] ABHIJNAN CHAKRABORTY, SAPTARSHI GHOSH, NILOY GANGULY, AND KRISHNA P GUMMADI. **Can trending news stories create coverage bias? on the impact of high content churn in online news media**. *Computation and Journalism Symposium*, 2015. 52

[26] CHALABI, MONA. **Why researchers fear the Gaza death toll could reach 186,000**. *The Guardian*, July 2024. Available from: `https://www.theguardian.com/world/article/2024/jul/12/gaza-death-toll-indirect-casualties`. 20

[27] DAGBLADET. **Minst 22 drept i dobbelt bilbombeangrep i Somalias hovedstad**. *Dagbladet*, Februar 2018. Available from: `https://www.dagbladet.no/nyheter/minst-22-drept-i-dobbelt-bilbombeangrep-i-somalias-hovedstad/69550774`. 59

[28] DAGBLADET. **Dagbladet**, January 2024. Available from: `https://www.dagbladet.no/`. 17, 22

[29] DAGEN. **Dagen**, January 2024. Available from: `https://www.dagen.no/`. 17

[30] DAGSAVISEN. **Om Dagsavisen**, January 2024. Available from: `https://www.dagsavisen.no/om-dagsavisen/`. 22

[31] DATABRICKS. **What are DataFrames?** Available from: `https://www.databricks.com/glossary/what-are-dataframes`. 31

[32] THOMAS DAVIDSON, DEBASMITA BHATTACHARYA, AND INGMAR WEBER. **Racial bias in hate speech and abusive language detection datasets**. *arXiv preprint arXiv:1905.12516*, 2019. 15

[33] MIGUEL DELGADO-RODRIGUEZ AND JAVIER LLORCA. **Bias**. *Journal of Epidemiology & Community Health*, **58**(8):635–641, 2004. 13

[34] UNIVERSAL DEPENDENCIES. **UD for Norwegian**, March 2024. Available from: `https://universaldependencies.org/no/index.html`. 30, 39

[35] DOCUMENT.NO. **Om Oss**, January 2024. Available from: `https://www.document.no/om-oss/`. 17

[36] DODD, VIKRAM. **More than 700 arrests made and 302 people charged over riots in England**. *The Guardian*, August 2024.

Available from: https://www.theguardian.com/uk-news/article/2024/aug/09/uk-riots-arrests-made-and-about-people-charged-say-police. 20

[37] Cora Alexa Døving. "Muslims Are...". Contextualising Survey Answers. *The Shifting Boundaries of Prejudice*, page 254–273, Feb 2020. 3

[38] Olav Elgvin. **Dialog, interessekamp og integrasjon: Islamsk Råd Norge og det norske samfunnet siden 1993**. *Kirke og kultur*, (4):380–400, 2021. 54

[39] European Commission against Racism and Intolerance. **ECRI REPORT ON NORWAY (sixth monitoring cycle)**. February 2021. Available from: https://rm.coe.int/6th-report-on-norway-/1680a17dd8. 9

[40] Antske Fokkens, Nel Ruigrok, Camiel Beukeboom, Gagestein Sarah, and Wouter van Atteveldt. **Studying Muslim Stereotyping through Microportrait Extraction**. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 2018. European Language Resources Association (ELRA). Available from: https://aclanthology.org/L18-1590. 13, 14, 20, 33, 39, 40, 63, 65

[41] The Centre for Equality and Diversity. **Welcome to KUN Centre for Equality and Diversity**, January 2024. Available from: https://www.kun.no/english.html. 8, 10

[42] The Norwegian Center for Holocaust and Minority Studies. **About the Norwegian Holocaust Center**, January 2024. Available from: https://www.hlsenteret.no/english/about/. 5

[43] Foss, Andreas Bakke and Moe, Ingeborg and Kagge, Gunnar. **Statsministeren fikk Sandberg-varsel i starten av juni**. *Aftenposten*, August 2018. Available from: https://www.aftenposten.no/norge/politikk/i/J1AXM6/Statsministeren-fikk-Sandberg-varsel-i-starten-av-juni. 55

[44] Foss, Andreas Bakke and Spence, Thomas . **Per Sandberg: Jeg fikk valget mellom å fortsette som statsråd og være kjæreste med Bahareh**. *Aftenposten*, August 2018. Available from: https://www.aftenposten.no/norge/politikk/i/l1jo2M/per-sandberg-jeg-fikk-valget-mellom-aa-fortsette-som-statsraad-og-vaere-kjaereste-n 54, 55

REFERENCES

[45] OLE FOSSHEIM. **Et ekkokammer og en liberalkonservativ riksavis En komparativ analyse av fremstillingen av islam og muslimer i Human Rights Service og Aftenposten, fra 2001-2017**. 2018. Available from: `http://hdl.handle.net/10852/63076`. 7, 8, 18

[46] THE FRITT ORD FOUNDATION. **What is the Fritt Ord Foundation?**, January 2024. Available from: `https://frittord.no/en/about/what-is-the-fritt-ord-foundation`. 8, 19

[47] FREMSKRITTSPARTIET. **Om Fremskrittspartiet**, January 2024. Available from: `https://www.frp.no/om-frp`. 7

[48] FÆDRELANDSVENNEN. **Fædrelandsvennen**, January 2024. Available from: `https://www.fvn.no/`. 22

[49] VERDENS GANG. **Om VG**, January 2024. Available from: `https://www.vg.no/informasjon/om-vg`. 22

[50] GIMSE, LARS MARTIN. **De fire britene ble kalt jihadistenes «The Beatles». Nå er tre fanget og én drept**. *Dagbladet*, February 2018. Available from: `https://www.dagbladet.no/nyheter/de-fire-britene-ble-kalt-jihadistenes-the-beatles-na-er-tre-fanget-og-n-drept/69438878`. 52

[51] GOOGLE. **Google Colab**. Available from: `https://research.google.com/colaboratory/faq.html`. 31

[52] LARS GULE. **Fremmedkrigere og terror**. *Kritisk juss*, **45**(1):18–47, 2019. Available from: `https://www.idunn.no/doi/abs/10.18261/issn.2387-4546-2019-01-03`. 52, 53

[53] ELLIOTT D. HAMMER AND JANET B. RUSCHER. **Conversing dyads explain the unexpected: Narrative and situational explanations for unexpected outcomes**. *British Journal of Social Psychology*, **36**(3):347–359, Sep 1997. 10

[54] THOMAS HEGGHAMMER. **Norske fremmedkrigere i går og i dag**. *Norsk statsvitenskapelig tidsskrift*, **30**(4):277–288, 2014. 44, 52, 53

[55] CHRISTHARDS HOFFMANN AND VIBEKE MOE. **ATTITUDES TO-WARDS JEWS AND MUSLIMS IN NORWAY 2022 - Population survey and minority survey**. December 2017. Available from: `https://www.hlsenteret.no/forskning/jodisk-historie-og-antisemittisme/holdningsundersokelsene/holdningsundersokelsen-2017/attitudes-towards-jews-and-muslims-in-norway-2017.pdf`. 5, 6, 7, 9

[56] PER ANNAR HOLM AND LENE LI DRAGLAND. **Islamsk Råd fikk statlig støtte til brobygging og dialog. Ansatte medarbeider med niqab.** *Aftenposten*, March 2017. Available from: `https://www.aftenposten.no/kultur/i/xoBdQ/islamsk-raad-fikk-statlig-stoette-til-brobygging-og-dialog-ansatte-medarbeider-med-`. vii, 54, 57

[57] DET ISLAMSKE FELLESSJAP BOSNIA OG HERZEGOVINA I NORGE. **Home**. Available from: `https://izbih.no/`. 44

[58] AFSHIN ISMAELI, TOR ARNE ANDREASSEN, OLGA STOKKE, AND KJETIL HANSSEN. **Aftenposten fant søstrene fra Bærum i live i Syria. Passet på fem foreldreløse norske søsken**. *Aftenposten*, June 2019. Available from: `https://www.aftenposten.no/verden/i/pL9WWo/aftenposten-fant-soestrene-fra-baerum-i-live-i-syria-passet-paa-fem-foreldreloese-n`. 53

[59] ISMAELI, AFSHIN AND ANDREASSEN, TOR ARNE AND SKOGSTRØM, LENE. **Råd fra ekspert: – De fem søsknene må ikke skilles. Deres viktigste støtte er hverandre.** *Aftenposten*, June 2019. Available from: `https://www.aftenposten.no/norge/i/LA995x/raad-fra-ekspert-de-fem-soesknene-maa-ikke-skilles-deres-viktigste-stoette-er-hvera`. 53

[60] JAIN, ABHISHEK. **TF-IDF in NLP (Term Frequency Inverse Document Frequency)**. *Medium*, February 2024. Available from: `https://medium.com/@abhishekjainindore24/tf-idf-in-nlp-term-frequency-inverse-document-frequency-e05b65932f1d`. 40

REFERENCES

[61] SEBASTIAN TJELLE JARMER. **Mediatisering, mangfold og megling. Fremstillinger av islam i etablerte og alternative medier. En studie av VG og Resett**. 2022. Available from: `https://hdl.handle.net/1956/22607`. 6, 7, 8, 9, 18

[62] VIBEKE JOHANNESSEN. **Resett as an alternative news outlet - A content analysis of Resett and its reporting on immigration and Islam**. 2020. Available from: `https://oda.oslomet.no/oda-xmlui/bitstream/handle/11250/2771668/Johannessen_MJ5900_2020.pdf`. 7, 17, 18

[63] JØRGEN SAND JOHANSEN. *Terrorismens sosiale drama: En analyse av sentrale narrativ, manuskript og fremføringer av den islamistiske terroren*. Master's thesis, University of Bergen, 2016. Available from: `https://hdl.handle.net/1956/12807`. 52, 53

[64] JØRGENSEN, FREDRIK AND AASMOE, TOBIAS AND VELLDAL, ERIK. **NorNE: Norwegian Named Entities**, April 2024. Available from: `https://github.com/ltgoslo/norne`. 30

[65] JØRGENSEN, FREDRIK AND AASMOE, TOBIAS AND VELLDAL, ERIK AND HUSEVÅG, ANNE-STINE RUUD AND ØVRELID, LILJA . **NorNE: Annotating Named Entities for Norwegian**, 2020. Available from: `https://doi.org/10.48550/arXiv.1911.12146`. 30

[66] KAALSTAD, JØRN E. AND LIEN, MARTHE S. AND TAHSEEN, RANO AND ROGNSTRAND, ANDREA. **Minister: Angrepet bygningskompleks i Nairobi er sikret**. *Verdens Gang*, January 2019. Available from: `https://www.vg.no/nyheter/utenriks/i/9maX8p/minister-angrepet-bygningskompleks-i-nairobi-er-sikret`. 59

[67] YURI KASAHARA, TONE LIODDEN, ANIS YAZIDI, STIAN LID, RAMESH UPRETI, ABOLPOUR ASIEH MOFRAD, AND BIRGITTA ANNE NILSEN. **Kartlegging av omfanget av hatefulle og diskriminerende ytringer mot muslimer**. April 2022. Available from: `https://hdl.handle.net/11250/2992278`. 7

[68] KHOURI, RAMI G. **Watching the watchdogs: Fear in newsrooms silences pro-Palestine voices**. *Al Jazeera*, November 2023. Available from: `https://www.aljazeera.com/opinions/2023/11/1/watching-the-watchdogs-fear-in-newsrooms-silences-pro-palestine-voices`. 20

[69] ANNE GRETHE KJELLING. **Å forhandle om identitet – unge muslimer om mediebildet av islam**. 2009. Available from: `https://hdl.handle.net/11250/3029325`. 7, 9, 10

[70] AGNIESZKA KLUSKA. **Adapting Microportrait Extraction for Queer Stereotype Identification in Polish Online News**. June 2023. 17, 33, 39

[71] KRISTIANSEN, ARNHILD AASS AND MOGEN, TRYM . **- En av Norges beste politikere må i dag gå av på grunn av en heksejakt**. *Dagbladet*, August 2018. Available from: `https://www.dagbladet.no/nyheter/en-av-norges-beste-politikere-ma-i-dag-ga-av-pa-grunn-av-en-heksejakt/70099905`. 55

[72] KUIPER, KATHLEEN . **Norwegian Language**. *Britannica*, August 2024. Available from: `https://www.britannica.com/topic/Norwegian-language#ref291223`. 35

[73] JONAS R. KUNST, HAJRA TAJAMAL, DAVID L. SAM, AND PÅL ULLEBERG. **Coping with Islamophobia: The effects of religious stigma on Muslim minorities' identity formation**. *International Journal of Intercultural Relations*, **36**(4):518–532, 2012. Available from: `https://www.sciencedirect.com/science/article/pii/S0147176711001362`. 7, 9

[74] ANDREY KUTUZOV AND JEREMY BARNES. **NorBERT**. Available from: `https://github.com/ltgoslo/NorBERT`. 63

[75] ANDREY KUTUZOV, JEREMY BARNES, ERIK VELLDAL, LILJA ØVRELID, AND STEPHAN OEPEN. **Large-Scale Contextualised Language Modelling for Norwegian**. In *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)*, pages 30–40, Reykjavik, Iceland (Online), May 2021. Linköping University Electronic Press, Sweden. Available from: `https://aclanthology.org/2021.nodalida-main.4`. 63

[76] VÅRT LAND. **Om Vårt Land**, January 2024. Available from: `https://www.vl.no/om/`. 17

[77] LOFSTAD, RALF. **Bahareh Letnes: - Mange usanne ting er blitt sagt og skrevet**. *Dagbladet*, August 2018. Available from: `https://www.dagbladet.no/nyheter/bahareh-letnes---mange-usanne-ting-er-blitt-sagt-og-skrevet/70072565`. 54, 55

REFERENCES

[78] LOFSTAD, RALF AND BRUSTAD, LINE . **- I beste fall svært dårlig dømmekraft**. *Dagbladet*, August 2018. Available from: `https://www.dagbladet.no/nyheter/i-beste-fall-svaert-darlig-dommekraft/70073007`. 55

[79] LORENTSEN, ANDRÉ AND NORSK TELEGRAMBYRÅ. **Stor narkoliga rullet opp – 36 pågrepet**. *Stavanger Aftenblad*, January 2017. Available from: `https://www.aftenbladet.no/innenriks/i/Kgzg5/stor-narkoliga-rullet-opp-36-paagrepet`. 60

[80] LOTFI-AL. **GitHub - Lotfi-AL/norwegian-stopwords: Norwegian stopword list**, November 2023. Available from: `https://github.com/Lotfi-AL/norwegian-stopwords`. 36

[81] C. NEIL. MACRAE AND GALEN. V. BODENHAUSEN. **Social cognition: thinking categorically about others**. *Annual Review of Psychology*, **51**:93–120, 2000. 13

[82] ALICE MARTINI. **Making women terrorists into "Jihadi brides": an analysis of media narratives on women joining ISIS**. *Critical Studies on Terrorism*, **11**(3):458–477, 2018. Available from: `https://doi.org/10.1080/17539153.2018.1448204`. 53

[83] MCGREAL, CHRIS. **CNN staff say network's pro-Israel slant amounts to 'journalistic malpractice'**. *The Guardian*, February 2024. Available from: `https://www.theguardian.com/media/2024/feb/04/cnn-staff-pro-israel-bias`. 20

[84] VIBEKE MOE. **DISKRIMINERINGSERFARINGER BLANT MUSLIMER I NORGE**. 2022. Available from: `https://www.imdi.no/contentassets/81b9c85aed6e42c994cf812e6f491fc3/rapport_diskrimineringserfaringblantmuslimer.pdf`. 9, 10

[85] VIBEKE MOE. **ATTITUDES TOWARDS JEWS AND MUSLIMS IN NORWAY 2022 - Population survey, minority survey and youth study**. April 2023. Available from: `https://www.hlsenteret.no/forskning/jodisk-historie-og-antisemittisme/holdningsundersokelsene/holdningsundersokelsen-2022/attitudes-towards-jews-and-muslims-in-norway-2022.pdf`. 5, 6, 7, 9

[86] TRYM MOGEN. **Norsksomaliske IS-søstre fra Bærum funnet i live i Syria**. *Dagbladet*, June 2019. Available from: `https://www.dagbladet.no/nyheter/` `norsksomaliske-is-sostre-fra-baerum-funnet-i-live-i-syria/71146279`. 53

[87] MOLONEY, CHARLIE . **Northern Irish first minister condemns mosque attack and other violence**. *The Guardian*, August 2024. Available from: `https://www.theguardian.com/politics/article/2024/aug/11/` `northern-irish-first-minister-condemns-mosque-attack-and-other-violence`. 20

[88] MARJAN NADIM AND AUDUN FLADMOE. **Silencing Women? Gender and Online Harassment**. *Social Science Computer Review*, **39**(2):245–258, 2021. Available from: `https://doi.org/10.1177/0894439319865518`. 9

[89] NORDLYS. **Nordlys**, January 2024. Available from: `https://www.nordlys.no/`. 22

[90] ISLAMSK RÅD NORGE. **Islamsk Råd Norge - Om Oss**, January 2024. Available from: `https://irn.no/om-oss/`. 10

[91] NORSK TELEGRAMBYRÅ. **Frp vil gi arbeidsgivere rett til å forby hijab på jobb**. *Stavanger Aftenblad*, March 2017. Available from: `https://www.aftenbladet.no/innenriks/i/LzqwP/` `frp-vil-gi-arbeidsgivere-rett-til-aa-forby-hijab-paa-jobb`. 59

[92] NORSK TELEGRAMBYRÅ. **Iraks statsminister erklærer IS' undergang**. *Aftenposten*, June 2017. Available from: `https://www.aftenposten.no/verden/i/` `99qel/iraks-statsminister-erklaerer-is-undergang`. 51

[93] NORSK TELEGRAMBYRÅ. **20 drept i sammenstøt med al-Shabaab**. *Dagbladet*, July 2018. Available from: `https://www.dagbladet.no/nyheter/` `20-drept-i-sammenstot-med-al-shabaab/69989409`. 59

[94] NORSK TELEGRAMBYRÅ. **Ordren om tilbaketrekning fra Syria er signert**. *Aftenposten*, December 2018. Available from: `https://www.aftenposten.no/verden/` `i/yvBoVr/ordren-om-tilbaketrekning-fra-syria-er-signert`. 51

[95] NORSK TELEGRAMBYRÅ. **Ni personer drept i bombeangrep i Somalia**. *Aftenposten*, May 2019. Available from: `https://www.aftenposten.no/verden/i/` `EWvMxG/ni-personer-drept-i-bombeangrep-i-somalia`. 59

REFERENCES

[96] NORSK TELEGRAMBYRÅ. **Terrorforsker stiller spørsmål ved Krekar-dom**. *Aftenposten*, July 2019. Available from: `https://www.aftenposten.no/norge/i/6jyk38/terrorforsker-stiller-spoersmaal-ved-krekar-dom`. 60

[97] NORSK TELEGRAMBYRÅ AND ANDREASSEN, TOR ARNE . **Norskiraner fengslet for attentatplaner i Danmark**. *Aftenposten*, October 2018. Available from: `https://www.aftenposten.no/verden/i/6nQ7V8/norskiraner-fengslet-for-attentatplaner-i-danmark`. 60

[98] NORSK TELEGRAMBYRÅ AND HAMMERSTRØM, INGER LISE. **Minst 300 drept i selvmordsangrep mot hotell i Somalia**. *Aftenposten*, October 2017. Available from: `https://www.aftenposten.no/verden/i/gVwVa/minst-300-drept-i-selvmordsangrep-mot-hotell-i-somalia`. 59

[99] NORSK TELEGRAMBYRÅ AND HUSØY, EIRIK. **Syrias hær: – IS' siste store by er gjenerobret**. *Aftenposten*, November 2017. Available from: `https://www.aftenposten.no/verden/i/vEpqV/syrias-haer-is-siste-store-by-er-gjenerobret`. 51

[100] NORSKGPT. **NorskGPT**. Available from: `https://www.norskgpt.com/`. 64

[101] ISLAMIC CULTURE CENTRE NORWAY. **Om ICC**, January 2024. Available from: `https://www.islamic.no/om-icc-2/`. 10

[102] NLP NORWAY. **Norwegian NLP Resources**. Available from: `https://github.com/web64/norwegian-nlp-resources`. 63

[103] "STATISTICS NORWAY". **4 prosent muslimer i Norge?**, January 2024. Available from: `https://www.ssb.no/befolkning/artikler-og-publikasjoner/4-prosent-muslimer-i-norge`. 4

[104] "STATISTICS NORWAY". **Innvandrere og norskfødte med innvandrerforeldre, totalt, etter kjønn og landbakgrunn 1970 - 2023**, January 2024. Available from: `https://www.ssb.no/statbank/table/05183`. 5, 38

[105] DAGENS NÆRINGSLIV. **Dagens Næringsliv**, January 2024. Available from: `https://www.dn.no/`. 22

[106] UNIVERSITY OF BERGEN. **Clarino - Corpuscle - Aviskorpus**, January 2024. Available from: `https://clarino.uib.no/korpuskel/corpus/concordance`. 21

[107] MINISTRY OF CULTURE. *Action plan to combat discrimination and hatred towards Muslims 2020-2023.* 2020. Available from: `https://www.regjeringen.no/contentassets/b2a6fd21c6a94bae83d5a3425593da30/action-plan-to-combat-discrimination-and-hatred-towards-muslims-final-version.pdf`. 17

[108] NATIONAL LIBRARY OF NORWAY. **Norwegian Newspaper Corpus**, January 2024. Available from: `https://www.nb.no/sprakbanken/en/resource-catalogue/oai-nb-no-sbr-4/`. 21, 35

[109] OSGOOD, BRIAN. **Israel's war on Gaza updates: Israeli tanks in Deir el-Balah amid attacks**. *Al Jazeera*, August 2024. Available from: `https://www.aljazeera.com/news/liveblog/2024/8/21/israels-war-on-gaza-live-israel-bombards-gaza-lebanon-us-says-deal-near`. 20

[110] LILJA ØVRELID AND PETTER HOHLE. **Universal Dependencies for Norwegian**. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Paris, France, may 2016. European Language Resources Association (ELRA). Available from: `https://aclanthology.org/L16-1250`. 39

[111] PANDAS. **pandas - Pythan Data Analysis Library**, April 2024. Available from: `https://pandas.pydata.org/about/`. 30

[112] POLARS. **Polars - About Us**, April 2024. Available from: `https://pola.rs/about-us/`. 30

[113] MARKUS KARLSTAD POLD. **Oppfatninger om muslimer og islam. En kritisk og relevansteoretisk analyse av meningsbærende tekster i alternative og tradisjonelle medier**. 2022. Available from: `https://hdl.handle.net/11250/3000822`. 8, 9, 10, 17

[114] POLITIETS SIKKERTSTJENESTE (PST). **Nasjonal trusselvurdering 2019**, 2019. Available from: `https://www.pst.no/alle-artikler/trusselvurderinger/trusselvurdering-2019//`. 52, 53

[115] RETRIEVER. **Retriever**. Available from: `https://www.retrievergroup.com/no/`. 64

REFERENCES

[116] RETRIEVER. **Islam og muslimer i norske medier i 2016**. 2017. Available from: `https://frittord.no/attachments/6821a1e981bafa24fe55aeb2b093b08ddad942f8/204-20220411151715188036.pdf`. 8, 19

[117] JANINE RICH. **'Saving'Muslim Women: Feminism, US Policy and the War on Terror**. *International Affairs Review*, **2**:1–6, 2014. 54

[118] NORSK RIKSKRINGKASTING AND NORSK TELEGRAMBYRÅ. **Kvinne i nikab er Islamsk Råds nye fjes**. *Norsk Rikskringkasting*, March 2017. Available from: `https://www.nrk.no/norge/kvinne-i-nikab-er-islamsk-rads-nye-fjes-1.13447885`. 54

[119] NICHA RUCHIRAWAT. **6 Tips for Interpretable Topic Models**. *Medium*, November 2020. Available from: `https://towardsdatascience.com/6-tips-to-optimize-an-nlp-topic-model-for-interpretability-20742f3047e2`. 40

[120] NICHA RUCHIRAWAT. **Optimizing LDA Topic Model for Interpretability**, February 2024. Available from: `https://nicharuc.github.io/topic_modeling/`. 40

[121] SOLVEIG RUUD. **Abid Raja mener Bærums-søstrene i Syria må komme seg hjem på egen hånd. KrF-topp uenig.** *Aftenposten*, June 2019. Available from: `https://www.aftenposten.no/norge/i/RR4oOa/abid-raja-mener-baerums-soestrene-i-syria-maa-komme-seg-hjem-paa-egen-haand-krf-topp-` 53

[122] DENISE SEKAQUAPTEWA, PENELOPE ESPINOZA, MISCHA THOMPSON, PATRICK VARGAS, AND WILLIAM VON HIPPEL. **Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination**. *Journal of Experimental Social Psychology*, **39**(1):75–82, 2003. Available from: `https://www.sciencedirect.com/science/article/pii/S0022103102005127`. 13, 14

[123] HUMAN RIGHTS SERVICE. **Om HRS**, January 2024. Available from: `https://www.rights.no/om-hrs/`. 17

[124] SETTEN, KAREN. **Italiensk avis: «Den fremmedfiendtlige ministeren fant kjærligheten med iransk flyktning»**. *Dagbladet*,

August 2018. Available from: `https://www.dagbladet.no/nyheter/`
`italiensk-avis-den-fremmedfiendtlige-ministeren-fant-kjaerligheten-med-iransk-flykt`
`70100867`. 55

[125] Setten, Karen and Lofstad, Ralf. **Sandberg deltok i feirin-
gen av islamsk revolusjonsdag som minister**. *Dagbladet*, Au-
gust 2018. Available from: `https://www.dagbladet.no/nyheter/`
`sandberg-deltok-i-feiringen-av-islamsk-revolusjonsdag-som-minister/`
`70086107`. 55

[126] Gersome Shimi, Jerin Mahibha, and Durairaj Thenmozhi. **Automatic
Classification of Gender Stereotypes in Social Media Post**. *Working Notes
of CLEF*, 2024. 15

[127] SIAN. **Stopp islamiseringen av Norge (SIAN)**, January 2024. Available from:
`https://www.sian.no/`. 7

[128] Skarvøy, Lars Joakim . **Frp-Sylvi vil forby barnehijab i
skolen: – Seksualiserer små jenter**. *Verdens Gang*, July 2017.
Available from: `https://www.vg.no/nyheter/innenriks/i/qqdpO/`
`frp-sylvi-vil-forby-barnehijab-i-skolen-seksualiserer-smaa-jenter`.
59

[129] Smilden, Jan-Erik. **Danmark har fått sin Khashoggi-sak – nesten**.
Dagbladet, October 2018. Available from: `https://www.dagbladet.no/kultur/`
`danmark-har-fatt-sin-khashoggi-sak--nesten/70400948`. 60

[130] Merete Kjølberg Solum. *Terror på dagsorden En analyse av norske nyhetsme-
diers dekning av IS-henrettelser av James Foley, Steven Sotloff, David Haines og
Alan Henning*. Master's thesis, University of Oslo, 2019. Available from: `http:`
`//hdl.handle.net/10852/70328`. 51, 52

[131] SpaCy. **Linguistic Features · spaCy Usage Documentation**, February 2024.
Available from: `https://spacy.io/usage/linguistic-features`. 36

[132] spaCy. **Norwegian Bokmål - spaCy Models Documentation**, April 2024.
Available from: `https://spacy.io/models/nb#nb_core_news_lg`. 30

[133] spaCy. **spaCy 101: Everything you need to know**, February 2024. Available
from: `https://spacy.io/usage/spacy-101`. 30, 36

REFERENCES

[134] COUNTY GOVERNOR STATSFORVALTEREN. **Religious and faith communities**. Available from: `https://www.statsforvalteren.no/en/portal/People-and-society/Religious-and-faith-communities/`. 54

[135] SISSEL SOMMER STENEBY. *I mine øyne: Muslimsk slør–et stoff med mening*. Master's thesis, Høgskolen i Sørøst-Norge, 2016. 54

[136] STOLT-NIELSEN, HARALD AND KALLELID, MAGNUS. **Dette er de norsk-tilknyttede medlemmene i mulla Krekars påståtte terrornettverk**. *Aftenposten*, July 2019. Available from: `https://www.aftenposten.no/norge/i/4qVBla/dette-er-de-norsk-tilknyttede-medlemmene-i-mulla-krekars-paastaatte-terrornettverk`. 60

[137] TANTESH, MALAK A, AND GRAHAM-HARRISON, EMMA. **Gaza rubble likely to conceal untold horrors to swell 40,000 death toll**. *The Guardian*, August 2024. Available from: `https://www.theguardian.com/world/article/2024/aug/15/gaza-death-toll-hits-40000-with-thousands-more-yet-to-be-counted`. 20

[138] NORSK TELEGRAMBYRÅ. **Nikabkledde Leyla Hasic er Islamsk Råds nye fjes**. *Verdens Gang*, March 2017. Available from: `https://www.vg.no/nyheter/innenriks/i/LzbaQ/nikabkledde-leyla-hasic-er-islamsk-raads-nye-fjes`. 54

[139] NORSK TELEGRAMBYRÅ. **Én kvinne og åtte menn inn i Islamsk Råds nye styre**. *Aftenposten*, March 2017. Available from: `https://www.aftenposten.no/norge/i/Lzz2q/en-kvinne-og-aatte-menn-inn-i-islamsk-raads-nye-styre`. 54

[140] ROCHELLE TERMAN. **Islamophobia and Media Portrayals of Muslim Women: A Computational Text Analysis of US News Coverage**. *International Studies Quarterly*, **61**(3):489–502, 11 2017. Available from: `https://doi.org/10.1093/isq/sqx051`. 6, 7, 9, 10, 15, 54

[141] THOMAS, MERLYN. **Israel Gaza What Gaza's death toll says about the war**. *BBC*, December 2023. Available from: `https://www.bbc.com/news/world-middle-east-67764664`. 20

[142] BERGENS TIDENDE. **Bergens Tidende**, January 2024. Available from: `https://www.bt.no/`. 22

[143] Tjernshaugen, Karen Torgersen, Hans O. and Norsk Tele-grambyrå and Stokke, Olga . **Norsk statsborger pågrepet for bestillingsdrap på ektefelle i Pakistan**. *Aftenposten*, July 2018. Available from: `https://www.aftenposten.no/norge/i/QlAV4x/norsk-statsborger-paagrepet-for-bestillingsdrap-paa-ektefelle-i-pakistan`. 60

[144] Natural Language Toolkit. **Documentation**. Available from: `https://www.nltk.org/`. 30

[145] Torgersen, Hans O. and Stolt-Nielsen, Harald and Riaz, Wasim K. and Virk, Arshad M. . **Pakistansk politi til Aftenposten: Mener at denne kvitteringen er betaling for drapet**. *Aftenposten*, July 2018. Available from: `https://www.aftenposten.no/norge/i/wEQWgd/pakistansk-politi-til-aftenposten-mener-at-denne-kvitteringen-er-betaling-for-drape`. 60

[146] Amna Veledar and Rasmus Burkal. **Hatefulle ytringer i offentlig debatt på nett**. 2018. Available from: `https://www.ldo.no/globalassets/_ldo_2019/arkiv/publikasjonsarkiv/fb-og-hatytring/ldo-fb_rapporten_2018.pdf`. 9, 11

[147] Kari Vogt, Dagfinn Rian, and Knut Axel Jacobsen. **Islam i Norge**. In *Verdensreligioner i Norge*, Oslo, Norway, May 2018. Universitetsforlaget. 3

[148] The Muslim Brotherhood's Official English web site | Ikhwanweb. **History of the Muslim Brotherhood in Egypt**, June 2007. Available from: `https://ikhwanweb.com/history-of-the-muslim-brotherh/`. 44

[149] Daniël H. J. Wigboldus, Gün R. Semin, and Russell Spears. **How do we communicate stereotypes? Linguistic bases and inferential consequences.** *Journal of Personality and Social Psychology*, **78**(1):5–18, 2000. 10

[150] Williams, Meleri. **Teen Muslims say they fear for safety over riots**. *BBC*, August 2024. Available from: `https://www.bbc.com/news/articles/cp3ddykeenlo`. 20

[151] Gunilla Zitawi-Koch, Sabina Rommelhoff Høy, Sabrina Anker Bue Jensen, Sofie Fromberg Støttrup, and Therese Bødker Andersen. **Islamistisk radikalisering med fokus på kvinder**. 2018. Avail-

REFERENCES

able from: `https://rucforsk.ruc.dk/ws/portalfiles/portal/62537585/BP2_` `_Islamistisk_radikalisering_med_fokus_pa_kvinder_152022_anslag.pdf`. 53

[152] ØVERGÅRD, MARTHE. **Seierstad ber norske myndigheter om å hente hjem de to søstrene fra Bærum**. *Aftenposten*, June 2019. Available from: `https://www.aftenposten.no/norge/i/BR6kv0/` `seierstad-ber-norske-myndigheter-om-aa-hente-hjem-de-to-soestrene-fra-baerum`. 53

Appendix

The appendix includes tables and large figures which were not included in the main text due to them just displaying the data which is already discussed in text.

9.2 Population in Norway

The table below presents the immigrant population in Norway in the years 2017, 2018 and 2019. This table is describes and discussed in chapter 2 *Context Analysis*, in section 2.1.1 *Statistics of Muslims in Norway*.

Table 9.1: Number of immigrants and children of immigrants in Norway in 2017-2019 by country of origin. Muslim majority countries are highlighted.

2017		2018		2019	
Polen	108255	Polen	110509	Polen	111985
Litauen	42491	Litauen	43930	Litauen	45415
Somalia	41463	**Somalia**	42406	**Somalia**	42802
Sverige	39266	Sverige	38870	Sverige	38770
Pakistan	36700	**Pakistan**	37412	**Pakistan**	38000
Irak	32304	**Irak**	33416	**Syria**	34112
Tyskland	27593	**Syria**	29889	**Irak**	33924
Eritrea	23618	Tyskland	27542	Eritrea	27855
Filippinene	22892	Eritrea	26218	Tyskland	27770
Vietnam	22658	Filippinene	23972	Filippinene	25078

9.3   HL-Center Survey Questions and Results

The HL-center conducted two surveys in Norway about people's perception of what Norwegians think of Muslims, one in 2017 and i 2022. The survey statements in Norwegian and in English are presented in the table below, and are stereotypical ideas of Muslims. Respondents could answer on a customized scale of 1 to 5 of agreement or sentiment. The "Answer" column refers to the scale value which most perpetuates the statement. This table, the survey and the context around it is discussed in chapter 2 *Context Analysis*, in section 2.2 *Stereotypes and Perceptions of Muslims in Norway*.

Table 9.2: Overview of statements and questions used by HL-senteret in their surveys

Statements	Answer	2017	2022
"What would you feel about having Muslim neighbors?" ("I hvilken grad ville du like/mislike at muslimer ble naboene dine?"	Dislike it a little/ a lot	37.8%	26.8%
"What would you feel about Muslims being brought into your friend group?" ("I hvilken grad ville du like/mislike at muslimer ble bragt inn i din vennekrets?"	Dislike it a little/ a lot	24.5%	21%
"Muslims are good Norwegian citizens" ("Muslimer er gode norske borgere")	Somewhat/ totally disagree	21.1%	13.1%
"I feel some disdain towards Muslims" ("Jeg føler en viss motvilje mot Muslims")	Somewhat/ totally agree	30.4%	26.4%
"Muslims are more violent than others" ("Muslimer er mer voldelige enn andre")	Somewhat/ totally agree	28.9%	25.6%
"Muslims have most of the blame for increasing anti-Muslim hate" ("Muslimer har selv mye av skylden for økende muslimhets")	Somewhat/ totally agree	47.4%	44.1%
"Muslims do not fit into a modern Western society" ("Muslimer passer ikke inn i et moderne vestlig samfunn")	Somewhat/ totally agree	36.4%	24.1%
"Muslims view themselves are morally superior to others" ("Muslimer ser på seg selv som moralsk overlegne andre")	Somewhat/ totally agree	45.1%	36.6%
"Muslims oppress women" ("Muslimer undetrykker kvinner")	Somewhat/ totally agree	69.1%	67.4%
"Muslims are a threat against Norwegian culture" ("Muslimer utgjør en trussel mot norsk kultur")	Somewhat / totally agree	42%	39.7%
"Muslims want to take over Europe" ("Muslimer ønsker å overta Europa")	Somewhat/ totally agree	30.3%	24.7%

9.4 Data

The tables below are discussed in chapter 4 *Data.*

Table 9.3 is presented in section 4.2 *Norwegian Newspapers.* It includes the news sites the corpus data was retrieved from, their abbreviations and their local reach. Tables 9.4 and 9.5 below are both discussed in section 4.3 *Exploratory Data Analysis.* They present the number of articles published in 2017 to 2019 for each news site in the corpus and which month of each year most articles were published. Finally, the table 9.6 is discussed in section 4.3.2 *Common Words and Phrases* and showcase the most common bigrams and trigrams from the relevant articles in the corpus.

Table 9.3: Overview of newspapers in the Norwegian Review Corpus

Name	Abbreviation	Reach
Adresseavisen	aa	regional
Aftenposten	ap	national
Bergen Tidende	bt	regional
Dagsavisen	da	regional
Dagbladet	db	national
Dagens Nærlingsliv	dn	national
Fædrelandsvennen	fv	regional
Nordlys	nl	regional
Stavanger Aftenblad	sa	regional
Verdens gang	vg	national

Table 9.4: Articles per year per news site

	All articles				Relevant articles		
Newspapers	**2019**	**2018**	**2017**		**2019**	**2018**	**2017**
Aftenposten	27200	35148	32422		567	2276	2610
Dagbladet	22178	27572	29730		487	2449	3041
VG	20302	20365	21988		522	2106	2513
Stavanger Aftenblad	14154	15819	18570		221	1044	1062
Dagens Næringsliv	13217	12923	14194		81	1147	1583
Dagsavisen	13065	12262	13717		383	1031	1357
Adresseavisen	11611	13147	28702		121	878	4501
Bergens Tidende	6099	8237	10034		109	634	814
Nordlys	5209	4255	5380		52	496	464
Fædrelandsvennen	2779	4694	5134		44	260	243
Total articles	135814	154433	179872		2587	12321	18188

Table 9.5: Articles per year per month

	All articles				Relevant articles		
Month	**2019**	**2018**	**2017**		**2019**	**2018**	**2017**
January	12749	15344	18334		219	1181	1831
February	11168	13564	15972		179	907	1660
March	12707	13786	17006		271	909	2005
April	9759	13152	13731		219	910	1586
May	9343	13444	15306		205	1107	1468
June	11085	12380	15135		184	1103	1664
July	10657	11906	11848		208	1036	1157
August	11917	13191	14721		322	1343	1481
September	11978	12173	14538		235	902	1431
October	12493	12915	14864		187	1161	1632
November	12015	11952	14546		196	903	1284
December	9943	10586	13829		162	859	986

Table 9.6: The 10 most frequently occurring bigrams and trigrams in relevant articles

Top	Bigrams	Trigrams
1	Malika Bayan	Predikant Fethullah Gülen ("Preacher Fethullah Gülen")
2	Sverget troskap ("Swore allegiance")	AlNoor Islamic Center
3	Human rights	Recep Tayyip Erdogan
4	Stopp islamiseringen ("Stop Islamization")	President Recep Tayyip,
5	Bahareh Letnes	Stopp islamiseringen Norge ("Stop Islamization Norway")
6	Fethullah Gülen	USA's president Donald
7	Erna Solberg	President Donald Trump
8	Boko Haram	Ytterliggående islamistgruppa IS ("Outlier Islamist group IS")
9	Alnoor Islamic	Råd Norge IRN
10	Recep Tayyip	Islamsk Råd Norge

9.5   Definition of terms

Table 9.7 was discussed in chapter 6 *Methodology*, in section 6.2.1 *Definition of Terms*, and describes the terms used for microportrait extraction.

Table 9.7: Overview of terms in Norwegian and English

Target terms	islam, islamsk, muslimsk, muslim, hijab, koran, moské, moske, quran, nikab, niqab, burka, burqa
	(Islam, Islamic, Muslim (noun), Muslim (adjective), hijab, quran, mosque, niqab, burqa)
Target terms nationalities	pakistaner, pakistansk, somalier, somalisk, marokkansk, marokkaner, iraner, iransk, irakisk, iraker, tyrker, tyrkisk, afghaner, afghansk
	(Pakistani (noun), Pakistani (adjective), Somali (noun), Somali (adjective), Moroccan (noun), Moroccan (adjecitve), Irani (noun), Irani (adjective), Iraqi (noun), Iraqi (adjective), Turkish (noun), Turkish (adjective), Afghani (noun), Afghani (adjective)
Target terms norsk	norskpakistaner, norskpakistansk, norsksomalier, norsksomalisk, norskmarokkansk, norskmarokkaner, norskiraner, norskiransk, norskirakisk, norskiraker, norsktyrker, norsktyrkisk, norskafghaner, norskafghansk
	Norwegian-Pakistani (noun and adjective), Norwegian-Somali (noun and adjective) *Norwegian-Moroccan (noun and adjective), Norwegian-Irani (noun and adjective)* *Norwegian-Iraqi (noun and adjective), Norwegian-Turkish (noun and adjective* *Norwegian-Afghani (noun and adjective)*
Counter terms	kristen, kristne, liberal, norsk, nordmann, ateist, sekulær, feminist, vestlig
	(Christian (adjective/noun), Christian (noun), liberal, Norwegian (adjective), Norwegian (noun), Atheist, secular, feminist (noun), feminist (adjective), Western)

9.6 *spaCy attributes*

Table 9.8 below considers an example sentence "The smiling woman", and details the *spaCy* attributes the study's pipeline would extract from it. This is explained in *spaCy* terms and from the example sentence. The table explains the method used in pre-processing, as explained in chapter 6 *Methodology*, section 6.1.2 *NLP Pre-processing*.

Table 9.8: Attributes collected per token

Example sentence	Attribute	spaCy	Value
"The smiling woman"	Token	token.text	"smiling"
	Lemma	token.lemma_	"smile"
	Part-of-speech	token.pos_	Verb
	Morphological syntax	token.morph	Present tense, participle verb
	Head	token.head	"woman"
	Head lemma	token.head.lemma_	"woman"
	Depencency	token.dep_	Adjectival modifier

9.7   Microportraits

This section displays the top 10 microportraits found by the pipeline for each term category. The retrieval of these is described in chapter 6 *Methodology*, section 6.2.2. The results are described in chapter 7 *Results*, section 7.1 and discussed in chapter 8.

Figure 9.9 lists the top microportraits found for *target terms*, while figure 9.10 list those for *target terms nationalities* and figure 9.11 lists those for *target terms norsk*. Figure 9.12 lists the top microportraits for the *counter terms*.

Table 9.9: Top 10 microportraits for target terms

count	query term	description	description type
89	islamsk/Islamic	stat/State	labels
43	muslimsk/Muslim (adj)	verden/World	labels
41	muslimsk/Muslim (adj)	land/Country	labels
39	muslimsk/Muslim (adj)	miljø/Environment	labels
37	muslimsk/Muslim (adj)	kvinne/Woman	labels
30	islamsk/Islamic	krig/War	labels
30	muslimsk/Muslim (adj)	jente/Girl	labels
24	islamsk/Islamic	fellesskap/Fellowship	labels
24	muslimsk/Muslim (adj)	brorskap/Brotherhood	labels
22	muslimsk/Muslim (adj)	samfunn/Society	labels

Table 9.10: Top 10 microportraits for target terms nationalities

count	query term	description	description type
70	tyrkisk/Turkish (adj)	myndighet/Authority	labels
65	iransk/Irani (adj)	myndighet/Authority	labels
62	tyrkisk/Turkish (adj)	styrke/Power	labels
62	tyrkisk/Turkish (adj)	regjering/Government	labels
41	tyrkisk/Turkish (adj)	politiker/Politician	labels
41	irakisk/Iraqi (adj)	soldat/Soldier	labels
40	tyrkisk/Turkish (adj)	minister/Minister	labels
38	iransk/Irani (adj)	luftrom/Air space	labels
36	tyrkisk/Turkish (adj)	offensiv/Offensive (as in combat)	labels
36	tyrkisk/Turkish (adj)	politi/Police	labels

Table 9.11: Top 10 microportraits for target terms norsk

count	query term	description	description type
8	norskiransk /Norwegian-Irani (adj)	kvinne/Woman	labels
7	norskpakistansk /Norwegian-Pakistani (adj)	miljæ/Environment	labels
6	norskpakistaner /Norwegian-Pakistani (noun)	nordmann/Norwegian (noun)	labels
5	norskpakistaner /Norwegian-Pakistani (adj)	gammel/Old or age	properties
5	norsksomalier /Norwegian-Somali (noun)	eie/Own	roles
5	norsksomalier /Norwegian-Somali (noun)	utvalg/Selection	properties
5	norsksomalier /Norwegian-Somali (noun)	redd/Save or afraid	roles
4	norsksomalisk /Norwegian-Somali (adj)	fremmedkrige/Foreign fighter	labels
4	norskiransk /Norwegian-Irani (adj)	miljø/Environment	labels
4	norskiransk /Norwegian-Irani (adj)	bahareh /Bahareh Letnes	labels

Table 9.12: Top 10 microportraits for counter terms

count	query term	description	description type
207	vestlig/Western	land/Country	labels
116	liberal/Liberal	demokrati/Democracy	labels
78	nordmann/Norwegian (noun)	vinne/ To win	roles
60	vestlig/Western	verden/World	labels
60	nordmann/Norwegian (noun)	million/ A million	properties
59	nordmann/Norwegian (noun)	spille/To play	roles
58	liberal/Liberal	parti/(political) party	labels
49	kristen/Christian (adj)	verdi/Value	labels
45	nordmann/Norwegian (noun)	ung/Young	properties
44	kristen/Christian (adj)	tro/Belief	labels

9.8   Evaluation

The tables and figures below include the results related to the evaluation of the pipeline. The method of evaluation is described in chapter 6 *Methodology*, section 6.3 *Evaluation*. The results are presented in chapter 7 *Results*, section 7.2, and discussed in chapter 8, section 8.3. The results were extracted from 30 randomly selected articles used for evaluation.

9.8.1   Microportraits

Tables 9.13, 9.14 and 9.15 present the results from microportrait extraction for the *target terms*, *target terms nationalities* and *target terms norsk*, respectively. These are discussed in chapter 7 *Results*, section 7.2.1.

Table 9.13: Microportraits extracted from selected articles by *target terms*

count	query term	description	description type
3	islamske	stat/State	label
1	burka	forby/Forbid	role
1	burka	redusere/Redusere	role
1	burka	ramme/Affect	role
1	hijab	hodeplagg/ Head garment	property
1	hijab	forby/Forbid	role
1	hijab	lære/Teach	role
1	hijab	vogn/Wagon	label
1	hijab	skole/School	label
1	islam	assosiasjon/Association	property
1	islam	miljø/Miljeu	property
1	islam	kvinneundertrykkend/ Women-oppressive	property
1	islam	ekstrem/Extreme	property
1	islam	prege/ Affect	role
1	islam	knytte/ Connect	role
1	islam	forvrenge/ Twist	role
1	islam	blåse/ Blow	role
1	islamsk	opprørt/Upset	property
1	islamsk	innflytelse/Influence	label
1	islamsk	høyreekstrem/ Far-right	label
1	islamsk	fremvekst/Emergence	label
1	islamske	miljø/Miljeu	label
1	moské	arbeide/Work	role
1	muslim	vikeligheten/Reality	property
1	muslim	radikalisere/Radicalize	property
1	muslim	spania/Spain	label
1	muslim	president	label

1	muslimsk	utseende/Appearance	label
1	muslimsk	jente/girl	label
1	muslimske	plagge/Clothe	label
1	nikab	skyte/Shoot	role
1	nikab	ram/Affect	role
1	nikab	opprør/Rebellion	role
1	nikab	forby/Forbid	role
1	nikab	delta/Participate	role
1	nikab	dekke/Cover	role
1	nikab	bruke/Use	role
1	nikab	bruk/Using	label

Table 9.14: Microportraits extracted from selected articles by *target terms nationalities*

count	query term	description	description type
3	iraker	bosatt/Settled	role
3	iraker	gammel/Old	property
2	iraker	kurdisk/Kurdish	property
2	irakere	kosovoalbaner/Kosovo-Albanian	label
2	iransk	bakgrunn/Background	label
2	iranske	ambassadør/Ambassador	label
1	afghansk	opprinnelse/Origin	label
1	afghansk	mann/Man	label
1	afghansk	kvinne/Woman	label
1	irakisk	kurdistan	label
1	irakisk	arabistan	label
1	iranernes	forsøk/Try	label
1	iransk	tvstasjon/ TV station	label
1	iranske	myndighet/Authority	label
1	iranske	revolusjonsgarden/ The revolution army	label
1	iranske	opposisjonell/Opposing	label
1	iranske	oljeinstallasjon/ Oil installation	label
1	iranske	myndighetsperson/ Authority figure	label
1	marokkansk	kvinne/Woman	label
1	pakistansk	politi/Police	label
1	pakistansk	militærhelikopter/ Military helicopter	label
1	pakistansk	kashmir	label
1	somalier	anklage/Accuse	role
1	somalisk	etnisitet/Ethnicity	label

1	somaliske	spesialstyrke/ Special force	label
1	somaliske	islamistgruppe/ Islamist group	label
1	somaliske	gruppe/Group	label
1	tyrker	bety/Meaning	role
1	tyrkisk	statsborgerskap/Citizenship	label
1	tyrkisk	offensiv/Offensive	label
1	tyrkiske	valgkampmøte/ Election campaign meeting	label
1	tyrkiske	terrorlovgivning/ Terrorism legislation	label

Table 9.15: Microportraits extracted from selected articles by *target terms norsk*

count	query term	description	description type
3	norskiraker	gammel	property
3	norskiraker	bosatt	role
1	norskiraner	hyre	role
1	norskpakistaner	politi	property
1	norskpakistaner	gammel	property
1	norskpakistaner	nordmann	label

9.9   Topic modelling

Figures 9.1 to 9.6 are graphs of the topic modelling results. The methodology behind the process is described in chapter 6, section 6.3.1. The results are presented in chapter 7 *Results*, section 7.2.2 and discussed in chapter 8, section 8.3.

Figures 9.1 and 9.2 show results extracted from the *target terms* by the LDA and TF-IDF models respectively. Figures 9.3 and 9.4 for *target terms nationalities*, and figures 9.5 and 9.6 for *target terms norsk*.
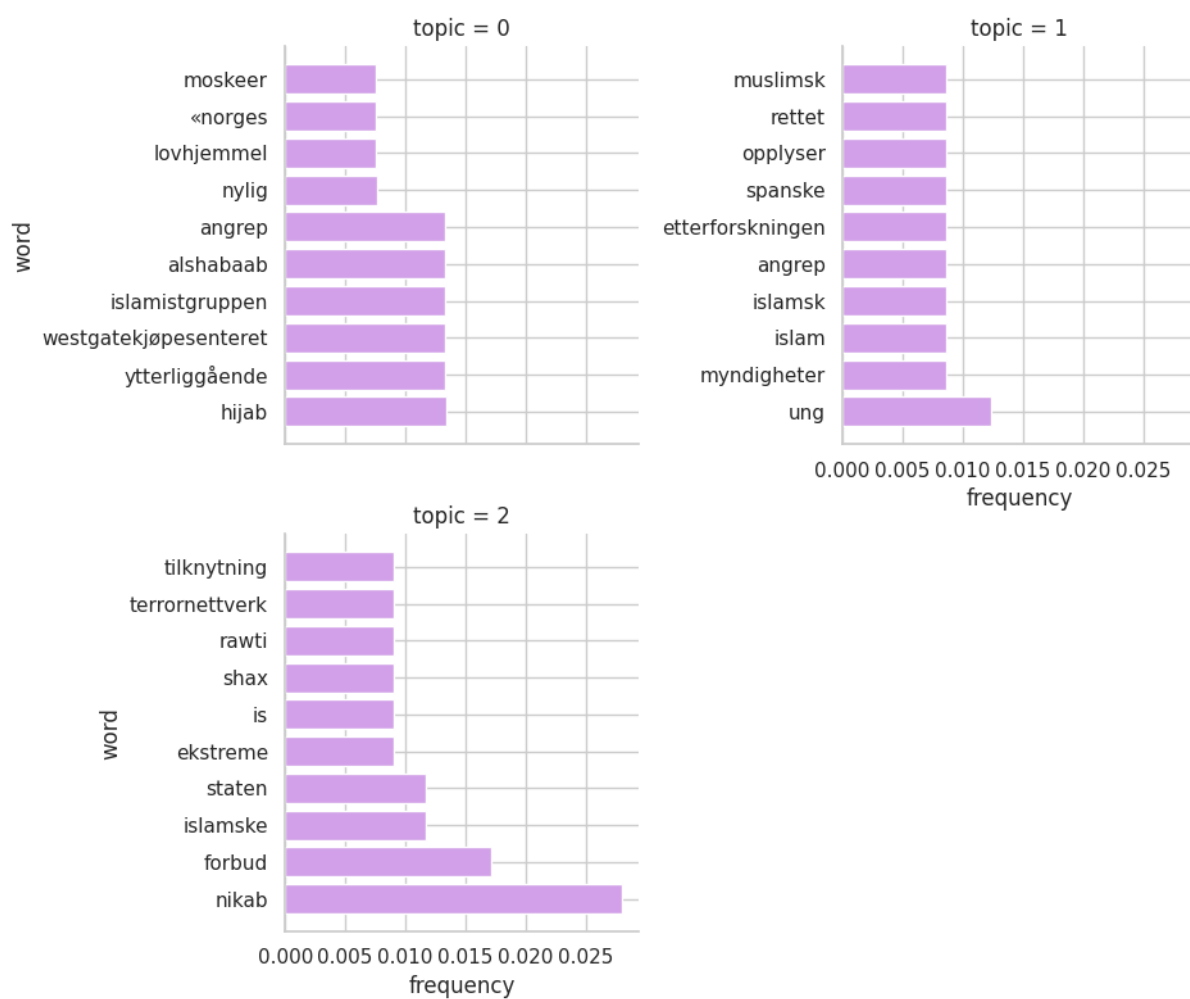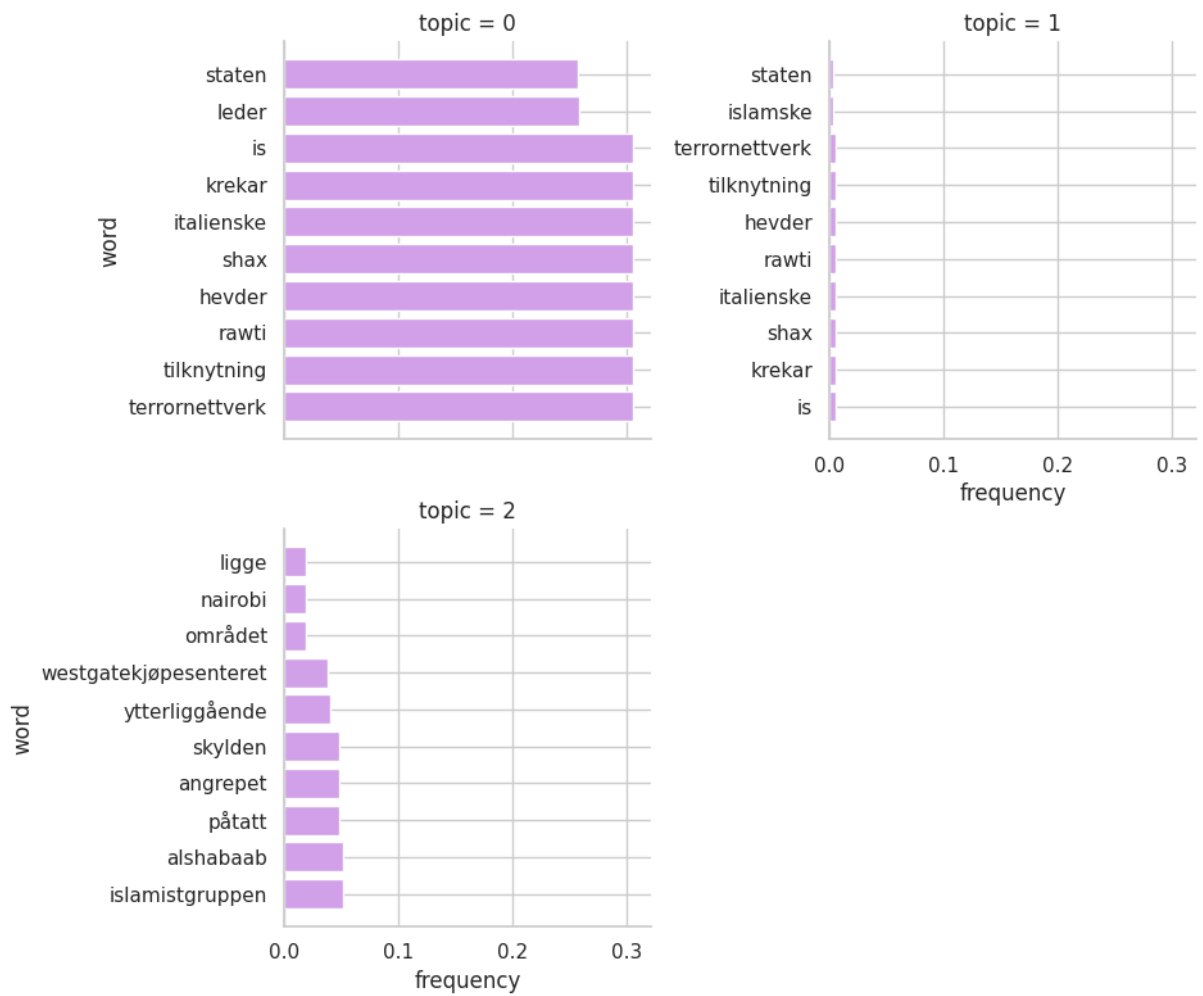
Figure 9.1: LDA results for target terms
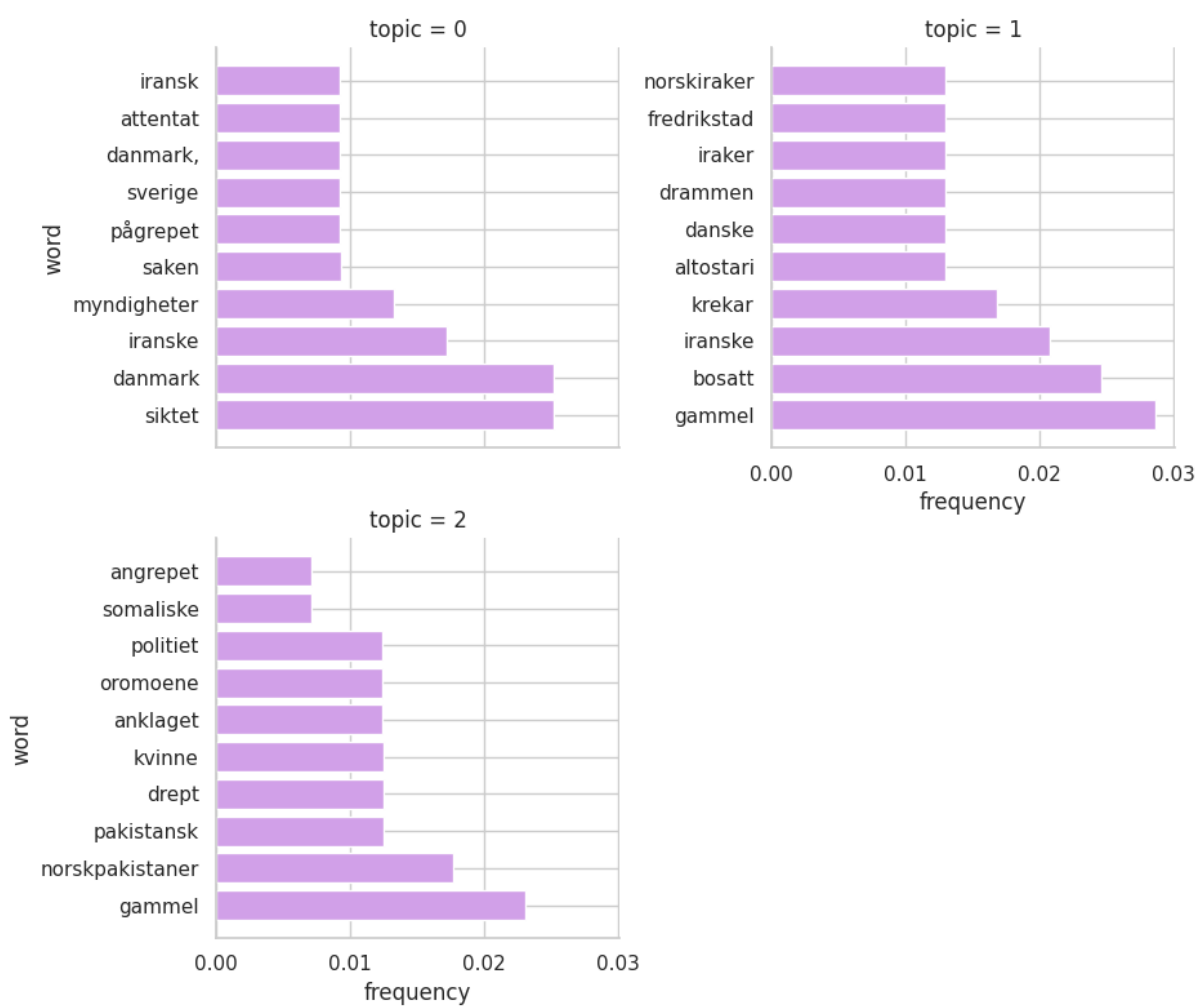
Figure 9.2: TFIDF results for target terms

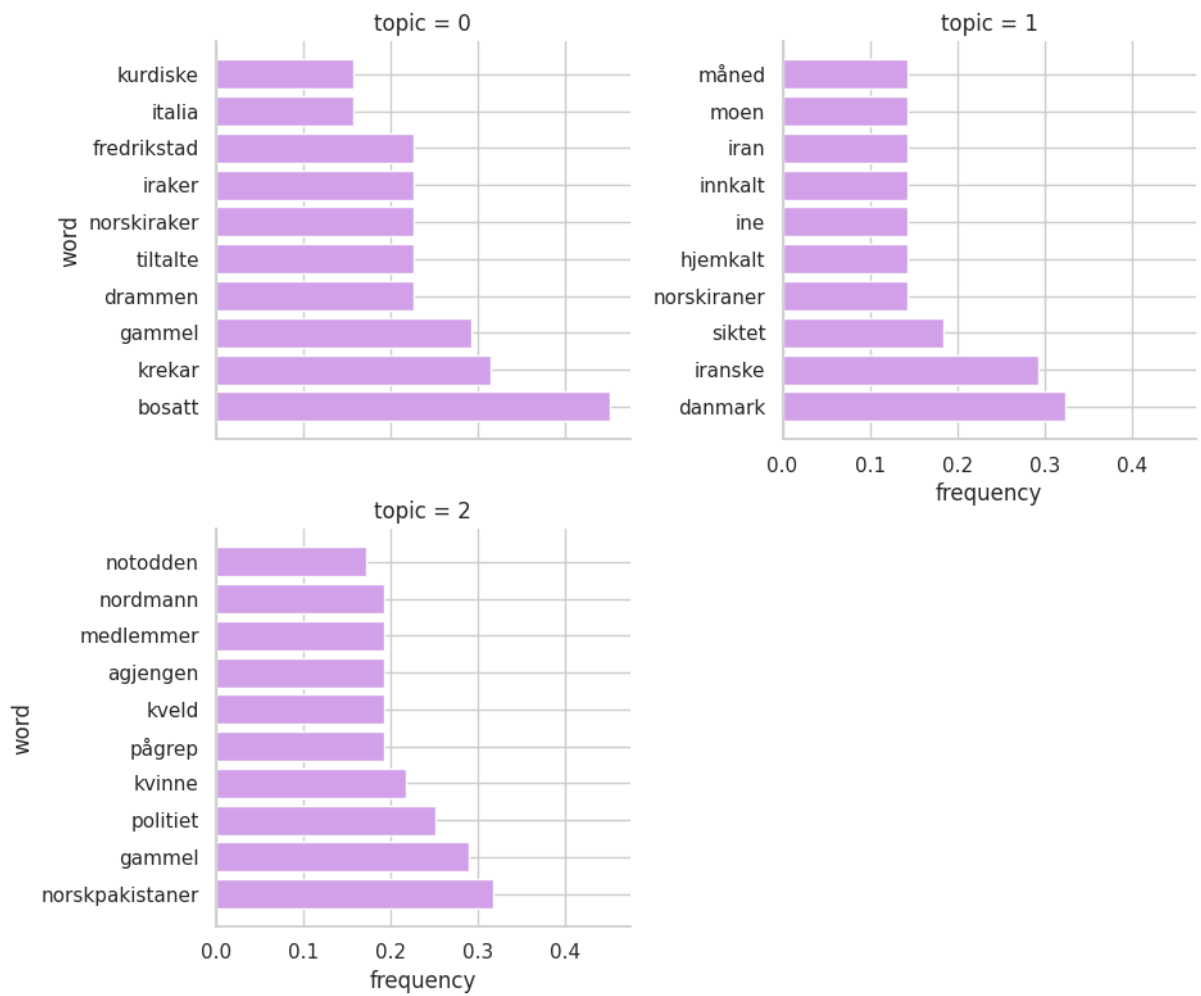**Figure 9.3:** LDA results for target terms nationalities

Figure 9.4: TFIDF results for target terms nationalities
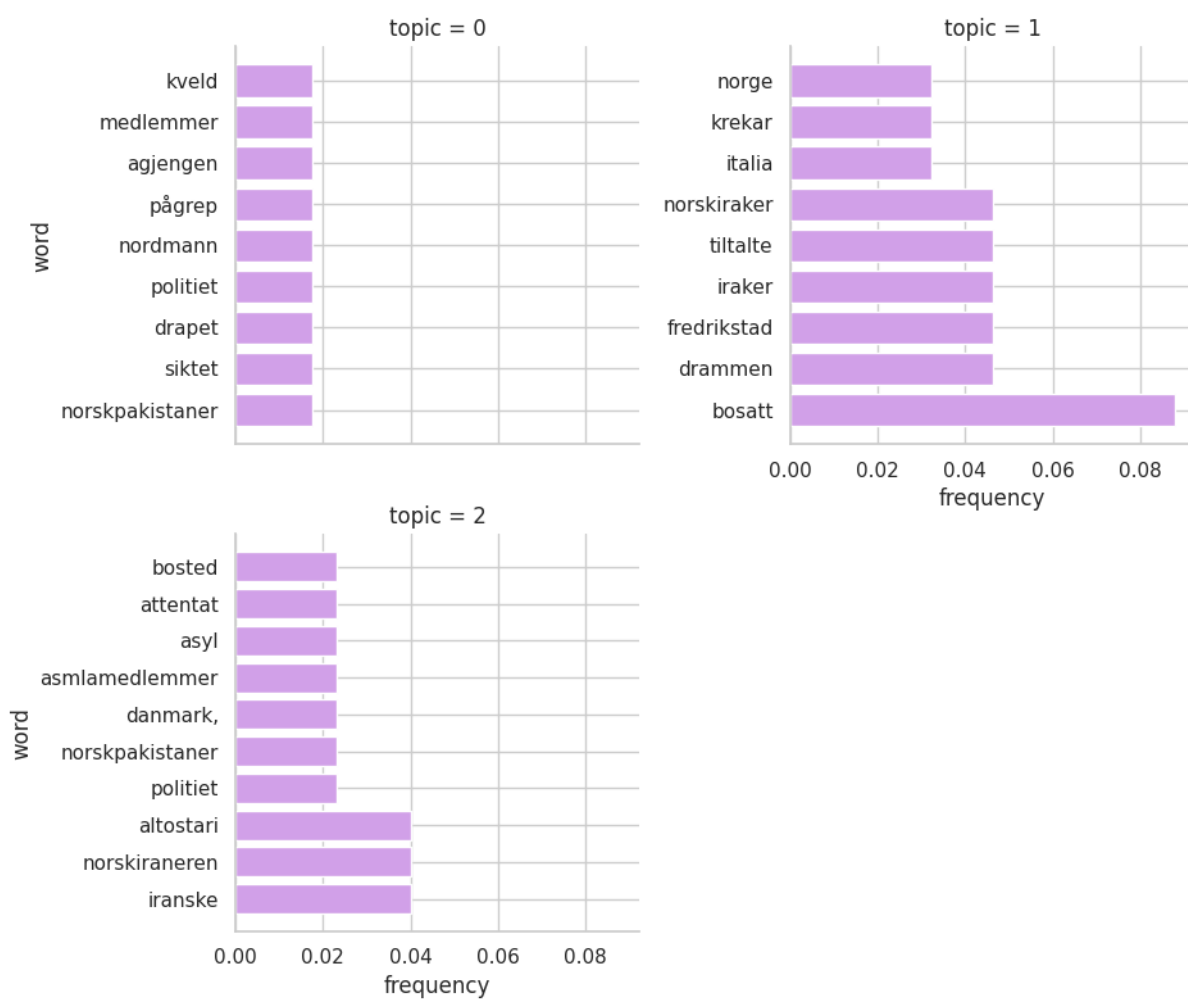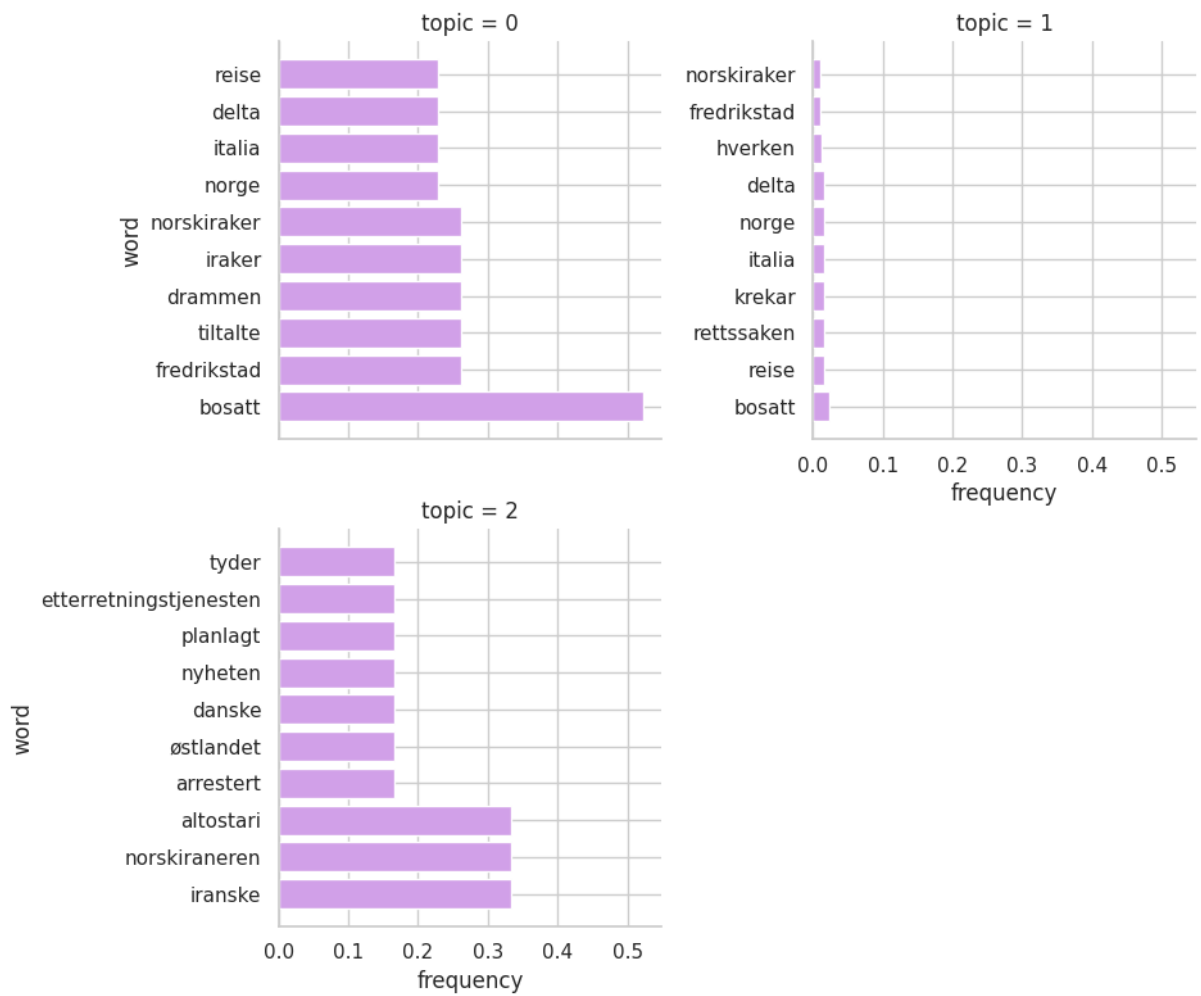
Figure 9.5: LDA results for target terms norsk

Figure 9.6: TFIDF results for target terms norsk

REFERENCES

Acronyms

AA Adresseavisen. 21, 23

AP Aftenposten. 17, 18, 19, 21, 23, 53

BoW Bag-of-Words. 15, 40

bs4 Beautiful Soup 4. 30, 35

BT Bergens Tidende. 22

CNN Convolution Neural Network. 15

DA Dagsavisen. 22

DB Dagbladet. 17, 22, 23, 44

DN Dagens Næringsliv. 22

FV Fædrelandsvennen. 22

HL-center The Norwegian Center for Holocaust and Minority Studies. 5, 8, 9, 10, 89

HRS Human Rights Service. 17, 18, 19

ICC Islamic Culture Centre Norway. 10

IRN Islamsk Råd Norge. 10, 23, 44, 54

KUN The Centre for Equality and Diversity. 8, 10

LDA Latent Dirichlet Analysis. 16, 17, 40, 46, 48, 59, 60, 61, 99

LDO The Equality and Anti-Discrimination Ombud. 9

Acronyms

MLP Multi-layer Perceptron. 15

NL Nordlys. 22

NLP Natural Language Processing. 2, 13, 17, 20, 30, 33, 36, 61, 65, 66

NLTK Natural Language Toolkit. 30, 40

PMI Pointwise Mutual Information. 40

PST Police Security Service. 52, 55

SA Stavanger Aftenblad. 22

SCSC Social Categories and Stereotypes Communication. 14, 29

SIAN Stopp Islamiseringen av Norge. 7, 23

SVM Support Vector Machine. 14

TF-IDF Term Frequency-Inverse Document Frequency. 14, 15, 40, 46, 48, 49, 59, 99

UD Universal Dependencies. 30, 39

VG Verdens Gang. 9, 18, 19, 22, 54