Muslims in social media discourse: Combining topic modeling and critical discourse analysis

Anton Törnberg [a,*], Petter Törnberg [b]

[a] *Department of Sociology and Work Science, Gothenburg University, Skanstorget 18, Box 720, 405 30 Gothenburg, Sweden*
[b] *Complex Systems Group, Physical Resource Theory, Chalmers University of Technology, Gothenburg, Sweden*

ABSTRACT

This article combines topic modeling and critical discourse analysis to examine patterns of representation around the words Muslim and Islam in a 105 million word corpus of a large Swedish Internet forum from 2000 to 2013. Despite the increased importance of social media in the (re)production of discursive power in society, this is the first study of its kind. The analysis shows that Muslims are portrayed in the forum as a homogeneous outgroup that is embroiled in conflict, violence and extremism: characteristics that are described as emanating from Islam as a religion. These patterns are strikingly similar to – but often more extreme versions of – those previously found in analysis of traditional media. This indicates that, in this case, the internet forum seems to serve as an "online amplifier" that reflects and reinforces existing discourses in traditional media, which is likely to result in even stronger polarizing effects on public discourses.

1. Introduction

In September 2012, the posting of the anti-Islamic film on YouTube entitled "Innocence of Islam" sparked a wildfire of manifestations of indignation all over the world, including both peaceful demonstrations and violent riots. These events were followed by intense discussions on the portrayal of Muslims and Islam in social media. Indeed, this film is only one of innumerable examples of what has been described as a trend toward an increasingly harsh online climate for Muslims, and a general growth of Islamophobic content in social media (Awan, 2016a; Matters, 2015; Oboler, 2013).

While the representation of Muslims and Islam in traditional media, such as newspapers, has been relatively well-studied (Baker et al., 2013a, 2013b; Hafez, 2000; Moore et al., 2008; Poole, 2002; Richardson, 2004), there are significantly less studies that focus on these processes within social media. This does not only apply to Muslims – social media discourse in general constitutes a bit of a blind spot for academic research. And it is a fast-growing blind spot: social media is becoming an important source for the (re)production of discursive power in society, while simultaneously constituting a unique source for studying everyday discourses outside the scope of mass media. A central reason for this striking lack of studies is the methodological difficulties relating to handling and analyzing the large amounts of unstructured textual data that often characterizes social media.

Based on the above, the purpose of this article is to investigate the representation and discursive construction of Muslims and Islam in social media. This article is part of a larger project that focuses on the representation of feminism and Islam in social media. When using "representation" here, we refer to how the meanings of these words are constructed in relation to, for instance, the linguistic context in which they are used. We have previously explored how discourses around Muslims and feminism are constructed, and investigated how these discourses are interrelated, both discursively and through user discursive networks (see Törnberg and Törnberg, 2016). In this particular article we focus on Flashback, which is the largest internet forum in Sweden and has a reputation for right-leaning bias. In the time of writing, the forum has 1 025 264 registered users and about 2.3 million unique visitors per week; figures that put it among one of the largest online forums in the world.[1] Due to its size and

---

* Corresponding author. Tel.: +46 739566502.
E-mail addresses: anton.tornberg@gu.se,
tornberganton@gmail.com (A. Törnberg), pettert@chalmers.se (P. Törnberg).

---

[1] This can be compared with the population of Sweden, which is currently about 9.5 million.

scope, this forum is argued to have a function comparable to that of traditional newspapers when it comes to producing and spreading societal discourses. Indeed, according to available figures, Flashback has more unique visitors per week than Sweden's two leading daily press newspapers, even when adding the number of their paper edition subscribers.[2]

By using Critical Discourse Analysis (CDA), we investigate how Muslims and Islam are represented in this forum compared with how they are represented in traditional, Western media. To resolve the problems that have so far limited this type of study due to large quantities of data, we complement CDA with topic modeling, which is a type of statistical model using hierarchical probabilistic modeling that was developed within the field of computer science (Blei et al., 2003). By providing an overview or "content map" of the corpus, topic modeling provides an enriching complement to CDA by aiding discovery and adding analytical rigor. This approach enables us to reveal linguistic patterns in extensive corpuses that would be too labor intensive to uncover by hand (Baker, 2006). Thus, beyond the empirical results, this article also has a methodological contribution and explores the mutual benefits of combining CDA and topic modeling.

The analysis shows that Muslims are portrayed in the forum as a homogeneous outgroup, embroiled in conflict, violence and extremism: characteristics that are described as emanating from Islam as a religion. Despite that the language usage tends to be aggressive and conflict-oriented in the forum, these general patterns are strikingly similar to those found in traditional media by previous research. These results contribute to the ongoing academic discussion on the relation between social media and traditional media, concerning whether social media constitutes mainly a "safe haven" for developing and sustaining alternative discourses, or rather as an "online amplifier" that reflects and reinforces existing hegemonic discourses, which may result in even stronger polarizing effects on public discourses.

The disposition of this paper is as follows. We start by positioning our study within existing research on the discursive representation of Muslims and Islam in media. We then assess the relevance of studying online Islamophobia. Following this, we present our methodological approach and explain how topic modeling can allow researchers to inductively structure large quantities of texts. Here, we also describe how we have collected our data and how it was analyzed. This is followed by our analysis in which we identify the discursive landscape surrounding and (re)contextualizing Muslims and Islam.

2. The representation of Islam in media

A vast amount of academic literature and reports show how Muslims and Islam are often reported and represented negatively and stereotypically in Western news media. In a comprehensive book, Baker et al. (2013a) investigated the representation of Muslims and Islam in British press from 1998 to 2009, which shows the presence of both explicit Islamophobic representations, as well as what seems to be more common – a more subtle, implicit and ambivalent picture, which indirectly contributes to negative stereotypes. These representations are manifested in various ways. One example is the high use and concurrence of the words "terrorism" and "extremism" with Muslims and Islam,

which places Muslims in a context of conflict. In particular, the word "Islamic" was shown to carry extremely negative discourse prosody, as it was heavily associated with religious and political extremism, militancy and terror. Furthermore, Muslim women are often portrayed as victims and Muslim men as potential aggressors.

Other studies show similar patterns. Moore et al. (2008) used content analysis of 974 articles on Islam in British press from 2000 to 2008, and found a general increase over time of articles that focus on extremism and differences between Islamic culture and "the West". At the same time, stories about attacks against Muslims and other problems that they face decrease in the same period. Poole (2002) analyzed all articles on British Muslims in the *Guardian/Observer* and the *Times/Sunday Times* from 1993 to 1997, showing that Muslims are frequently represented as irrational and antiquated, threatening to liberal values and democracy, extremists and fanatical. Manan (2008) shows similar results when analyzing the magazines *Time* and *Newsweek*, and argues that through the strategy of the negative "other" presentation, the media represented "Islam and its many followers as deviant, volatile, evil, and anti-modern" (p.124). Richardson (2004) finds four argumentative themes consistently associated with Islam in British broadsheets in 1997: military threat, association with terrorists/extremists, a threat to democracy and a sexist/social threat. Hafez (2000) shows that similar negative patterns also go back well before the September 11 terrorist attacks of 2001.

However, considerably less scholarly work has focused on anti-Muslim and anti-Islamic sentiments in social media, such as Facebook, Twitter and internet forums. As Copsey et al. (2013: 10) and Awan (2016a: 32) explicitly state, online Islamophobia remains under-researched, both on a policy level and an academic level. Most existing research is located within the field of cyber hate, where Islamophobia is generally regarded as only one topic among several types of hate alongside, for example, anti-Semitism, misogyny, homophobia and general racism. Although limited in its scope, the existing academic literature seems to suggest similar tendencies of a general negative bias within social media. Statistics from monitoring organizations, such as Tell Mama, reports of a significant increase in Islamophobic discourses and abuse online, indicating a growing trend of hardening attitudes towards Muslims (Feldman and Littler, 2014; Copsey et al., 2013).

In a newly published book, which according to the authors is "the first comprehensive critique of online Islamophobia" (Awan, 2016a: 5), the authors compile several studies on this topic, and focus on both the victims and the perpetrators of abuse. For instance, Oboler (2016) investigates how Facebook is being used to normalize Islamophobia. Through a qualitative analysis of 349 posts on Facebook, he finds several themes that depict Muslims as a security threat (e.g. represented as terrorists or rapists), as a threat to our way of life and as manipulative and dishonest (see also Oboler, 2013). In another chapter, Awan (2016b) investigates both how Muslims are viewed on Twitter and the characteristics of the offenders. By examining 500 separate tweets between January 2013 and April 2014, he found a common reappearance of words used to describe Muslims, including "terrorist", "pedos", "scum", "Yusrats", and "Pigs". Similarly, Copsey et al. (2013) have studied how far-right groups on social media represent Muslims as a dangerous, untruthful "out group", and commonly refer to them in terms of "dirt" and "filth", along with accusations of rape, pedophilia, incest and terrorism.

3. The relevance of online islamophobia

There are three main reasons for studying online Islamophobia. Firstly, we are currently seeing a trend in which traditional news

---

[2] This is intended to serve only as a loose approximation, since getting accurate and trustworthy figures on website statistics is close to impossible. The numbers for the two largest daily newspapers (Dagens Nyheter and Svenska Dagbladet) come from the KIA-index, a website that measures media impact. For offline editions, we rely on the Orvesto Konsument Index. As Flashback is not included by KIA, we are forced to rely on the figures that Flashback have provided themselves.

sources and mass media give way to emerging social media platforms. Since the last decade or so, there has been an explosive growth in the use of social media and social networking sites and they are progressively becoming both an important platform for social interaction, as well as an increasingly important source for receiving and producing daily news. While new social media clearly shares many traits and functions often ascribed to traditional mass media, not least by framing issues and events and thus shaping people's perceptions of reality and of social and political issues (Moscovici and Duveen, 2000), there are also important differences. Social media builds upon interaction between networked individuals who collaboratively share their narratives by creating and managing content (Klinger and Svensson, 2014). In this sense, the increase of social media usage also marks a shift from media consumers and passive observers to content creators. Through this shift towards user-driven, participatory information exchange, there are reasons to assume that the growth of new social media may bring with it new media dynamics, which could relocate the construction and formulation of societal discourses. Thus, traditional media no longer has a "privileged access to discourse" (Van Dijk, 1993b: 255), where power and discursive dominance are simply imposed from the elite who use mass media as channels. Therefore, we need to pay more attention to how power and dominance can be jointly produced through mass-interaction from the bottom up.

Secondly, the interplay between traditional and new social media is becoming a genre in itself and the literature is beginning to examine the dynamics and consequences of this emerging relationship, and particularly its effects on the classic agenda-setting power of the media (Klinger and Svensson, 2014; Sayre et al., 2010; Van Dijck and Poell, 2013). There is intense discussion on whether social media functions primarily as a "safe haven" for groups and individuals who use it to produce and spread counter-narratives in conflict with hegemonic values and ideologies of mainstream society (see e.g. Brunsting and Postmes, 2002; Cammaerts, 2012; Diani, 2000; Koster and Houtman, 2008), or mainly as an "online amplifier" that mirrors and reinforces offline social structures and the prevailing discourses in traditional media (for an overview, see Lilleker et al., 2011). Most research so far is based on speculation rather than broad-scope empirical work, and more attention is needed to address this evolving relationship (Sayre et al., 2010).

Last but not least, cyberbullying and online abuse tend to have a detrimental impact on the victims and their families, which often leads to anger, shock, anxiety, depression and fear that the verbal threats will materialize into physical action (Allen et al., 2013; Awan, 2016b). More research is needed on how to confront these pressing issues.

To sum up, online Islamophobia clearly contributes to legitimizing xenophobic discourses, normalizing abusive behavior and providing discursive opportunities that encourage physical attacks by propelling and fueling hatred, thus the risk of turning opinions into action (Koopmans and Olzak, 2004). While many would agree that more effort needs to be put into studying online Islamophobia, and online hatred in general, interested scholars have often been limited by multiple methodological challenges, with perhaps the foremost being the sheer amount of unstructured textual content that characterizes social media (Bail, 2014; Karpf, 2012). Even relatively small data sets can be difficult to approach as it is hard to delineate, select and confine materials from millions of texts, posts or tweets. Making matters worse, these texts are often short, lack discursive context and vary in complex and highly non-linear ways, which makes them difficult to extract and study using established methodological and analytical approaches. As a consequence, the rather few existing studies using such data have generally been limited to either qualitative studies of small samples, or relatively simple quantitative studies that focus on word frequency lists.

4. Design and method

To address these issues, this article combines a corpus-linguistic (CL) approach with Critical Discourse Analysis (CDA) to investigate the representation of Muslims and Islam in social media. With this approach, we [i] will enable the study of this vast amount of unstructured data and [ii] intend to address some methodological weaknesses often raised concerning CDA, such as lack of academic rigor in that the analyst's subjective preconceptions and desired results may affect the outcome of the analysis (Fowler, 1996; Orpin, 2005; Widdowson, 1996; Widdowson, 2000). For instance, discourse analysis is often criticized for the imminent risk of "cherry picking" (i.e. that the author "picks a text to prove a point"), which leads to obvious problems relating to representativeness and generalizability (Baker et al., 2008; Stubbs, 1994, 1997). Another criticism that is often raised concerns small data sets, which creates the risk of neglecting linguistic patterns that are less frequent or only cumulatively frequent. Many documents often only contain bits and pieces of ideologies, arguments and discourses – small but systematic patterns and tendencies that may not be visible to the naked eye when restricted to small-n studies (Stubbs, 1994). Due to the large data quantities, these issues are likely to be aggravated in the context of social media.

Since this CL approach is combined with CDA, this study is placed within the relatively recently developed but quickly growing field of Corpus-Assisted Discourse Studies (CADS). This is basically an umbrella term for approaches that combine Discourse Analysis and techniques for corpus enquiries from CL (Cheng, 2014; Partington, 2006; Wodak and Meyer, 2009), either in the form of a methodological synthesis (Baker et al., 2008), or as separate components combined in a triangular way (Baker and Levo, 2015). CL consists of various empirical methods with the aim of finding "probabilities, trends, patterns, co-occurrences of elements, features or groupings of features" (Teubert and Krishnamurthy, 2007: 6). While there are several different techniques, the focus generally lies in the analysis of words and their textual context, using word lists, keywords, collocations and concordances, backed up with measures of statistical significance (Cheng, 2014).

Instead of using the conventional tools and techniques within CL, this study uses Topic Modeling, which is a technique developed in computer science that inductively finds recurring clusters of co-occurring words in a text (see section below). While sharing many common traits with other, more common techniques in CL and CADS, an important difference is that unsupervised topic models inductively structure the data *without* using any pre-set keywords. This means that it is corpus-driven (i.e. the analysis is driven by whatever patterns are salient in the data itself) (Tognini-Bonelli, 2001). While there are other unsupervised CL-techniques (e.g. cluster analysis and word frequency lists), it is more common to search for certain key words and study them using, for example, collocation analysis (Pollach, 2012).

This unique feature of topic modeling allows us to investigate different topics that relate to Muslims and Islam without the limitations of any predefined hypotheses. Of course, this does not mean that the analysis is free from the authors' biases and unconscious preconceptions. While association patterns represent quantitative relations and hence can be measured, interpretation is fundamentally a qualitative enterprise, involving interpretation, expert knowledge of the specific text material in question, and calibration of the tool to produce coherent and interpretable topics. Or, as Baker (2006: 179) puts it: data often needs to be

"subtly massaged" in order to produce desired results. Because of these factors, this process can never claim full objectivity, but it *can* allow higher levels of intersubjectivity and systematicity; for example, consistency in the use of tools (Marko, 2008: 92). Thus, the approach taken here can be described as a form of qualitatively-informed corpus-based analysis.

When it comes to the use of CDA, we follow Baker et al. (2013b) in that we do not aim to simply replicate conventional CDA-practices and modes of analysis on larger data sets. Instead, we focus on using topic modeling to achieve aims compatible with those of CDA, departing from the notion that CDA is a heterogeneous research program (Wodak and Meyer, 2009), not limited to certain concepts or modes of analysis, and united in a shared purpose which is to study "the way social power abuse, dominance, and inequality are enacted, reproduced, and resisted by text and talk in the social and political context" (Van Dijk, 2003: 353). Our analysis also makes use of certain concepts and practices often used in CDA. Most notable of these is Van Dijk (2009) definition of discourses as "(re)contextualization of practices". In other words, social practices happening in the "material world" are reformulated in texts and discussions. While this reformulation or recontextualization is always selective and from a certain point of view, it is not a matter of a pure articulation of subjectivity or ideology – discourses always work with the material of the world. Thus, elites and other groups attempting to assert hegemonic dominance use discourses to recontextualize material events in line with their own interests. We thus use topic modeling to analyze how Muslims and Islam are recontextualized through various "topics".

4.1. Topic modeling

Topic modeling is a collection of methods and algorithms that uncover the latent thematic structure in document collections by revealing recurring clusters of co-occurring words. This methodology allows quantification and visualization of themes to arise inductively from texts. The algorithm that is currently the most used for topic modeling, which is also the algorithm applied in this paper, is Latent Dirichlet Allocation (LDA) (Blei et al., 2003).

The logic behind topic modeling is that a document about a certain topic is more likely to contain words associated to that particular topic. For example, if a document is about brain surgery, it is more likely to include words like "neocortex" than a document about, say, cats. Correspondingly, it is less likely than the latter to contain words like "purr" or "feline". Words that are not related to either of these topics, such as "is" or "and", should appear with more-or-less the same frequency in both documents. Topic modeling is basically an algorithmic operationalization of this simple logic, which defines a topic as a list of words with different assigned probabilities and attempts to find the set of topics that best capture the documents. It starts by counting the words in each document and listing how many times each word appears. It then seeks a pre-selected number of topics within each document in order to reduce them to as few topics as possible, without "cheating" by allowing the topics to include too many different words. These rules result in topics that are useful summaries of the topics in the documents.

The algorithms that LDA uses are based on Bayesian statistical theory (Gelman et al., 2014), where the topics and the per-document topic proportions are seen as latent variables in a hierarchical probabilistic model. The conditional distribution of those variables is approximated, given an observed collection of documents. When applied to the documents in a corpus, inference produces a set of topics and for each document, an estimate of its topic proportions and to which topic each observed word is assigned. For a more technical description of LDA, we refer the reader to Blei (2012) and Blei et al. (2003).

4.1.1. Data and procedure

The corpus for this study is extracted from Flashback, which is currently one of the largest web forums in the world. At the time of writing, there are 1 025 264 registered users and 53 601 028 posts that are contained within different sub-forums and threads, that are about a broad range of subjects. The forum is fully open to the public and only modestly moderated by a set of privileged users. The forum is expanding with about 15 000–20 000 new posts per day, and has around 2.3 million unique visitors per week. As mentioned, if it were part of traditional media then these figures would make it the largest daily mail newspaper in Sweden. While the forum is indeed very broad and includes discussions on many different topics, it has a reputation for leaning towards extreme-right opinions and it is often mentioned in Swedish media as a hub for online hatred and xenophobia.

By using customized web crawlers, we downloaded and anonymized the entire content of the forum between May 2000 and May 2013, which comprised 50 million posts in total, and placed it into a local database. To construct the corpus for the analysis, we selected all posts containing the case-insensitive search terms "muslim" and/or "islam", which resulted in a corpus of 746 062 posts (about 105 million words).

There is obviously a risk that a number of potentially relevant posts could accidentally be excluded because they either discuss the specific topic indirectly or refer to it by using slang and common derogatory terms such as "paki", "raghead", "sand nigger" or "culture-enricher". However, we consider this a minor problem since [i] there are few users that only use such slang words, most combine them with, for example, the word "muslims", and [ii] most posts – as conventional in discussions on internet forums – cite a previous submission that will in turn most likely use the specific terms, and they will therefore be included in the analysis. Furthermore, because of the large data quantities, small data losses are not likely to be a problem. A close review of a sample of the selection indicates that our selection criteria and keywords seem to capture the intended data.

For technical reasons, standard LDA generally works best for documents with a size of at least 1000 words. We therefore aggregate all posts from each individual user in a specific thread within the same time period into chunks of 1000-word documents. Posts that are significantly longer than 1000 words are split up into smaller chunks.

Furthermore, we excluded a number of stop words from the analysis by using off-the-shelf lists for Swedish, Norwegian, Danish and English, further extended with slang and abbreviations commonly used on the forum. While lemmatizing the corpus can often be useful for producing better topics, this proved difficult in this case due to the often informal language and the amount of slang, abbreviations and misspellings. To investigate changes over time, we performed a separate LDA each year, which enabled the study of both consistent topics and new, salient topics that emerge in a specific year. An alternative to this would be to perform a single run for the entire material, which would indeed facilitate comparison over time. However, this approach would be too fixed and inflexible since we would thereby most likely miss smaller topics that are unique for a shorter time period, but still not sufficiently salient in the corpus overall for them to appear in the model. To perform the LDA model we used Big Text Tool[3], which is an online-based application that includes various tools and techniques for automated text analysis and graphical illustrations. This

---

[3] Big Text tool is available for download: http://www.toernberg.com.

application is free, easy to use and is particularly customized for social scientific studies that use large corpuses.

5. Methods of analysis

As stated above, the corpus consists of all posts containing keywords relating to Muslims and Islam. In the analysis, we use topic modeling to inductively find topics within this corpus. This means that we may also capture separate topics that, while not explicitly relating to Muslims, are nonetheless often raised and discussed in parallel to discussions on Muslims/Islam. In this sense, topic modeling captures the linguistic discursive landscape surrounding Muslims and Islam (see Fig. 1). This discursive landscape is argued to (re)contextualize Muslims and Islam, and thus influence the understanding and meaning of these words (Van Dijk, 1993a, 2009). In other words, the topics produced through topic modeling arguably represent events or issues in response to which images of Muslims and Islam are reproduced and reinvented. For instance, if discussions on "terrorism" or "sexual abuse" are frequently raised in relation to Islam, this inevitably affects the meaning of the words Muslims and Islam. Thus, this way of applying topic modeling has certain similarities with the methodological approach taken by Baker et al. (2013b), where common collocates to the word Muslims are examined and put into thematic categories through manual concordance analysis. The difference is that this process is automatic when using topic modeling.

The focus in the subsequent analysis is to investigate the discursive contexts in which Muslims and Islam are discussed and whether these contexts are changing over time. We particularly look for shared, underlying strategies and tendencies within and between these different contexts. We do this by focusing on central topics relating to Muslims and Islam, and investigate how these evolve over the years. We use topic modeling to generate the topics and then qualitatively analyze both their constituent words and the documents relating to these topics. This means that the qualitative and quantitative aspects of the analysis cannot be neatly separated, but are rather integrated as two sides of the same coin.

The emphasis in the analysis lies on the larger, more significant topics that are consistent throughout time, as well as large and temporally salient topics that are characteristic of certain periods. We distinguish between these types of topics as *continuous topics* and event-oriented *episodically salient topics.*[4] For instance, while topics about sexual abuse and violence remain as constant topics throughout the years, more specific topics about mosques and minarets, for example, vary greatly between the years, and are driven by certain events such as the construction of a new mosque. Large, background topics (consisting mainly of common words) are left out of the analysis, as well as small and insignificant topics.

Furthermore, we distinguish between *topics* and *topic categories*. While the former refers to single topics, the latter is a group of topics that belong to a common subject area; that is, all topics that discuss an arguably similar issue or event, both in the same year and over time. These heuristic topic categories are hence manually constructed through close study of the constituting topics, with the purpose of enabling an overview of the results and to facilitate the analysis. It is thus important to note that there are often no strict, rigorous membership criteria for deciding which topic category a topic belongs to. Often the boundaries are rather
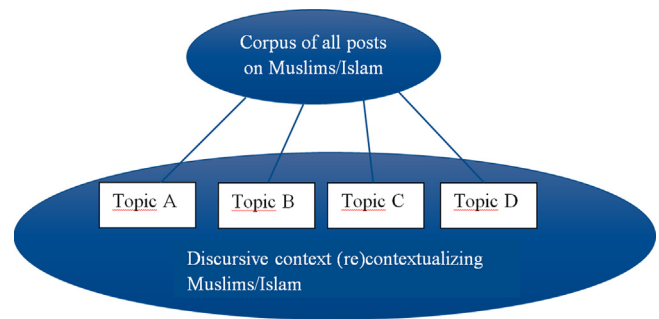

Fig. 1. This figure illustrates how the topics generated using LDA capture the discursive landscape surrounding Muslims and Islam, thus serving to (re)contextualizing these words.

ambiguous and topic categories may be more-or-less overlapping, where certain topics may be classified within more than one category. Additionally, this means that the decision on how wide and inclusive the categories are set to thereby depends on the level and resolution of the analysis. Thus, a broad topic category consists of a number of topics that in turn can be thematized within smaller, more specific topic categories.

Finally, we also analyze the development of the topics over time. We do this by first by adding the strength/size value[5] of all topics within each topic category for each year. In this way, we get the proportional size of the specific topic category in relation to the other topics in the same year. We then plot the results for the continuous topic categories in an area graph over time to explore changes in how salient the topics are each year.

It should be noted that this is an approximate measure, suitable for exploration and illustration rather than being seen as rigorous and statistically reliable. Since the construction of the topic categories is manual and subjective, certain topics may just as well be placed in another adjacent topic category, depending both on the purpose with the analysis and the subjective bias/interpretation of the analyst. This may have large effects when plotting the size of the topic categories, particularly if the specific topic has a very large proportional size in a certain year. Accordingly, this means that the comparison over time works best for more specific, event-based topics rather than broader, more abstract topic categories that consist of different topics that are joined together based on more implicit, underlying theoretical analogy. Noteworthy is also the rather "spiky" feature of the diagrams (see e.g. Figs. 3 and 4) with topics often peaking and declining rapidly between the years, which stems from the often event-focused discussions that characterize internet forums; that is, discussions and threads often focus on external events that provoke short, but intense discussions and debate.

6. Results and analysis

In the following analysis, we start by presenting the continuous topic categories separately, which are exemplified with short quotes/extracts from posts within each category. We also look at the broader changes of these topic categories over time. We then briefly focus on some of the temporally salient topics and investigate how these develop over the years. Finally, we analyze common trends, patterns and strategies used within and between these topic categories. What emerges from the analysis is a representation of Muslims as a homogeneous group, embroiled in

---

[4] However, it should be noted that this distinction between *continuous* and *episodically salient* topics is not analytically strict, since episodical topics also may be more-or-less persistent and range continuously over time.

[5] The size/strength value of a topic is defined as the sum of the connections of all the documents that are linked to the topic, as provided by the topic model run.
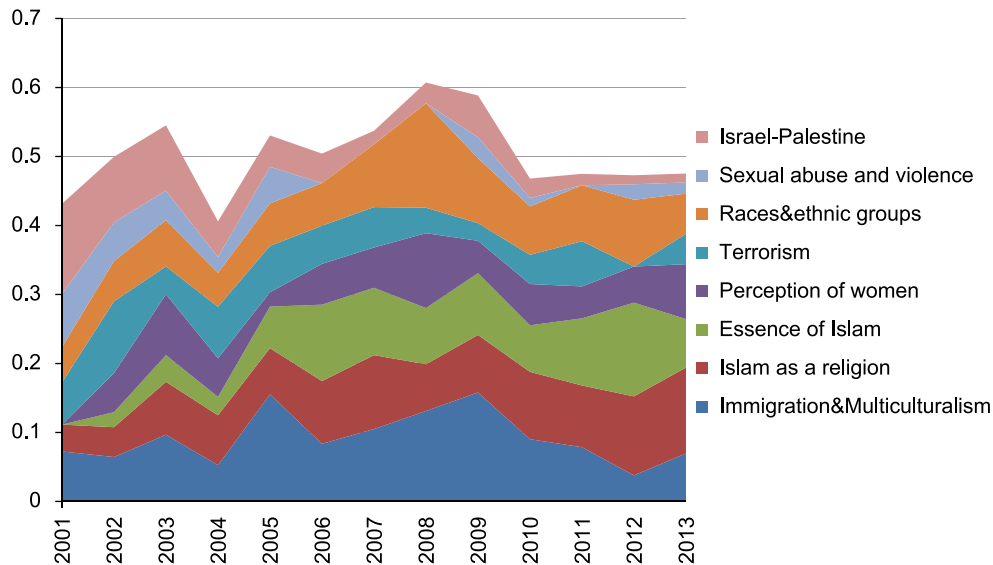
Fig. 2. This area graph plots the development over time for the continuous topic categories. The graph shows the proportional size of the topic categories in relation to all other topics in the same year and is calculated by adding the strength value of all topics within each topic category/year.

conflict and violence, which is described as stemming from the very essence of Islam as a religion.

As can be seen in Fig. S1 in the Supplementary material, the most significant, continuous topic categories are [i] Immigration and multiculturalism, [ii] Islam as a religion, [iii] The essence of Islam, [iv] The perception of women, [v] Terrorism [vi] Sexual assault and violent crimes, [vii] "Race" and ethnic groups, and the [viii] Israel and Palestine conflict.

6.1. Immigration and multiculturalism

The first topic category is rather broad, comprising topics that focus on various aspects and consequences of immigration and multiculturalism, and remains highly central throughout the time period. The focus lies on immigrants in general, but also on specific ethnic groups that are portrayed as particularly problematic. Much of the discussions circle around issues such as immigrant-dense suburbs, criminality and whether immigrants tend to overuse welfare benefits, evade taxes, steal jobs, not adjust to Swedish culture, and constitute a threat to "Swedish identities". As one user puts it:

> The high level of crime among these people [muslims] is just one of the problems. Another problem is their unwillingness to integrate with the swedes and their despise of swedish women (and in many cases also swedish men). Another problem is that they come from societies that are a 1000 years behind us in development and it's a mystery to me how politicians thought this kind of prehistoric people would be possible to integrate into our modern western hitech-society.

There are also discussions on aspects that are claimed to clash with what is considered traditional Swedish culture and values, including, for example, religious private schools, genital mutila-tion, Ramadan, veils, a democratic and secular society, and "Swedish" traditions such as Christmas. When looking more clo-sely at the words associated to the topics and related documents, we may discern a possible shift in focus over time, with increasing focus on Muslims and Islam in this category (see Fig. S1 in the Supplementary material). The focus in the first years lies foremost on immigrants in general: "blacks", "Romani people", "Turks", "multiculture", "Arabs" and "refugees", but also on derogatory terms such as "trash", "dreg" and "vermin". However, from around

2005 there seems to be an increased particular focus within these topics on Arabs and Muslims – a tendency that is even more apparent during the following years. From around 2009, it appears that this topic category might be decreasing somewhat in the corpus overall; albeit, this is difficult to statistically confirm for reasons stated above (see Fig. 2).

6.2. Islam as a religion

The second topic category collects topics that focus on *Islam as a religion*; that is, both in comparison with other religions and with a particular focus on religious practices and various orientations of Islam (e.g. Sunni, Shia and political Islam). These topics are central and also remain internally consistent over time, although we can perhaps discern a gradual increase of these topics starting around 2005 (see Fig. 2). While some discussions in this category seem to be more-or-less explicitly related to conflict, many are also rather balanced discussions that compare and discuss different religions. A central reason for this is likely to be that these discussions come from the large sub-forum Religion, which focuses particularly on different religions, beliefs, and practices. This is rather interesting since it seems like topic modeling tends to separate these dis-cussions from the more critical ones that are categorized in the topic category called *the essence of Islam*.

6.3. The essence of Islam

The topic category that we call *the essence of Islam* is somewhat connected to the previous category but focuses on the claimed immanent characteristics or "nature" of Islam; particularly in relation to fundamentalism, intolerance, terrorism and physical aggression. Discussions here focus on whether Islam is inherently and inevitably brutal and violent, which is well-illustrated in this post:

> *Fundamental is that islam is, whatever form it takes, a complete political system that through aggressive expansion has the goal to replace existing societal and political structure in the west with a form of government that is complete incompatible with western values.*

Similar to the categories above, there are some indications that we can observe a gradual increase of these topics starting from

around 2005, with increasing focus on Islamism and radical Islam, sharia laws and whether Islam is consistent with democracy and human rights.

6.4. Perception of women

The fourth topic category focuses on the *perception of women* among Muslims and within Islam, and comprises broader topics on the perception of women in general among Muslims, and specifically attitudes toward Swedish women. Many discussions here focus on various parts of the Quran, which is claimed to be hostile towards women. Connections are frequently drawn between Islam, oppression of women, pedophilia and child marriage, and to circumcision and genital mutilation. As one user puts this:

> It will never be possible… to integrate muslims. Take their perception of women for example. If a muslim marries a swedish woman, he can seem fine and willing to adapt to swedish rules in the beginning, but when he has more power over her, for example when they've had kids, he'll deep-rooted perception of women will come out and he'll turn completely.

Other central issues that are discussed in this context are homosexuality, sharia laws, honor killing and execution by stoning.[6] Overall, this category appears to remain relatively stable over time.

6.5. Terrorism

The topic category on *terrorism* focuses, naturally, to a large extent on the events after the 2001 attacks in the US. There are a number of different topics relating to this, including both conspiracy theories and later also the subsequent "war on terrorism" in Iraq and Afghanistan. Other specific terror attacks are represented as topics in this category, such as Bali in 2002, in the London Underground in 2005, the Stockholm bombing in 2010 and Bali, yet again, in 2013 due to new conflicts in the geographical area. As exemplified in the excerpt below, the boundaries between this category and the one above on the essence of Islam is sometimes rather fuzzy due to the broad discussions on whether Islam and the Quran "support terrorism or not" and if Islam is immanently leaning towards terrorism or not.

> Yet again terrorists have bombed, this time in India. In the north parts of Thailand it's also being bombed. In the north parts of China as well. Yes, wherever you look there are bomb attacks and other terror activity. The common denominator is Islamic fundamentalism. Why are almost only Muslims committing these acts today?

6.6. Sexual abuse and violence

The topic category *Sexual abuse and violence* particularly comprises topics relating to rape and sexual assault, immigrants and Muslims, and whether these groups are overrepresented in the statistics. A close reading reveals that in most of the discussions, Muslims are portrayed as violent and the focus often lies on the alleged reasons for this (e.g. due to genetic/biological causes or if it is "inherent in their culture/religion"). As illustrated in the excerpt below, Swedish women are generally particularly portrayed as victims of sexual assault.

> The integration policy is fucked up. But you can't blame everything on Sweden? Just because Abdullah, Ahmed and Mohammed are marginalized they don't have the right to gang rape swedish women. I'm so fed up with people blaming sexual assaults on defects of the system. Shouldn't at least a little personal responsibility be demanded from our "cultural enrichers"?!

As can be seen in Fig. 2, these topics are particularly central in the first years, but seem to decline from 2006 to 2008, and recur from about 2009 to 2013.

6.7. Race and ethnic groups and the Israel–Palestine conflict

Finally, the last two continuous topic categories relate to the *Israel–Palestine conflict*, which remains central in the discussions of all years, and to *race and ethnic groups*. We choose to discuss these categories together here, due to their similarity. The latter comprises a spectrum of topics on race, ethnic groups and alleged differences between these. There are also clearly distinct topics on Jews. Discussions generally focus on how/if race and ethnicity are connected to issues such as IQ, individual and social behavior, predispositions to violent behavior and various disabilities, which is clearly illustrated in the next post:

> Somalis and other aethiopides are another story. In part because somalis actually have a fair share of negroid blood in them, although no pureblooded caucasiod. Aethiopides are caucasiod morphologically (i.e. skeleton), and even some outside the skeleton (certain facial features for example, and in some cases also hair), but somalis are hardly any true caucasoids, so it's possible they'll spare somalis.

Other issues are also represented here, such as "race mixing" and which ethnic groups belong to which race. A close reading of the documents reveals that, similarly to several of the topics above and starting around 2006 and 2007, we may discern something of a shift from a focus on race and different ethnic groups in general, to more of a particular focus on Arabs and Muslims. We can also observe a peak in 2008 – a partial explanation for this is likely to be that the topic category on International conflicts declines in the same period, and there are indications that these categories share some rather ambiguous topics.

Besides the continuous topic categories, there are also a number of relevant episodically salient topic categories that focus on *mosques and minarets*, *cultural aspects of Islam* (e.g. food, eating meat, alcohol, halal), *extreme-right parties*, *freedom of expression and caricatures*, *international conflicts*, a rather diverse topic category on *media*, *the political establishment and net-hatred*, and finally a collection of minor topics (e.g. the ghetto/suburb riots, Anders Breivik etc.). While analyzing all these topics would require more space than available here, we choose to focus on analyzing the five that we deem as most interesting. For the same reason, we choose not to include any extracts from these topic categories, but instead focus more extensively on their development over time.

The topic category on *freedom of expression and caricatures* emerges around 2005 and is spurred primarily by the publications of a number of caricatures of the prophet Mohammed in the Danish broadsheet newspaper Jylland-Posten at the end of the same year. The events further escalated in 2006 when this was spread internationally, and in 2007 these discussions were connected to the publication of a number of pictures of Mohammed by the artist Lars Vilks in the newspaper Nerikes Allehanda. It decreases significantly in the period for 2008 to 2009, only to emerge yet again in 2010 – this time propelled by Vilks publication of a number of new and much-debated paintings of Mohammed as a "roundabout dog".

---

[6] As a minor note, we can also see a topic in 2009 relating to "Buried Alive" ("Gömda"), a novel by Liza Marklund that is claimed to be based on a true story, and deals with a woman who is abused by her immigrant boyfriend. This book received much public attention and generated heated debate on the subject.

The topic category that focuses on *International conflicts* collects a number of different topics on various international conflicts, ranging from the Iraq war to the Arab spring. It emerges around 2005, and focuses particularly on the Iraq war and conflicts between the US and middle-eastern countries and peaks again around 2009 – this time primarily sparked by the presidential elections in Iran and the following protests often referred to as the Green revolution. There are also discussions on the situation of the Kurds and the Kurdish protests in Turkey at the end of the same year. As can be seen in Fig. 3, this topic category again increases dramatically from 2011 to 2013, and focuses extensively on the events relating to the Arab spring, the events in Egypt that began in late 2010, the Muslim brotherhood, Mubarak, etc. In the same year, we can see another topic in the same category about Libya and Syria, the rebels and Gaddafi. The topic category about Syria remains in 2012 and 2013, and focuses on the conflicts between the rebels and the regime (and includes Lebanon by the end of 2013).

The topic category that focuses on *extreme-right parties* seems to generally increase over time, and clearly peaks in the time of the Swedish government election in 2010. Most discussions are about the extreme-right party the Sweden Democrats, but after the election the discussions are increasingly shifted towards the Swedish Party (SvP), a small neo-Nazi party.

Let us now take a look at the topic category of *mosques and minarets* (Fig. 4). This topic category peaks in 2003, driven by discussions particularly on mosques and an arson attack on a Mosque in Malmö the same year.
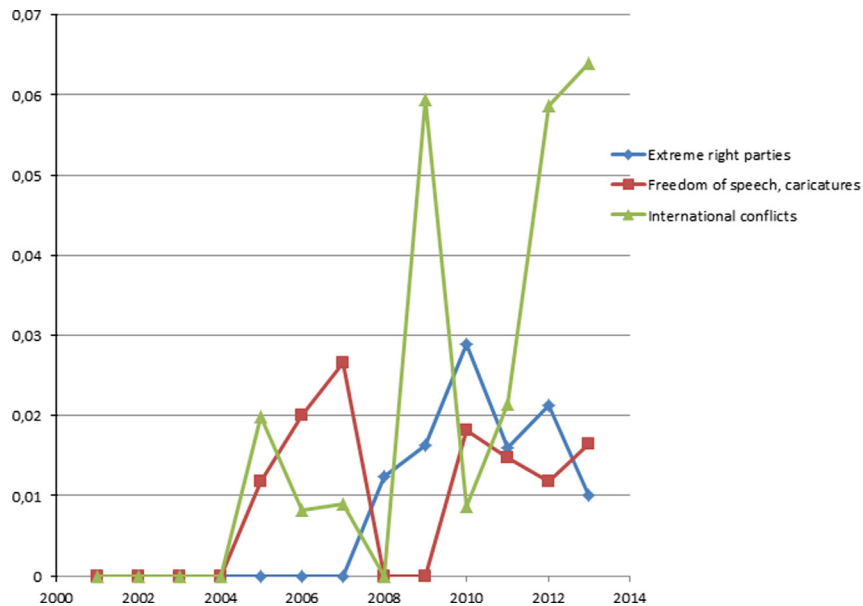

Fig. 3. This graph plots the development over time for the episodically salient topic categories Extreme-right parties, Freedom of speech and International conflicts. The graph shows the proportional size of the topic categories in relation to the other topics in the same year and is calculated by adding the strength value of all topics within each topic category/year.

Fig. 4. This graph plots the development over time for the episodically salient topic categories Mosques and minarets, and Cultural aspects. The graph shows the proportional size of the topic categories in relation to the other topics in the same year and is calculated by adding the strength value of all topics within each topic category/year.
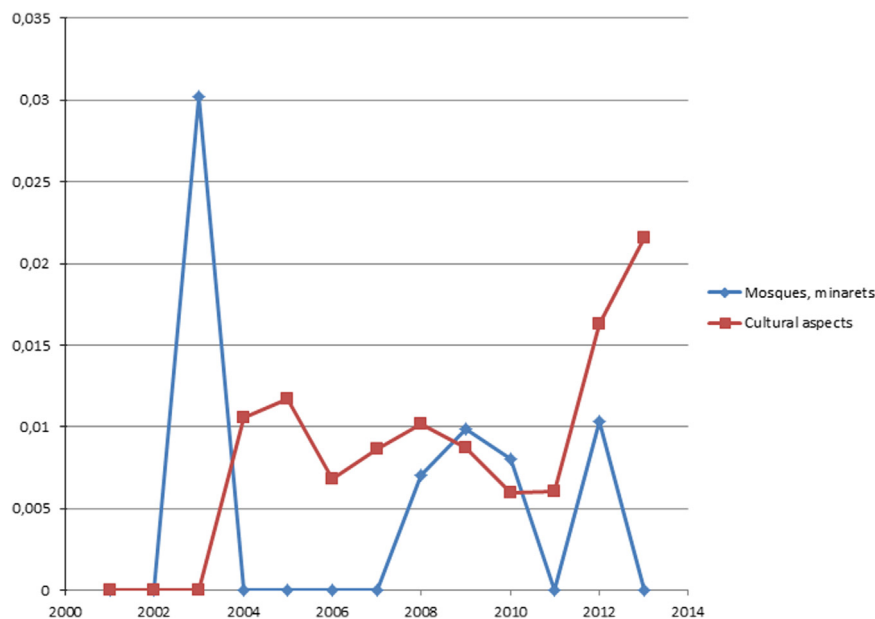
The discussions increase again in 2008, and focus on critical discussions on mosques in general, often described as an "infiltration of Islam", with a particular focus on Malmö and Rosengård. There is also a specific topic about minarets the following year (2009), mainly driven by the focus on these issues in an election in Switzerland the same year. In 2012 discussions increasingly focus on adhan/call of muezzin, which was actualized by the Islamic society/association in the county of Botkyrka that applied for permission to have adhan once a week in the mosque in the suburb Fittja – this caused heated discussions in the forum.

Finally, the rather broad topic category on *Cultural aspects* of Islam emerges around 2004 to 2005 and focuses primarily on pork and Halal slaughter, which is generally portrayed as inhumane and problematic. Later topics focus on Ramadan, dogs and what are considered to be "clean" and "dirty" animals within Islam, but also on personal hygiene and humor. In 2012, there is a topic that focuses particularly on halal food in schools that was triggered partly by Jönköping municipality, which was reported for special treatment of Muslim students (getting time off during Ramadan, not having to stay overnight on school trips, being allowed to wear veil at sport activities, etc.).

7. General patterns

By taking a broad perspective on the patterns that appear in the corpus we can observe some general trends in the representation of Muslims and Islam.

Firstly, we overall distinguish a general discursive shift from a focus on immigrants in general, to Muslims in particular. In this sense, Muslims are increasingly defined as a distinct – and as we will see also a homogeneous – social group *within* immigrants. As we have seen above, this tendency appears particularly clear in the broader topic category that focuses on multiculturalism and clashes with Swedish traditions, but also in discussions on race and ethnic groups, and regarding sexual abuse and violence. In the latter, we seem to discern a certain shift from "immigrants" being portrayed as the perpetrators, to an increasing focus on particularly Muslims and Arabs. However, we should note that this tendency appears despite the fact that we specifically focus on the corpus of posts that include keywords about Muslims and Islam and not immigrants in general. A more thorough analysis of this would require a larger, more inclusive corpus consisting of posts that do not only contain keywords relating to Muslims and Islam.

Furthermore, Muslims are clearly represented as a group embroiled in conflict. This can be seen both explicitly and in more subtle, latent ways that pervade throughout most, or even all, topics. Perhaps the most obvious is the dominance of topics that focus on terrorism, sexual abuse and violence, and the perception of women in Islam. While there are certain indications that at least the topic categories on terrorism and sexual abuse and violence might decline somewhat over time after the September 11 attacks in 2001, the more general discussions on the perception of women within Islam seems to be relatively stable and remains highly central over time. These discussions generally focus on veils and burqas, the alleged negative perception of women among Muslims in general, and in particular attitudes towards Swedish women. A very central discussion in this topic category, which also seems to increase over time, is how women are portrayed in the Quran, and whether gender oppression is an inherent trait of Islam. This also shows that the representations of Muslims and Islam are gendered. Muslim men are generally portrayed as aggressors and extremists, while Muslim women are often portrayed as the victims of oppression (see e.g. the veil-related topics), and Swedish women as the victims of sexual assault.

This conflict dimension is also apparent in the large topic category about immigration and multiculturalism. Discussions within this topic category clearly reveal how Muslims are generally represented as a homogeneous group, rather than a highly diverse population pertaining equally in diverse practices and beliefs that vary from culture to culture.[7] More specifically, these discussions describe a claimed internal conflict or incompatibility between Islam and what is depicted as Swedish culture and values, where cultural/religious festivals such as Christmas and midsummer are described as being challenged or threatened by foreign culture and customs such as the Ramadan. This claimed incompatibility with Swedish culture is also manifested in some of the more specific temporarily salient topic categories that focus on mosques and minarets, and freedom of speech, where the latter is claimed to be in danger because of Islam; for example, as with the case of the Mohammed caricatures.

This element of conflict and claimed inherent violent nature of Islam seems generally to be portrayed as emanating from the essence of the religion itself. This is apparent in that the topic model, rather interestingly, distinguishes a number of topics that we have chosen to classify under the category *The essence of Islam*. As we have seen above, this category focuses on the nature or essence of Islam and whether violence, fundamentalism, sexism, and incompatibility with democracy are implicit in the religion itself. While not statistically validated, we seem to discern a certain increase of this topic category over time (see graph 2).

Ultimately, this essentialist depiction of Muslims and Islam that permeates through most predominant topics and over time contributes to constructing them as the "other": an immature and even backward ethnic group that exists in tension – or even incompatibility – with what is depicted as Swedish/Western culture. This tendency pervades through most of the topics and is also blatant in the temporarily salient topics, where foreign cultural aspects in Islam are contrasted against what is considered Swedish values. This is decidedly apparent in the discussions in the topic category of *cultural aspects*, which focuses on halal meat, alcohol, religious festivals and food.

8. Discussion

As we have argued, the topics analyzed here capture the discursive landscape of Islam and Muslims, and serve to (re)contextualize and thus infuse these words with meaning. We have seen that while this landscape is indeed relatively broad, and consists of discussions that focus on various issues, most topics are nonetheless conflict-oriented and characterized by a distinction of Muslims as the other: an outgroup that is in conflict with Swedish culture and values. By being (re)contextualized in these ways, Islam is thus represented as an inherently violent religion, and Muslims are frequently represented as fundamentalist, terrorist, sexist, militant, undemocratic, violent and fanatical. As we have seen, these results are in-line with established research on the representation and discursive constructions of Muslims and Islam in mainstream Western media. The results also fit well with existing research on online Islamophobia, although this field is far less investigated. While it is hard to draw any definite conclusions here, due to the lack of research on the discursive representations of Muslims and Islam in Swedish media in particular, a recent report from the United Nations Human Right Committee has

---

[7] An interesting exception of this tendency towards homogenization of Muslims is the topic category Islam as a religion, which focuses on different ethnic Muslim groups and religious orientations within Islam. However, a closer analysis of documents within these topics reveals that most of these discussions seem to stem particularly from the religious sub-forum in Flashback.

directed firm criticism against Sweden for what they call the "chronic negative portrayal of Muslims in the media" (CCPR, 2016: 4). This indicates that discursive patterns above are also present in Swedish media.

We can also note that, compared with mainstream media, there is overall a more aggressive and conflict-oriented language in the forum and derogatory and racist terms such as "raghead" and "sand nigger" are commonly used. The main reasons for this are likely to be user-anonymity, lack of press-ethical considerations and the absence of moderation.

Nonetheless, these strikingly similar patterns indicate that, at least in this particular case, social media does not serve simply as an alternative arena or a free space for counter-discourses that challenge the dominating discourses of traditional media. Rather, in this case, the discourses in both these media types seem to reflect each other. This indicates that instead of existing in a strained or competing relationship, traditional and social media co-exist in a symbiotic relationship: they are knitted together in a symbolic web, where the discourses and representations of Muslims and Islam may reinforce each other. The existence of such a reciprocal/consensual relationship is further sustained by research that shows that 50% of journalists globally consider social media as the "main source of information" (ING, 2014), and that participants in social forums and blogs often consume, spread and discuss news that is consistent with their own views.

This relationship between traditional and social media thus results in a feedback loop, where a message that originated as news can be caught up in social media and "go viral", thus intensifying the coverage dramatically (or vice versa). In this sense, when the representations of Muslims and Islam in both social and traditional media correspond to each other, this likely results in an even stronger polarizing effect on public discourses. Rather than constituting free spaces, social media thus become online amplifiers that serve to reinforce representations in traditional media. However, this still constitutes rather uninvestigated territory and there is clearly a great need of more scholarly work in order to shed light on how traditional and social media work together, and how such feedback processes (re)produce discourses and shape public opinion.

A necessary and essential step here is to develop new methodologies and technical solutions that enable researchers to approach the vast and growing archive of social texts that constitute social media. The approach exemplified in this article of combining discourse analysis and powerful data mining techniques, such as topic modeling, demonstrates one such particularly promising avenue. More steps in this direction are crucial and highly relevant, not only to allow for a critical investigation of the construction of discursive power in society, but also since social media offers a unique entrance into the everyday discourses and the otherwise often impenetrable world of kitchen-table discussions: a world that has previously been beyond the reach of both traditional media and social analysis.

Appendix. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dcm.2016.04.003.

References

Allen, Chris, Isakjee, Arshad, Ögtem, Özlem, 2013. Maybe we are hated. The experience and impact of anti-Muslim hate on British Muslim Women. Institute of Applied Social Studies, University of Birmingham.

Awan, Imran, 2016a. Islamophobia in Cyberspace: Hate Crimes Go Viral. Routledge.

Awan, Imran, 2016b. Virtual islamophobia: the eight faces of anti-muslim trolls on Twitter. In: Awan Imran (Ed.), Islamophobia in Cyberspace: Hate Crimes Go Viral. Roughledge, New York.

Bail, Christopher A., 2014. The cultural environment: measuring culture with big data. Theory Soc. 43 (3–4), 465–482.

Baker, Paul, 2006. Using Corpora in Discourse Analysis. A&C Black.

Baker, Paul, Gabrielatos, Costas, Khosravinik, Majid, et al., 2008. A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. Discourse Soc. 19 (3), 273–306.

Baker, Paul, Gabrielatos, Costas, McEnery, Tony, 2013a. Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press. Cambridge University Press.

Baker, Paul, Gabrielatos, Costas, McEnery, Tony, 2013b. Sketching Muslims: a corpus driven analysis of representations around the word 'Muslim'in the British press 1998–2009. Appl. Linguist. 34 (3), 255–278.

Baker, Paul, Levo, Erez, 2015. Picking the right cherries? A comparison of corpus-based and qualitative analyses of news articles about masculinity. Discourse Commun. 1750481314568542

Blei, David M., 2012. Probabilistic topic models. Commun. ACM 55 (4), 77–84.

Blei, David M., Ng, Andrew Y., Jordan, Michael I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022.

Brunsting, Suzanne, Postmes, Tom, 2002. Social movement participation in the digital age predicting offline and online collective action. Small Ggroup Res. 33 (5), 525–554.

Cammaerts, Bart, 2012. Protest logics and the mediation opportunity structure. Eur. J. Commun. 27 (2), 117–134.

CCPR, 2016. CCPR – International Covenant on Civil and Political Rights.

Cheng, Winnie, 2014. Corpus-based linguistic approaches to critical discourse analysis. In: Chapelle, C.A. (Ed.), The Encyclopedia of Applied Linguistics.. Oxford, England

Copsey, Nigel, Dack, Janet, Littler, Mark, et al., 2013. Anti-Muslim Hate Crime and the Far Right. Teeside University: Centre for Fascist, Anti-Fascist and Post Fascist Studies, Teeside.

Diani, Mario, 2000. Social movement networks virtual and real. Inf. Commun. Soc. 3 (3), 386–401.

Feldman, Matthew, Littler, Mark, 2014. Tell Mama Reporting 2013/14 Anti-Muslim Overview, Analysis and 'Cumulative Extremism'. Teeside University Centre for Fascist, Anti-Fascist and Post-Fascist Studies.

Fowler, Roger, 1996. On critical linguistics. In: Caldas-Coulthard, Carmen Rosa, Coulthard Malcolm (Eds.), Texts and Practices: Readings in Critical Discourse Analysis. Psychology Press, London, England, p. 1.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2014. Bayesian data analysis, vol. 2. Boca Raton, FL, USA, Chapman & Hall/CRC.

Hafez, Kai, 2000. Islam and the West in the Mass Media: Fragmented Images in a Globalizing World. Hampton Pr.

ING, 2014. Impact of Social Media on News. ING Group ⟨https://www.ing.nl⟩.

Karpf, David, 2012. Social science research methods in Internet time. Inf. Commun. Soc. 15 (5), 639–661.

Klinger, Ulrike, Svensson, Jakob, 2014. The emergence of network media logic in political communication: a theoretical approach. New Media Soc. 1461444814522952

Koopmans, Ruud, Olzak, Susan, 2004. Discursive opportunities and the evolution of right-wing violence in Germany. Am. J. Sociol. 110 (1), 198–230.

Koster, W.de., Houtman, Dick, 2008. Stormfront is like a second home to me. On virtual community formation by right-wing extremists. Inf. Commun. Soc. 11 (8), 1155–1176.

Lilleker, Darren G., Koc-Michalska, Karolina, Schweitzer, Eva Johanna, et al., 2011. Informing, engaging, mobilizing or interacting: searching for a European model of web campaigning. Eur. J. Commun. 26 (3), 195–213.

Manan, S.A., 2008. Through Western eyes: Covering Islam after September 11. Critical Perspectives on Language and Discourse in the New World Order, pp. 111–133.

Marko, Georg, 2008. Penetrating Language: A Critical Discourse Analysis of Pornography. Narr.

Matters, Faith, 2015. Facebook report: Rotherham, hate and the far-right online.

Moore, Kerry, Paul, Mason, Wren, Lewis Justin Matthew, 2008. Images of Islam in the UK: The representation of British Muslims in the national print news media 2000–2008.

Moscovici, Serge, Duveen, Gerard, 2000. Social representations: Explorations in social psychology. Polity Press Cambridge.

Oboler Andre. (2013) Islamophobia on the Internet: The Growth of Online Hate Targeting Muslims. In: Institute Online Hate Prevention (ed).

Oboler, Andre, 2016. The normalisation of islamophobia through social media: Facebook. In: Awan Imran (Ed.), Islamophobia in Cyberspace: Hate Crimes Go Viral. Routledge, New York.

Orpin, Debbie, 2005. Corpus linguistics and critical discourse analysis: examining the ideology of sleaze. Int. Journal. Corpus Linguist. 10 (1), 37–61.

Partington, Alan, 2006. The Linguistics of Laughter: A Corpus-Assisted Study Of Laughter-Talk. Routledge.

Pollach, Irene, 2012. Taming textual data: the contribution of corpus linguistics to computer-aided text analysis. Organ. Res. Methods 15 (2), 263–287.

Poole Elizabeth. (2002) Reporting Islam: Media Representations of British Muslims: IB Tauris.

Richardson, John E., 2004. (Mis) Representing Islam: The Racism and Rhetoric of British Broadsheet Newspapers. John Benjamins Publishing.

Sayre, Ben, Bode, Leticia, Shah, Dhavan, et al., 2010. Agenda setting in a digital age: tracking attention to California proposition 8 in social media, online news and conventional news. Policy Internet 2 (2), 7–32.

Stubbs, Michael, 1994. Grammar, text, and ideology: computer-assisted methods in the linguistics of representation. Appl. Linguist. 15 (2), 201–223.

Stubbs, Michael, 1997. Whorf's children: critical comments on critical discourse analysis (CDA). Br. Stud. Appl. Linguist. 12, 100–116.

Teubert, Wolfgang, Krishnamurthy, Ramesh, 2007. Corpus Linguistics: Critical Concepts in Linguistics. Routledge.

Tognini-Bonelli, Elena, 2001. Corpus Linguistics at Work. John Benjamins Publishing.

Törnberg, A., Törnberg, P., 2016. Combining CDA and topic modeling: Analyzing discursive connections between Islamophobia and anti-feminism on an online forum. Discourse & Society 0957926516634546.

Van Dijck, José, Poell, Thomas, 2013. Understanding social media logic. Media Commun. 1 (1), 2–14.

Van Dijk Teun, A., 1993a. Elite Discourse and Racism. Sage Publications.

Van Dijk, Teun A., 1993b. Principles of critical discourse analysis. Discourse Soc. 4 (2), 249–283.

Van Dijk, Teun A., 2003. 18 Critical discourse analysis. Handb. discourse Anal. 18, 352.

Van Dijk, Teun A., 2009. Society and Discourse: How Social Contexts Influence Text and Talk. Cambridge University Press.

Widdowson, Henry G., 2000. On the limitations of linguistics applied. Appl. Linguist. 21 (1), 3–25.

Widdowson, Henry George, 1996. Linguistics. Oxford University Press.

Wodak, Ruth, Meyer, Michael, 2009. Methods for Critical Discourse Analysis. Sage.