Jonas Svensson

The 'Radical Muslim' in the Swedish News

Using Computer-Assisted Methods to Map a Discourse

The aim of this article is to illustrate how highly rudimentary computer-assisted text analysis can aid the preliminary mapping of public discourse. The case that has been chosen is the use of a particular combination of words, 'radical' and 'Muslim', in a corpus of articles published in the Swedish printed news media which have been accessed through the database *Mediearkivet*. Texts in the corpus were accessed and analysed using custom-made code in the programming language Python. The article shows how even limited skills in computer-assisted methods can aid scholars in discovering patterns and trends in a large collection of texts which is difficult, if not impossible, to process manually.

This article is based on a paper which was presented at an academic workshop at which participants were asked to prepare an answer to the question: 'How do we define a radical Muslim?'. I understood, of course, that the task was to participate in sharp, intellectual conceptual analysis with some aspirations to academic normativity – that is to ask the question: 'How should we define a radical Muslim?'. Nevertheless, I decided to approach the task literally, and in so doing to expand the community of 'we' beyond the assembled researchers in the field of Islamic studies.

Jonas Svensson is a professor in the study of religion specialized within Islamic studies at the Linnaeus University. Early in his career, he has primarily been interested in gender and Islam, modernist Islam interpretations, Islamic education in East Africa, and coultural historical perspectives on the Quran and Mohammad. Lately, he has primarily been occupied with theory and method in the study of Islam and more specifically research on cognitition with the study of relgion and its consequences for the study of Islam; in addition to this is datadriven methods that can enrich the latter

Hence, my contribution to the workshop took as its starting point not the normative question of how we ought to use (or not to use) 'radical Muslim' in our analysis of contemporary Islam, but instead the empirically oriented question of how the term is actually used in the public discourse in which academics partake. This was not merely an attempt to be witty. There is a serious point at play: while academic conceptual analysis and attempts to provide stable working definitions of concepts are important, they are in some cases – when the terminology in question is already in use outside the world of academia – difficult. They become even more difficult if the use of these concepts in public discourse has not been systematically investigated. But how can one carry out such an investigation in a time-efficient manner?

It is this latter, methodological question that I focus on in the following. As scholars, we find ourselves in an era in which new technologies for producing and disseminating text, images and sound have made the amount of data available for the analysis of public discourse overwhelming. The researcher is faced with challenges on all levels of enquiry: collecting, sampling and analysing. Fortunately, new technologies also offer new tools that can be employed on all these levels, tools that may previously have required expert knowledge, but are now becoming increasingly accessible.

In a sense, the following falls within the larger field of 'digital humanities' (see e.g. Berry & Fagerjord 2017; Muhanna 2016; Hansson & Svensson 2020), or more precisely, the aspect of digital humanities that concerns how new technologies can assist with performing what is essentially traditional humanities research. Some technologies offer radically novel ways of carrying out humanities research and allow for the execution of research which, without digital tools, was impossible to perform. However, there are also less revolutionary ways in which computers can aid the humanities scholar: by performing fairly simple and easy-to-grasp tasks which, while *possible* to perform manually, are highly time-consuming, tedious, or both. The following is an example of the latter, using the question asked at the workshop: 'how do we define a radical Muslim?' as a case study.

The Task and the Method

The analysis of a segment of Swedish public discourse below uses rudimentary custom-made code in the programming language Python. While Python is far from the only language available, it is, due to its simple syntax and extensive resources in terms of community-created code, which can be reused in the creation of new code, particularly apt for what the computer scientist Nick Montfort has termed 'exploratory programming' in the humanities (Montfort 2016).

The data consisted of a corpus of texts which was constructed using the Swedish database *Mediearkivet*, a searchable database of material produced by a large selection of Swedish media outlets (over 3,500). The database allows for the gathering of rudimentary statistics on media coverage of particular words or collocations, but – more importantly for what follows below – it also provides downloadable metadata for all posts and, in many cases, the full texts of articles.

In creating this corpus, I limited myself to material from printed daily media, in the form of 80 established newspapers. The limitation here is the assumption that that the character of this material is somewhat uniform. It should be remembered, however, that the main focus is not on the corpus and its content as such, but on the computer-assisted methods which were employed. The selection was *not* based on the assumption that this particular subset of public discourse was in any way representative of the whole discourse. On the contrary, a more extensive mapping of how 'radical Muslim' is used in Swedish public discourse would have to consider other sources besides daily newspapers and beyond those available in Mediearkivet.

In addition, the focus of this investigation was on the use of a particular word combination (or, more accurately, a set of word combinations) in public discourse, not on the larger issue of media representations of Islam and Muslims. The latter is an established field of scholarly inquiry which often takes the form of critical discourse analysis. It has a fairly long history (for classical works see Said 1981; Poole 2002; Baker, Gabrielatos & McEnery 2013), and research in the field is constantly being produced (for recent additions see e.g. Hervik 2018; Pennington 2018, and for some critical notes, see Faimau 2015). Given its narrow scope, this article is of marginal, if any, relevance to the

subject matter of this particular area of research, although the question of how the term 'radical Muslim' is employed in Swedish printed news media has not (to my knowledge) been addressed previously.

Constructing the corpus

Within the subset of the 80 Swedish daily newspapers, I searched for the word stem and 'wildcard' combination 'radikal*' and 'muslim*'. This identified articles which contained any combination of these two words stems in the given order and adjacent to each another. The articles identified contained word combinations such as 'radikala muslimer' (radical Muslims), 'radikal muslimsk' (radical Muslim [adjective form]), 'radikaliserade muslimer' (radicalised Muslims). I will return to the issue of different word combinations below. In total, the search returned references to 1,317 articles published in the selected media outlets between 1991 and 2019.

At this early stage, some questions arose. Is 1,317 articles over a period of 29 years a lot? Is the use of the expression 'radical Muslim' particularly prominent in public discourse? A comparison with another word combination, radical* and islam*, was enlightening. While the combination of radikal* and muslim* returned a total of 4,437 hits in all media outlets (i.e. not only newspaper articles) from 1945 to 2019, the combination of radikal* and islam* returned 28,853 hits. It would appear that the particular word combination of radikal* and muslim* was not very common. On the other hand, compared to corresponding word combinations of radikal* and words referring to adherents of other religious traditions, the dominance of 'muslim*' was apparent. 'Radical*' and 'kristen/kristna'[Christian] returned 353 hits, 'hindu*' 320 and 'buddhist*' 142.

So while it may not be overtly commonplace, there is some justification for paying closer attention to the expression 'radical Muslim'. Data for the 1,317 posts in the search results was downloaded as a text file using *Mediearkivet's* export function. This file was parsed into separate posts, one for each article, using a customised Python script. The text file contained metadata for all the articles such as their date of publication and the newspaper in which they were published. For a large proportion of the articles, it also contained part of or the whole body of text.

Posts for each article containing the metadata and (when present) article texts were stored as separate items in a Python dictionary, each with keys and corresponding values to make the different kinds of data easily accessible. This 'database' of article posts was then further scrutinised. The set of 1,317 posts downloaded from *Mediearkivet* turned out to contain a fair number of duplicates. Removing these reduced the total set to 924 articles. An additional 100 articles were removed due to complete overlap of text content with another article in the set. Furthermore, it is common practice in news media to publish information from news agencies which has been rewritten to a greater or lesser extent. This means more or less the same article may reoccur in several newspapers, perhaps with slight variations. In order to exclude such variant versions of what is in practice the same text, I constructed a Python script that screened for versions like this. For this I used a simple measure of text similarity called the Jaccard index or the Jaccard coefficient (Vijaymeena & Kavitha 2016, 22) which compares texts on the lexical level with respect to unique words ('types'). If two texts contain the exact same set of unique words, they receive a JC of 1.0. If two texts have no overlap of unique words, their JC is 0.0. I set the threshold value to 0.5 and removed one of each set of two articles that displayed a JC above 0.5. This reduced the total number of articles further to 692. An additional 40 articles were excluded because, although returned by the search in Mediearkivet as containing a combination of the words 'radikal*' and 'muslim*', the downloaded texts did not contain that combination. This is due to the fact that not all texts whose content is searchable are accessible from Mediearkivet as full texts.

Ultimately, then, the following is based on further exploration of the texts of the remaining 652 articles, published in 59 newspapers between June 9, 1991 and November 12, 2019.

General Overview

The newspapers included both national and local press. Of the 59, only 12 had ten or more articles featuring a combination of radikal* and muslim*. Together, these 12 contained 79% of the total number of such articles. It should be noted that the set of articles which are analysed further below does not contain all
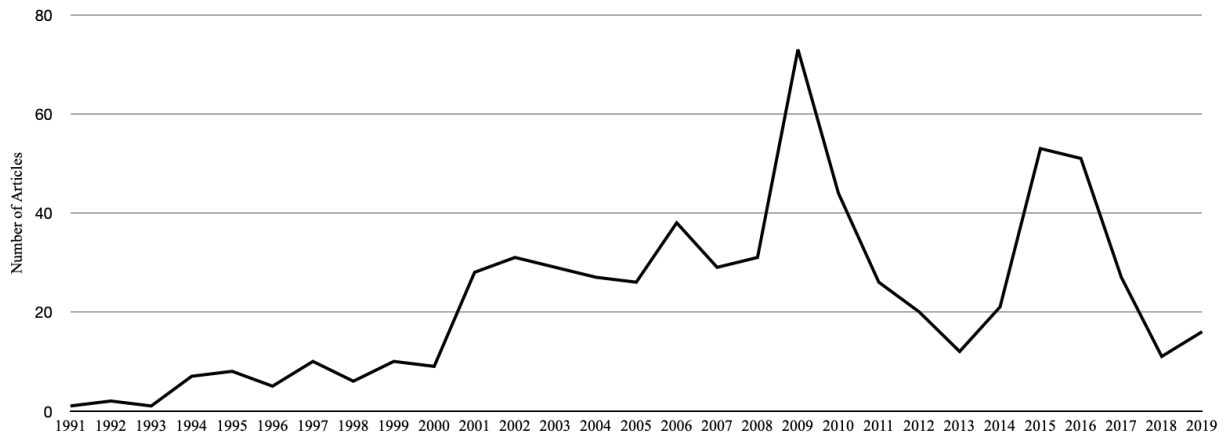
Figure 1. Frequency over time of articles containing the word combination 'radikal*' and 'muslim*'.

the articles containing radikal* and muslim* which had been published in the 80 newspapers originally included in the search results from Mediearkivet. The set which was analysed further comprises articles for which a body text containing radikal* and muslim* could be downloaded. Duplicates were also removed. In cases where two or more articles displayed a measure of similarity above the 0.5 threshold, only one version was kept.

The number of articles which contain the word combination being studied in the remaining set varies over time, as can be seen in Figure 2. There are significant increases and peaks in 2001, 2006, 2009, and lastly in 2015/2016.

Three of these four peaks are perhaps to be expected. The peak (or rather increase) in 2001 is attributable to the attack on the World Trade Centre. 22 of the total of 28 articles from that year were published after September 11, 2001. Based on the graph, it is possible to claim that this date is when the word combination of 'radikal*' and 'muslim*' made a significant entry into the Swedish newspapers, although there is a noticeable increase from 1994. A closer look at the article texts from 2006, 2009 and 2015/16 provides clues about the other peaks. Here, computer-assisted text analysis comes in handy. A frequency analysis of words used in the articles published in 2006 shows that the words 'Danmark' and 'dansk' ('Denmark' and 'Danish') together occupy the second position (after 'Muslims' and 'Muslim'), and that 'Jyllandsposten' (with variants) takes 11th place among the most common words. The peak in 2006 can thus be connected to a particular event of both regional and international significance: the publication of a set of cartoons in the Danish daily newspaper *Jyllandsposten*, some of which depicted the Prophet Muhammad, and the subsequent solidarity publica-

tions in some Swedish papers in late 2005 (Klausen 2009). In 2015 and 2016, the most common word is 'IS', with 'Syria' coming in fourth place. This indicates that during this period, the focus of articles containing 'radikal*' and 'muslim*' was on events in Syria and Iraq and the activities of the Islamic State.

But what about the even more significant peak in 2009? Here, too, a word frequency search provides some clues. Disregarding words containing the stems 'radikal*' and 'muslim*', the fourth most common word in the articles from 2009 is 'Rosengård', the name of a suburb of the southern city of Malmö. 16 of the 73 articles from that year mention 'Rosengård', and all but one were published during the first two months of 2009. On January 28, 2009, the Swedish Defense University (Försvarshögskolan) published a report on increasing 'radicalisation' of Muslims in Rosengård based on interviews with social workers, teachers and police who were active in the area (Ranstorp & Dos Santos 2009). The report was criticised by other researchers on methodological grounds, as well as on the grounds that the word 'radicalisation' was not defined. The critique, and responses to it, were highlighted in media.

Another event in early 2009 may have contributed to the increase in frequency. On January 9, the poet Mohamed Omar declared himself to be a 'radical Muslim' in an article in the tabloid *Expressen* and voiced support for the Palestinian Islamist movement Hamas (Omar 2009). Comments on this statement are a case of 'radical Muslim' being used, not as a descriptive journalistic term, but as someone's self-designation. The name Muhammad Omar appears in 15 articles from 2009.

What can be concluded from this very vague preliminary frequency analysis? Firstly, the word stem combination being analysed has existed in the media for over 30 years, and is hence not a recent phenomenon. However, there is an apparent increase in its use after September 11, 2001, as well as peaks associated with (1) events involving religiously framed violence, (2) public discussions about the perceived increase of certain forms of religious activism in Sweden (framed as a social problem), and (3) a single case of self-designation as a 'radical Muslim'. There is also an indication of a decrease in use in 2018 and 2019 compared to previous years after 2001.
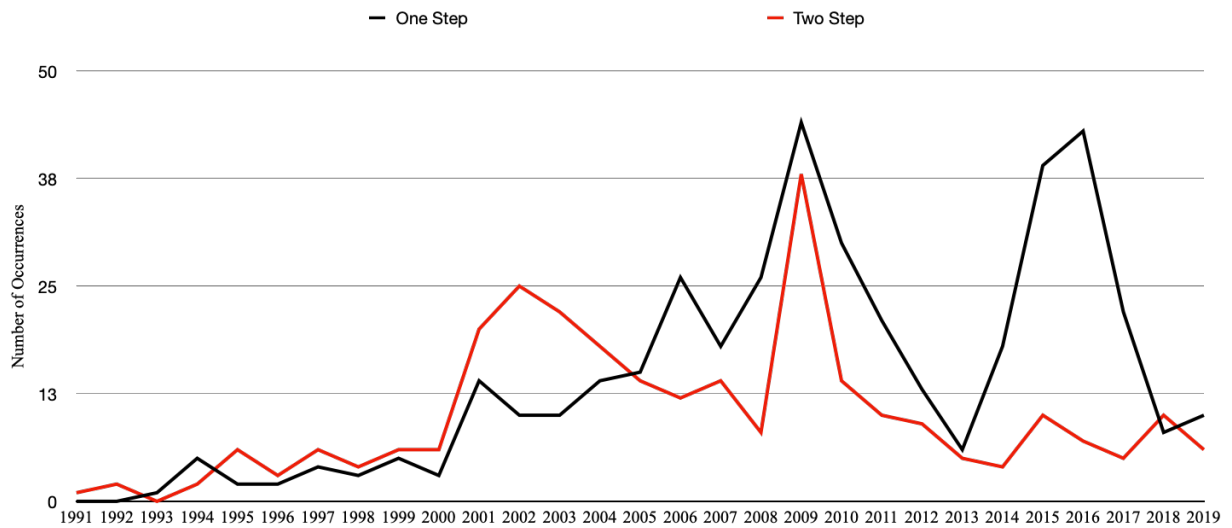
Figure 2. 'One-Step' and 'Two-Step' adjective use in word combinations over time.

Word Combinations

The texts included in the final analysis all contained the word stem combination 'radikal' and 'muslim', but the actual word combinations varied. Using the Python library TextBlob and its function for extracting so-called n-grams, i.e. word combinations where 'n' denotes the number of words adjacent to one another, I collected all two-word combinations from all the articles and extracted from these all those in which the first word contained 'radikal' and the second 'muslim'. Excluding word combinations that occurred only once in the material, and not distinguishing between singular and plural forms, all occurrences could be sorted into one of the three following categories:

1) Cases where words were built on radikal*, which served as modifier of muslim*, which in turn modified a noun – as in the phrase 'radikal muslimsk grupp' (radical Muslim group) or 'radikal muslimsk rörelse' (radical Muslim movement). There were a total of 284 such combinations in the material.
2) Cases where the muslim* part of the combination was the noun and radikal* was the specifying adjective, for example, 'en radikal muslim' (a radical Muslim) or 'radikaliserade muslimer' (radicalised Muslims). There were 412 such cases.
3) Lastly, there was a small number of cases where radikal* formed part of a verb, with muslim* as an object, for example 'radikalisera muslimer' (to radicalise Muslims).
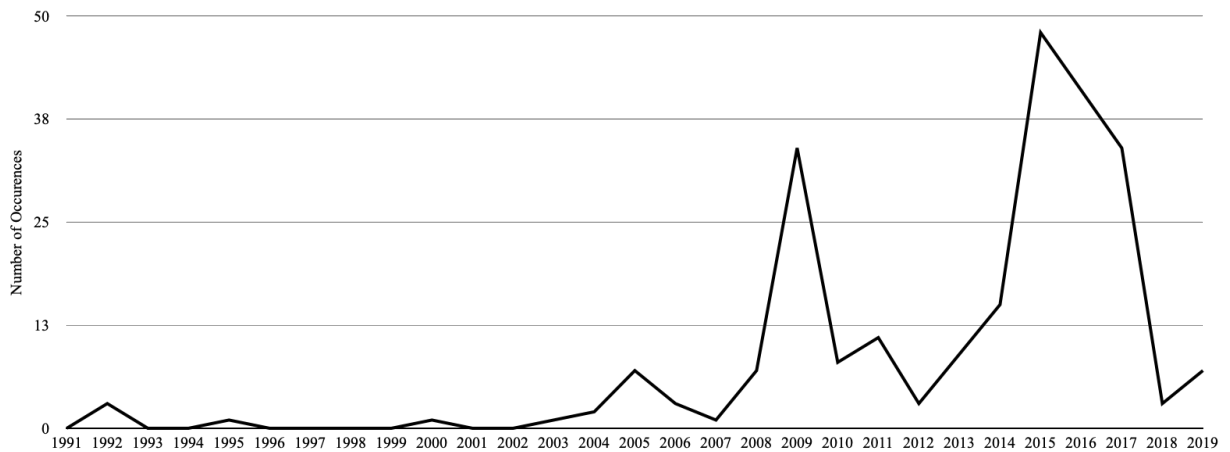
Figure 3. Frequency of words with the word stem 'radikaliser*' over time.

Only 13 cases like this with more than one occurrence were found in the material.

There was thus a dominance of the use of 'radikal*' as an adjective modifying muslim* which will be given closer attention below. Focusing purely on the two cases of 'two-step' (i.e. 1 above) and 'one-step' (i.e. 2 above) adjectival use, Figure 2 plots change in frequency over time.

With a few exceptions (1993, 1994), there appears to be a tendency of dominance of two-step over one-step adjectival use before 2005. The reverse is true after 2005 (except in 2018). Making reservations for the fact that the overall number of instances is low, there is some room for speculative probing for a pattern.

The two-step adjectival use relates to nouns. But which nouns? Of a total of 279 occurrences, 88, or 32%, related to 'radical Muslim *group/s*'. There was a tendency for the word combination to specify *collective* entities. Other nouns with relatively high frequencies were 'movement/s' (19 instances), 'organisation/s' (15), and 'party/parties' (7). Cases referring to individuals, as opposed to collectives, were rarer. The two main examples were 'leader/s' (16) and 'preacher/s' (9). It can be concluded that the two-step adjectival use of the word combination radikal* and muslim* refers, to a great extent, to the *programme, ideology* or *activities* of a group.

The apparent shift in use from two-step to one-step adjectival use of radikal* and muslim* may also indicate a shift in the conceptualisation of the word 'radical'. Over time, 'radical' has become more of something that Muslims *are,* rather than a spec-

ification of contexts (organisations, groups, parties) in which Muslims are *found*. 'Radical Muslim' sticks out as a particular sub-category of the larger category of 'Muslim' (implicitly different, then, from the sub-category of non-radical Muslims). Further data supports the notion of such a shift. Becoming a 'radical Muslim' implies a process of change. Hence, it is not just that 'radical' becomes more of a character trait of certain Muslims after 2005, but there is also an increase in references to a *process* of Muslims becoming 'radical' – i.e. 'radicalisation' – in the set of articles. A search for the word stem 'radikaliser*', which covers words such as 'radikalisera' (to radicalise), 'radikalisering' (radicalisation) and 'radikaliserad' (radicalised), produced 236 hits. Of these, 230 (97%) were in articles published after 2005. Figure 3 shows the distribution over time.

The highly speculative claim concerning a possible pattern in the use of the words radical* and muslim* over time made here should not overshadow the main point of this section. The analysis of the texts that underlie the claim is one that *could* be carried out without the help of computers: extracting bigrams from the texts that contain radikal* and muslim*, organising these bigrams into groups, and then counting change in frequency over time. However, the total number of words in all the articles is 412,304. Assuming an overall reading pace of 300 words per minute, just reading through all the texts in the collection (which would be a necessary first step) would take a total of 22 hours. Using a computer, and rudimentary Python code, the entire process of identifying bigrams, extracting those bigrams that are relevant to the main question and counting changes in frequency takes just seconds. The power of computer-assisted text analysis becomes even more evident in the next section.

Contextualising Radical Muslim/s

It is possible to further investigate the context of instances where 'radikal' serves as an adjective specifying 'Muslim'. In order to find the immediate contexts, I once again used TextBlob, this time to divide the texts into sentences. I then extracted all sentences which contained any one of 12 different word combinations in which 'radical' was a specification of a Muslim or of

Muslims. The total number of such sentences was 389 (of a total of 25,516 sentences in all articles). The basic assumption, commonplace in computer-assisted text analysis, was that the words used in the textual surroundings of the word combinations would provide indications of the semantic framing. For comparative material, I used the complete texts of all the articles.

To do this, I created two texts: one which contained the text from the extracted sentences (referred to as the 'sentences text' from now on), and one which contained all the text from all sentences in all articles (referred to as the 'articles text' from now on). During this process, I once again removed sentences that were similar to one another within both sets, setting the threshold for the Jaccard index to 0.5. This reduced the number of sentences in the sentences text to 369.

The next step was a comparative word frequency analysis of the two compiled texts. After restricting myself to words that occurred more than twice in the sentences text, and excluding common words in Swedish (so-called 'stop words') and all words containing 'radikal*' or 'muslim*', the remaining number of what can be termed 'top words' was 232.

I then extracted the 232 most common words from the articles text (again, excluding words containing 'radikal*' or 'muslim*' and stop words). Within these two sets of words, I carried out two searches. First, I identified words that occurred among the top words in the sentences text but did not occur among the 232 most common words in the articles text. I termed this set of words 'sentence-unique words'. Secondly, I compared words that occurred in both sets of words, noting their relative frequencies in the texts from which they were extracted.

The total number of sentence-unique words was 119. Most frequently occurring among these was 'Mohamed' (11 occurrences). Again, the event in 2009 when the poet Mohamed Omar declared himself to be a 'radical Muslim' turned out to have a significant effect on the overall analysis.

Among the sentence-unique words, some, perhaps unsurprisingly, indicated a textual context for 'radical Muslim/s' associated with conflict, violence and war. These were words such as 'terror' (3), 'extremister' (extremists) (3), 'våldsamma' (violent) (3), 'oskyldiga' (innocent) (5), 'slåss' (to fight) (3), 'vapensmuggling' (arms smuggling) (3), jihad (3), 'mördade' (mur-

dered) (3) and 'misstänkta' (suspects) (3).

Another set of words may perhaps be less associated with violent action and more with 'radical Muslim/s' as referring to a particular ideology and ideological influence. Hence, there was a noteworthy occurrence among the sentence-unique words of variants of the word 'imam' ('imam' (3), 'imamer'(5)). Looking more closely at the context in which these words appear, it becomes clear that 'imam' here does not refer to the limited function of leader of prayer, but rather to a religious authority, mainly associated with a mosque. In four out of the five cases where the word 'imamer' (imams) appeared, it was in a context containing a notion that Saudi Arabian 'radical Muslims' were offering financial aid to Swedish Muslims in the form of imams' salaries.

'Slöjan' (the veil) was another 'unique' word that occurred three times. In two of these cases it co-occurred, in the same sentence, with another word in the set of sentence-unique words: 'tvingas' (to be forced to) (3). This could indicate that in the media discourse being studied here, a recurring practice associated with 'radical Muslims' (presumably male) is forcing women to wear the veil.

Some patterns reoccurred in the differences between relative frequencies of words that were found in both texts. Of the 113 words that occurred in both, I limited myself to the 25% of the words which differed the most in their relative frequencies in favour of the sentences text. These were words with a relative frequency in the sentences text that exceeded 1.72 times their relative frequency in the articles text. Among these there was again a set of words clearly relating to violence: 'IS', 'terrordåd' (terrorist act), 'hot' (threat),'krig' (war) and 'våld' (violence). Mohamed Omar reoccurred, this time under his last name, 'Omar', which had a relative frequency in the sentence text which was 3.56 times higher than that in the articles text: this was the second largest difference in frequency.

The greatest difference, however, between a word's relative frequency in the two texts, was for the word 'unga'(young [pl.]). This word was 4.35 times more common in the sentences text than in the articles text. This could perhaps indicate that radical Muslims, at least in Swedish newspaper articles, are conceived of in an age-related way. In fact, searching in the sentences text for recurring three-word collocations (trigrams), the

combination 'unga radikala muslimer' (young radical Muslims) featured 15 times, as the seventh-highest among trigrams containing 'radikal' and 'muslim', following combinations in which the third word in question was one of the following: 'är' (are), 'de' (the), 'som' (who), 'av' (by), 'i' (in), and 'en' (one). The word 'ung' (young [sing.]) also occurs among the sentence-unique words at the relatively high frequency of five.

Unlike in the case of unique words, in the set of common words there were also high frequency differences between words referring to regions and countries (West, Europe, Syria and Irak) and nationalities (American and British), as well as two cities (London and Malmö). 'Grupp' (group) also occurred in the set, which may be an indication that although there may have been a shift over time – noted above – from the use of the combination of 'radikal*' and 'muslim*' in connection with collective nouns (group, party, movement) to 'radical' being an adjective specifying Muslim, there is also a tendency in the latter case for 'radical' to be perceived as something which Muslims are with in connection other Muslims.

It should be noted that the words 'kristna' (Christians) and 'judar' (Jews) were also among the words in the top 25% of words that were more common in the sentences text than in the articles text, and similarly, the words 'kristen' (Christian) and 'judarna' (the Jews) occured in the list of 'unique' words. This could indicate a certain connection between 'radical Muslim/s' and inter-religious relations, perhaps most likely in terms of religious conflict. A closer look, however, indicates that this should be qualified, at least regarding Christianity. 142 of the total 649 analysed articles contained the words Christian or Christians (or both). Of these, however, 29% had been published in just one newspaper: the Christian daily newspaper, *Dagen*. In fact, 74% of articles from that particular newspaper contained one or both of these words.

Concluding Words

As mentioned in the introduction, the background of this article was a workshop on the concept of a 'radical Muslim' and the definition of this. The explicit rationale for that workshop was the notion that the concept was widely used in Swedish public

discourse and was therefore a relevant subject of academic scrutiny. This notion had not been, as such, formed based on any systematic exploration of that discourse, either in its entirety nor of any particular subdivision of it. This article has provided an example of how such an exploration may be conducted, and how computer-assisted text analysis may be highly useful – if not necessary, given the amount of data to be processed – in this context.

The results of the analysis of a subdivision of the discourse (printed news media) does testify to the presence of the concept in this context, although these results cannot determine whether its level of prominence justifies further academic attention. If that is the case, the analysis provides some indications of when the concept came into use, which contexts it appears in most dominantly, and how its use has changed over time. Given the fact that the analysis is not complete (only about half of the articles indicated by the search on Mediaarkivet as containing a combination of the words 'radikal*' and 'muslim*' were included in the actual analysis), these results are tentative. Still, they are arguably better than guesses, and they point to potential patterns, trends and textual contexts that may merit closer study.

This article has exemplified some of the benefits of adding computer-assisted methods to the toolbox used by the humanities scholar. Even such rudimentary and conceptually simple operations as the ones outlined above, which require a minimum of programming skills, can help the scholar to gain a 'bird's-eye' view of a volume of material too great to process manually, hence saving time and energy which can be spent on the kinds of analysis that a humanities scholar does best: interpretation and cultural contextualisation.

References

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

Berry, David M. and Fagerjord, Anders. 2017. *Digital Humanities. Knowledge and Critique in a Digital Age*. Cambridge: Polity Press.

Faimau, Gabriel. 2015. 'The Conflictual Model of Analysis in Studies on the Media Representation of Islam and Muslims. A Critical Review.' In *Sociology Compass*, 9(5): 321–335.

Hansson, Joacim and Svensson,

Jonas. 2020. *Doing Digital Humanities. Concepts, Approaches, Cases*. Växjö: Linnaeus University Press.

Hervik, Peter. 2018. 'Ten years after the Danish Muhammad cartoon news stories: Terror and radicalization as predictable media events'. In *Television & New Media* 19 (2): 146-154.

Klausen, Jytte. 2009. *The Cartoons that Shook the World*. New Haven: Yale University Press.

Montfort, Nick. 2016. *Exploratory Programming for the Arts and Humanities*. Cambridge: MIT Press.

Muhanna, Elias. 2016. *The Digital Humanities and Islamic & Middle East Studies*. Berlin: Walter de Gruyter.

Omar, Mohammad. 2009. 'Gaza har gjort mig radikal'. *Expressen*, 2009-01-09.

Pennington, Rosemary (ed.) 2018. *On Islam. Muslims and the Media.* Bloomington: Indiana University Press.

Poole, Elizabeth. 2002. *Reporting Islam: Media Representations of British Muslims*. London: I.B. Tauris.

Ranstorp, Magnus and Dos Santos, Josefine. 2009. *Hot mot demokrati och värdegrund - en lägesbild från Malmö.* Stockholm: Försvarshögskolan.

Said, Edward W. 1981. *Covering Islam. How the Media and the Experts Determine How We See the Rest of the World*. New York: Pantheon.

Vijaymeena, M. K., & Kavitha, K. 2016). 'A Survey on Similarity Measures in Text Mining'. *Mach. Learn. Appl. Int. J,* 3: 19–28.

Digital resources cited in the text:

Retriever. *Mediearkivet.* (https:// www.retriever.se/product/ mediearkivet/)

TextBlob. *TextBlob: Simplified Text Processing* (https://textblob. readthedocs.io/en/dev/)