

Article:

Prayer, Dreams, and Sex: A Showcase of an AI-Augmented Distant Reading of 20 Thousand E-Fatwas

Jonas Svensson

Abstract

This article provides a methodologically oriented case study of the application of Large Language Models (LLMs) in Islamic studies. Its empirical focus is on 20 thousand e-fatwas from the website askimam.org. The study demonstrates how LLMs, here represented by the open-source tool BERTopic, can significantly augment a process of distant reading. The research focuses on a topic modelling revealing insights into recurring concerns evident from the questions posed by believers in fatwas. It also provides an example of further, post-topic modelling and analysis focusing on thematic variations across questions from different geographical regions. The article argues for the potential of LLMs to reshape research methodologies, offering faster, more comprehensive, and nuanced analyses of large datasets.

Keywords

Large Language Models (LLMs), fatwa, E-fatwa, distant reading, Islamic Studies, BERTopic

Introduction

Imagine, as an Islamic studies scholar, to be able to ask questions in plain language to someone who has read and memorized a very large amount of original works from Islamic intellectual history, and who will respond, in likewise plain language, to any question on content in an instant, without getting tired or bored. Imagine being able to, within a reasonable timeframe, transcribe thousands of audio-recorded contemporary *khutbas* and summarize their content. Imagine being able to compare original theological treatises on a given topic written in diverse “Muslim” languages that you have not mastered. Imagine being able to scan thousands and thousands of *fatwas*, mapping questions that beliefs pose to religious scholars in search of continuity and change in religious concerns over time and space.

In fact, there is no need to imagine. All of the above are already possible, due to recent advances in the technology of artificial intelligence within the fields of Natural Language Processing and Natural Language Understanding. The current article takes on the last of the imaginary tasks above to substantiate a forward-looking assertion: the advent of the technology of Large Language Models (henceforth LLMs), notably exemplified to the public by Open AI’s introduction of ChatGPT in late 2022, is poised to revolutionize numerous research domains. These domains include various aspects of the humanities, understood as the study of the expressions of the human mind, both contemporary and historical, including the Study of Religions and its subfield, Islamic studies.

LLMs represent a sophisticated branch of artificial intelligence focused on processing, comprehending, and generating natural human language. The *large* in their name refers to the enormous amount of data they have been trained on, which enables them to learn and draw inferences from intricate language patterns. LLM processes diverse and expansive datasets, predominantly textual, to predict language patterns and contextual nuances. Their capability extends beyond mere data processing; LLMs exhibit an almost human-like proficiency in language generation and understanding.¹

¹ For a good, nontechnical introduction, and historical contextualization of LLMs, see Manning (2022). For a more technical, but also more comprehensive overview, see Naveed et al. (2023). For an introduction to the larger issue of Generative AI, see Banh and Strobel (2023).

As has been suggested in the recent year and a half (late 2022 to mid-2024), LLMs hold great potential to augment research on language and by extension on culture (Meyer et al. 2023). Already at this point in their development (as of mid-June 2024) LLM shows astonishing capabilities in performing some research tasks basic to the humanities, such as reading, translating, and comparing texts, and extracting themes and key concepts. These are tasks that used to claim days, weeks, months, years, and even a large part of academic careers, but that now be achieved in minutes or even seconds. In addition, the technology is becoming increasingly accessible even to researchers with limited or no programming skills.

This article gives a limited, but easy-to-follow case of an LLM-augmented “distant reading” (Moretti 2013). It does so through a so-called “topic modelling” of 20 thousand online fatwas from the e-fatwa site *askimam.org*, employing the open-source tool BERTopic (Grootendorst 2022). This tool offers a relatively transparent means to showcase, but in no way exhaust, the capabilities of LLMs to augment humanistic research. By engaging with a limited set of “born digital,” naturalistic data (Potter and Shaw 2018), the article aims to contribute to the ongoing discussion on how LLMs can enhance research methods in Islamic studies and beyond. The article also makes some use of some of the capabilities of ChatGPT4, to exemplify how also this particular tool may be integrated into a researcher’s workflow.

Short on fatwa research

Although the aim of this article is primarily methodological, the analysis performed in the concluding case study does attach to a larger, established research area within Islamic studies, a research area that deserves some elaboration.

What is a fatwa?

Simply put, a *fatwa* is a formal interpretation or “opinion” on Islam, or more precisely on aspects of *Shari’a*, given by a legal scholar—or sometimes, particularly in modern times (Masud, Messick, and Powers 1996b, 27–28),

a collective of scholars—acting as a consultant, a *mufti*. It is issued in response to a question posed by a questioner (the *mustafti*) who may be an individual, a legal institution, or a political body, and it addresses specific issues, often related to daily life, morality, application of legal norms, or ritual. Fatwas are usually non-binding (i.e., they are not legal verdicts, albeit they can form a basis for such) but may be influential both on a personal and a societal level if the *mufti* (a person or an organizational body) in question is invested with personal or formal authority, in the latter case part of a state apparatus.

Fatwas are an old phenomenon.² In the sense of believers seeking and receiving advice from learned others, *fatwa* is probably as old as Islam as a religious tradition. According to Masud, Messick, and Powers (1996b, 9), the first written collections of fatwas started to appear in the second half of the tenth century C.E., and this activity of collecting fatwas has continued since then. In this sense, e-fatwa collections, such as the one in focus here, are but a contemporary development, and above all an upscaling of a long tradition. E-fatwas are here defined widely as fatwas in electronic form made publicly accessible to everyone on the internet. E-fatwas are often hosted on dedicated websites run by individuals, organizations, or even state bodies.

Research on fatwas

Fatwas, and fatwa issuing, have been objects of extensive research, and this in several respects. Scholars in the history of Islam have used fatwas as source material to map the theological thinking of individuals, institutions, or schools of thought, the development of Islamic law, or the cultural, social, or historical context in which the fatwas were issued (e.g., see Noralla 2023; De Angelo 2017; Caeiro 2006; Larsen 2018; Schussman 1998) Less common are more anthropologically oriented studies of, for example, views on the authority of fatwas and the role of *fatwas* in Muslims' religious meaning-making (Agrama 2010; Whyte 2022).

² For a recent overview of the cultural phenomenon of *fatawa* from classical times to the present, see Awass, (2023); for a number of case studies throughout the centuries, see chapters in Masud, Messick, and Powers (1996a).

Some well-known and influential fatwas in history have been studied in detail, both concerning content and impact (see e.g. Weiss 1996; Hamzah 2005; Copinger-Symes 2017). Research has also been conducted on the historical, social, and political role of different fatwa issuing bodies (also state-supervised or affiliated) or famous *muftis*, as well as on different modern means of dissemination (see e.g. Al-Marakeby 2022; Skovgaard-Petersen 1997; Gräf 2009; Pasuni 2018; Kemper 2022).

E-fatwas as digital religion

These general tendencies in the study of fatwas are also evident in studies of e-fatwas, albeit these have also generated a new set of questions related to the means through which they are disseminated. These questions are not exactly new. Every advance in communication technology, from the printing press, via radio and television, to contemporary social media, has stimulated research questions relating to changes in the production and dissemination of *fatwas* (Masud, Messick, and Powers 1996b, 30–31; Illar 2010). Hence, some contemporary studies (e.g. Ibahrine 2020; Bunt 2003; Caeiro 2013; Kutscher 2009; Marcotte 2016; Šisler 2011a; Illar 2010; Larsen 2016; Ragab 2023) have addressed e-fatwas as a phenomenon of particular interest within the wider research area of “digital religion” or “digital Islam.”³ Important questions raised here concern the effects of e-fatwas, and globally available “cyber Muftis,” on larger processes of modernity, such as the “fragmentation of religious authority” (Eickelman and Piscatori 2004, 70), Muslim diasporic identity construction, and radicalization (e.g. Whyte 2022; Rusli and Nurdin 2022; Weimann 2011; Šisler 2011b, 2009). Such questions are not addressed in the current article, which has a focus on e-fatwas as data, rather than on the phenomenon of production and distribution of fatwas online.

³ This research area is vast and continuously expanding, raising important questions of the effects of technology on Islamic religious beliefs, practices, and forms for organization. To provide an overview of the field here would be a formidable task, and divert attention from the main issue. For a very recent attempt by a pioneer in the field, see Bunt (2024, 16–23).

E-fatwas as data

E-fatwas constitute a public, easily available, varied, and large type of data. They have already been used as such in studies on a wide range of issues in contemporary Islamic discourse, for example, the legitimacy of tattoos (Larsson 2011), responses to Covid-19 (Svensson 2021), bioethics (Van den Branden and Broeckaert 2011), and Salafi thought and practice (Olsson 2017), to mention but a handful. Such studies often use e-fatwas alongside additional primary and secondary material, historical and contemporary.

Other studies using e-fatwas as data have been more directly targeting what could be termed dedicated e-fatwa sites, such as the one selected for this article, and e-fatwas they contain, but still often with a clearly limited focus on responses to particular issues. Researchers here tend to use the sites' internal search engines or topic trees of categories in purposive sampling of a small number of fatwas studied in depth (see e.g. Singer 2021; Wynn and Foster 2018; Ismail 2015; Kutscher 2011b; 2011a).

Search engines and topic trees are common features of dedicated e-fatwa sites, such as *askimam.org*. Some also allow for the visitor to pose questions. The question and the answer are then published and made part of the sites' continuously expanding database. These sites hence constitute examples of the kind of user-provider interaction and user cocreation of content that in research has been referred to as Web 2.0, as distinguished from the provider-only Web 1.0. Examples of this form of fatwa sites in English (online on November 17, 2023) besides *askimam.org*, are *islamweb.net* and *islamqa.info*. These could then be distinguished from sites that mainly disseminate fatwas, like the online services of fatwa-producing bodies such as the Saudi Arabian *Permanent Committee for Scholarly Research and Ifta* or the Turkish Presidency of Religious Affairs *Dinayet* (see Yakar 2022).

It should be remembered that the way e-fatwas are conceptualized in the present article does not limit them to fatwas present on dedicated fatwa sites. Not only are other, perhaps even more influential, platforms for the public distribution of fatwas available to scholars or institutions: YouTube, stand-alone applications

in smartphones, Facebook, Twitter–X, TikTok, et cetera. The conceptualization also allows for the transformation of non-e-fatwas into e-fatwas through the change of the medium for and scope of distribution. Hence, also digitized fatwas from the pre-internet era, and fatwas produced in, for example, TV or radio shows, but transferred onto for example YouTube, become e-fatwas when made generally available on the internet.

So far, I have not been able to locate any attempt within the academic study of Islam at large-scale collections of e-fatwas. This is perhaps understandable. Collecting data by manually copy pasting thousands and thousands of fatwas and their metadata from a diversity of platforms on the web in a document or spreadsheet for further analysis is not feasible. The number of e-fatwas that could potentially be housed in such a collection is large, indeed very large.

In an attempt to train a neural network to serve as a virtual mufti, Munshi et al. (2022) collected more than 850 thousand question-and-answer pairs from ten publicly available, online sources. The number of fatwas collected is impressive, but still limited. The authors explicitly restricted themselves to what they viewed as Islamically “trustworthy” sources in the Arabic language (Munshi et al. 2022, 1146), which in practice meant Sunni, Salafi-oriented, websites. It is not unreasonable, from this, to assume that the actual number of publicly available e-fatwas counts in millions. These millions of e-fatwas could prove to be a goldmine for systematic research, particularly if such research is augmented with computer-assisted methods and in particular LLMs that can read, understand, organize, and summarize their content.

Selecting and collecting the material

As mentioned above, I here limit myself to one e-fatwa website, *askimam.org*. This is purely out of convenience. The aim here is not to provide a large-scale overview of e-fatwas but to point to a method that can, if need be, be scaled up considerably.

Askimam.org is based in South Africa. It is a fairly large site but not massive. It is also well-structured. The fatwas are originally in English. This gives it

a potentially wider audience than fatwa sites in, for example, Arabic or Indonesian. It is also a site that has existed for a long time and has been cited in previous research (see e.g. Šisler 2009, 67; Kutscher 2009, 39; Mishra and Semaan 2010; Larsson 2011, 240; Maravia et al. 2021).

A master's thesis with a detailed analysis of six fatwas from *askimam.org* within a theoretical framework of female sexual agency, by the author Farhana Ismail from the University of Kwazulu Natal, provides some context (Ismail 2015). The founder, Ahmad Desai, who passed away in 2021, was a scholar of South Asian descent, belonging to the Hanafi *madhhab*, and more particularly the Deobandi movement. The Deobandi movement is a mixture of scripture-oriented 18th-century South Asian Islamic reformism and “sober” Sufism. In South Asia (and diaspora contexts), it stands most clearly in contrast to the more folk tradition-oriented and charismatic Barelvi movement (see e.g. Metcalf 1982; Usha n.d.; Ingram 2018). Trained both in India and South Africa, Desai was active in creating an infrastructure of religious education and training in South Africa (Ismail 2015, 18). The production of fatwas, and the creation of a fatwa providing local infrastructure, is part of Desai's institution *Darul-ifta al Mahmudiyyah*'s activities (Ismail 2015, 19).

I will not dwell on the Deobandi, Desai, and the local South African context here. All of these would have been important if the aim was to explain or understand the *fatwas* published on the website, particularly the answers provided by Desai and his students, but the focus does not lie there.

Collecting the fatwas

All fatwas publicly available on *askimam.org* were collected over 12 days in December 2022. For each fatwa, the following data and metadata were collected (if given): (1) the unique ID (a number) of the fatwa, (2) the date the question was received (3) the national origin of *mufti* (4) a descriptive title for the fatwa provided by the administrators (5) the “raw” question (6) the response given by a *mufti*, and (7) the category under which the administrators of the site had placed the fatwa. Instead of manually collecting the information from the web interface through a process of copy and paste, I used a custom-built script

in the programming language Python, iterating through the pages extracting the relevant information, with significant pauses in order not to overload the *askimam.org* server.

Overview of the fatwas

Temporal and geographical distribution

The total number of fatwas, after removal of duplicates, was 23,694. These stretch over a period of 23 years (with the question in the first fatwa received on June 27, 2000, and the last on December 9, 2022). A small number of fatwas (31) had the odd, and obviously erroneous, dating of the question to January 1, 1970. Figure 1 displays the distribution of fatwas over time.

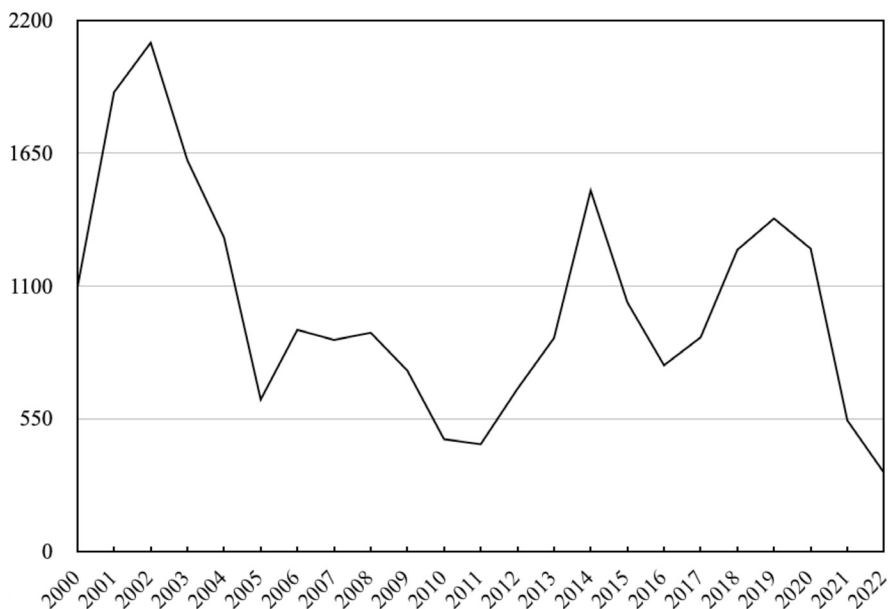


Figure 1. Number of fatwas over time.

While there is a peak in the early days of the site, the number appears to have stabilized somewhat over the period from 2003 to 2020, with a significant

drop from 2021, the year Desai passed away. It should be remembered that the graph shows the number of fatwas *published* on the site. The date given is also that of reception, not of publication. There is no way to determine whether the collection houses the total number of questions received. One could expect a process of selection. However, as was discovered in the topic modelling below, the collection contains some (48) fatwas with the title *unanswered*. This shows that more questions have been received than were attended to. In most cases, there is no reason given for the lack of a response. However, in a few cases, there are. The reasons involve the following (1) the fatwa site accepts only questions in English; (2) the question is “off topic;” (3) the question is about sexual etiquette. The single *unanswered* fatwa in the last category concerns masturbation and was received on April 17, 2008. It merits the comment: “After a Mashwarah [joint counselling] it has been decided that We will not be answering any more sexual oriented question [SIC!]” and a reference to an e-publication particularly addressing this issue of masturbation (*Ask Imam* 2008). The result from the topic modelling indicates that this decision was not followed through.

Figure 2 (constructed using ChatGPT4) shows the geographical distribution according to the information provided under the headline *Country*. Questions appear to come from all over the globe, but they are not equally distributed. The top ten countries and the number of fatwas from each is given in Table 1.

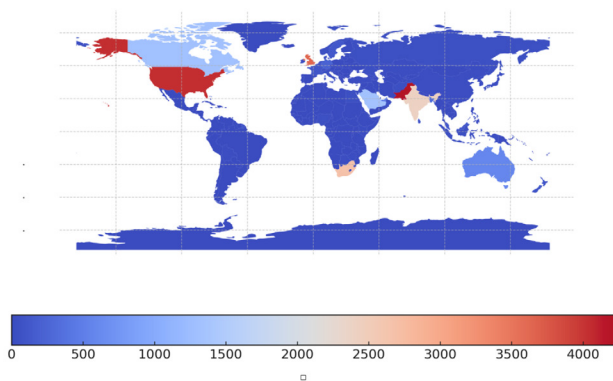


Figure 2. Number of fatwas by mustafti’s given country of origin. Generated with ChatGPT4.

Country	Number of fatwas
Pakistan	4,241
USA	4,077
United Kingdom	3,672
South Africa	2,673
India	2,419
Saudi Arabia	1,339
Canada	1,315
Australia	599
United Arab Emirates	523
Bangladesh	324

Table 1. Top ten countries.

South Africa’s position on the list is most likely a result of the location of the site and its organization. That the South Asian countries of Pakistan, India, and Bangladesh are on the list may be attributed to the site’s religious affiliation (Hanafi and Deobandi) but not necessarily so. Taken together, these countries house roughly a third of the world’s total Muslim population. They are also all former British colonies and English is an official language in both India and Pakistan. The fact that *askislam.org* is exclusively in English can account for the large proportion of questions from Great Britain, USA, Canada, and Australia. Ducking the task of deciding what constitutes “Western” countries, I asked ChatGPT4 to provide a list and calculated the percentage of questions originating from the countries on that list. The result was 44 percent. A ChatGPT4 list of countries in the “Global North” yielded the same result. Computer literacy and internet availability is another factor, besides the English language, that could account for these large numbers. This could also be a partial explanation for why both Saudi Arabia and the United Arab Emirates appear among the top

ten countries. According to statistics from the *World Bank Group*, Saudi Arabia and the UAE both belong to a group of six countries worldwide that claim that 100 percent of their population use the internet (*World Bank Group* 2024).

Internal Categorization

As mentioned above, *askimam.org*, like other fatwa sites, categorizes the fatwas in the database. In total, the fatwas in the collection are divided into 47 categories. Table 2 shows the ten largest.

Category	Number of fatwas
Misc. Fiqh	2,675
Jurisprudence and Rulings (Fiqh)	2,575
Finance and Commerce	2,476
Beliefs and Practices (Aqeedah)	2,286
Marriage (Nikah)	1,926
Prayer (Salaat)	1,656
Character and Morals	1,259
Divorce (Talaaq)	827
Ramadan	563
Miscellaneous	551

Table 2. Ten largest internal categories.

It can be noted that in total 5,700 fatwas, around 20 percent of the grand total, belong to the generic categories of *Misc. Fiqh*, *Jurisprudence and Rulings*

(*Fiqh*), and *Miscellaneous*. In addition, there is a category of *Unknown* (containing 40 fatwas). These labels are not informative.

While not all visible in the Table 2, there are several categories that could be seen if not as synonymous, then at least similar enough to be merged, such as *Fasting* and *Fasting (sawm)*; *Divorce* and *Talaq*; and *Prayer* and *Prayer (Salat)*. I decided to perform some mergers (see Appendix 1), thereby reducing the total number of categories to 29. These merged categories will in the following be referred to as the *internal categories*. It should be stressed that these are top-down categorizations. One may assume that they, at least in part, mirror what the site administrators consider to be appropriate fatwa categories. As will be shown, the topic modelling produces a slightly different categorization.

The topic modelling

Topic modelling is a common name for computer-assisted techniques to extract semantically coherent themes, mainly in texts or collections of texts. Several such techniques have been developed. Concerning text material, most methods have so far been based on word frequency analysis, word co-occurrence and diverse probabilistic computations. Topic modelling techniques pre-LLM are conceptually complex and often require extensive text preprocessing and domain knowledge in interpreting the results, which also tend to be sensitive to different initial settings. Some also provide results that are difficult to reproduce over consecutive runs (Golub 2020).

BERTopic uses a novel method of topic modelling which is conceptually, but by no means computationally, simpler than earlier methods, and somewhat more stable. It relies on a pre-trained and fine-tuned LLM to place chunks of texts (typically sentences or shorter paragraphs) in a multidimensional vector space based on their *semantic* content, in a manner so that semantically similar chunks are placed closer to one another. This is called *sentence embedding*. From this embedding, BERTopic identifies clusters of similar chunks and compares these clusters to one other based on relative word frequency. The result is a *topic* represented by the most significant words in the text chunks contained within a cluster.

The method has several advantages compared to many previous methods of topic modelling.⁴ It requires little, or no, preprocessing of the data (e.g., removing common words, or reducing words to basic forms). In fact, such preprocessing may obscure the semantics of the text chunks. Since BERTopic relies on semantics, rather than word occurrence—and the models it relies upon in a sense read and understand the content of the chunks—it is less sensitive to lexical variation, and also to the actual language. While the data for this article is exclusively in English, or perhaps more apt “Islamic English” (Al-Faruqi 1986) spiced with religious terminology, BERTopic, can handle also multilingual data.

Deploying LLMs, or more widely Generative AI, in research may give rise to ethical concerns. Some of these relate to how the models have been trained. Recently, the training procedures for LLMs, especially the most competent ones such as GPT-4, have come under some scrutiny, as some most likely have been trained on copyrighted data. The default version of BERTopic used in this article utilizes the model *all-MiniLM-L6-v2*, which has been trained on a wide array of publicly available and open internet data (for details of the training datasets, see *Hugging Face* n.d.).

Another ethical concern is the research data presented to the model. In the current article, the data used is publicly available and cannot be traced back to any specific individual. Furthermore, it is not collected on the model’s end and used for further training. In LLM-augmented research, however, care must be taken to protect sensitive or copyrighted data.

Focus

As mentioned above, most previous studies that focus on the content of fatwas have concerned muftis’ *responses*, particularly on specific issues. In the following, I will instead focus on the questions. This is because I assume

that the questions can provide interesting information on patterns in the religious concerns of those, assumed ordinary believers, who turn to the page for advice.

The average length of all questions is 131 words, but they range from a single word to 9,744 words. At present, a commonplace upper limit for successful text embedding is around 250 words. A text chunk exceeding the maximum number will often be truncated. A possible solution to the problem would be to utilize an LLM to summarize the questions to a maximum of 250 words. Text summarization is one of the tasks at which LLMs excel. Text length, however, is not the only problem. Not all fatwa pages on the site provide text under the headline *Question*. All do, however, provide text under the headline *Title*. This is where the site's administrators present the theme of the question, with a set of descriptive words, or through merely iterating the question in full or abbreviated. The lengths of these *Titles* are more consistent, with an average of 15 words and a maximum of 96. Hence, the text under the headline *Title* was used in the following.

The result

The procedure of embedding the texts and identifying topics in BERTopic took two minutes. Using the default settings, the number of topics returned was 320, hence by far exceeding the number of internal categories. The minimum number of fatwas placed under each topic was here set to ten. Changing this setting—raising the number—generally reduced the total number of topics. Topic modelling with BERTopic, unlike several other topic modelling methods, also produces outliers, in this case, titles that could not be identified as belonging to a certain cluster of similar texts, and that hence are not placed under a topic. The number of outliers produced also varies with the minimum number set.

After experimenting with different minimum topic sizes between 10 and 100, in steps of ten, I reached the conclusion that the least outliers were produced if the minimum size of topics was set to 50. This produced 53 topics. The number of outliers, albeit the lowest, was still quite large (6,595). The number

of outliers can be reduced in BERTopic, forcing them under topics, but for the following, I chose not to. The analysis from here on disregards the outliers, reducing the total number of fatwas analyzed to 16,937.

Interpreting the topics

The 53 topics returned by the topic modelling are displayed in the form of a dendrogram in Appendix 2. In the dendrogram, they are arranged not in order of size (indicated by their number from 0 to 52, with 0 being the topic housing the largest number of fatwas, and 52 the smallest, i.e., the minimum of 50 fatwas) but by similarity. In many cases, the grouping makes sense, for example, the topics 34, 39, 15, and 36 are grouped together and their representation does indicate a common theme of marriage.

With a few exceptions, the topics also make sense. In Table 3, I limit my main comments to the 15 largest topics (topics 0 to 14), covering roughly half of all the fatwas in the original collection. These are the following, with representations in the form of the most characteristic words for each topic, the default output of the topic modelling.

Topic	Topic representation	Number
0	[permissible; money; bank; halal; company; insurance; use; work; business; food]	4,493
1	[salaah; salah; pray; salaah; prayer; masjid; salat; prayers; praying; time]	1,903
2	[zakat; zakaat; pay; pay zakat; money; pay zakaat; gold; zakaah; land;	616
3	[meaning; names; meanings; child; know meaning; suitable; naming; know; tell meaning; baby]	575
4	[fast; fasting; fasts; break; ramadan; break fast; days; missed; month;	522
5	[allah; prophet; ibn; did; shias; ra; saw; muhammad; shia; ali]	476
6	[dua; duaa; duas; make dua; make; duaas; read; recite; making dua;	392
7	[hajj; umrah; haj; perform; mahram; perform hajj; travel; ihram; hajj year; going]	368

8	[quran; read; quraan; read quran; translation; reading; holy; books; english; arabic]	357
9	[sex; wife; masturbation; husband; intercourse; oral; parts; oral sex; masturbate; private]	355
10	[dream; dreams; interpretation; interpret; dream interpretation; interpret dream; saw; dream dream; istikhara; interpretation dream]	347
11	[pregnant; pregnancy; birth; abortion; baby; child; contraception; birth control; permissible; wife]	332
12	[islam; muslim; muslims; jihad; country; religion; islamic; nonmuslim; state; non]	322
13	[divorce; wife; husband; divorced; constitute divorce; divorces; divorce wife; constitute; divorce husband; divorce divorce]	317
14	[inheritance; property; estate; father; sons; passed; daughters; inherit; passed away; share]	310
Total		11,685

Table 3. Top 15 topics discovered by BERTopic with representative words and a number of fatwas.

Several of the topics appear, at first glance, to correspond quite well with the internal categories provided by *askimam.org* (see Appendix 1). This is so for *Prayer* (which corresponds with topic 1); *Fasting (Sawm)* (topic 4); *Charity, Zakat, Lillah, Sadaqah* (topic 2); *Wills and Inheritance* (topic 14); *Quraan* (topic 8); *Supplications (Duaas)* (topic 6); *Hajj and Umrah (Pilgrimage)* (topic 7); and *Divorce (Talaaq)* (topic 13). This impression is correct, to various degrees. The mapping is not total. On the one hand, 93 percent of the fatwas placed under the internal category of *Fasting (Sawm)* are also present under topic 4. On the other hand, only 39 percent of the fatwas under the internal category of *Divorce (Talaaq)* are found under topic 13. In the latter case, this may be at least partly attributed to the fact that the topic modelling resulted in at least three topics possibly related to divorce (18, 13, 27, see Appendix 1). This indicates that the topic modelling provides a more fine-grained categorization, alternatively, one that is sensitive to the terminology used (see below). If all these topics are included, the percentage rises to 76 percent.

The largest topic discovered, 0, does not appear, from its word list representation, to match any internal category, only a similar one. This internal category is incidentally also the largest (apart from the generic category *Miscellaneous*): *Finance and Commerce* (see Table 2 and Appendix 1). Some of the words in the topic representation of topic 0, such as “money,” “bank,” “insurance,” and “company” suggest an overlap. Of the 2,356 fatwas belonging to the internal category of *Finance and Commerce* (outliers not considered) 91 percent are also found under topic 0. But the topic covers almost the double number of fatwas. The other words in the topic representation, that is, “permissible,” “halal,” and “food” indicate that something else is hidden here. True enough. The internal category *Food and drink* has an even larger proportion (96 percent) of its fatwas under topic 0. Taken together with the (fused) category *Miscellaneous* the total number of fatwas in these three internal categories makes up 88 percent of the fatwas under topic 0. Performing a secondary topic modelling on the fatwa titles categorized under 0 reveals that there are important sub-topics relating to food, drink, and smoking. This also gives an indication as to what is housed under the internal category of *Finance and Commerce*, that is, fatwas that concern the legitimacy of certain commercial transactions in terms of *halal* (allowed) and *haram* (forbidden)—rather than general business advice.

Topic 0 is dominant. There are other topics that can also be assumed to cover issues of *halal* and *haram*, for example, the well-known issue of the permissibility of men wearing gold jewellery (topic 32). A possible explanation for this is that questions pertaining to *halal* and *haram* placed under topic 0 are too diverse to generate separate topics, given the minimum threshold of 50 fatwas per topic. If this minimum threshold is lowered, the topic distribution becomes more fine-grained, and the dominance of a single large topic 0 disappears.

Some of the representations (i.e., list of words) of the top 15 topics are not as easily interpretable. This is true for topic 5 and topic 12, which would appear to contain generic references to key concepts and names in Islam. Here, however, a closer look at the intersections with internal categories may provide some hints as to the content. Topic 5, for example, shares 45 percent of its fatwas with the internal category *Beliefs and Practices (Aqeedah)* and 24 percent with the internal category *Biography and History*. It could be assumed that within this

topic one finds general questions on Islam as a religious tradition. Similarly, fatwas placed under topic 12 are mainly found in the category *Miscellaneous* (31 percent), but also in two other large ones: *Beliefs and Practices (Aqeedah)* (24 percent), and perhaps more significantly in the category of *Islamic Politics* (21 percent). This latter placement indicates that the topic covers questions relating to contemporary issues in Islamic politics.

The four remaining topics among the top 15 are especially interesting because they, although quite easy to decipher, have no correspondence in the internal categorization. The most prominent, topic 3, concerns the meaning of personal names. Most of the fatwas under the topic are placed under *Miscellaneous* (41 percent), *Character, Morals (Akhlaq)* (30 percent), or *Child upbringing (Tarbiyyah)* (11 percent).

Topic 9, which, given the word representation, mainly concerns sexuality has a similarly wide distribution over the internal categories, but with a certain bias towards the category of *Marriage (Nikah)* (34 percent), *Miscellaneous* (28 percent), and *Character, Morals* (19 percent). Topic 11, which also could be construed as related to issues of sexuality, is mainly distributed over the same three categories, but in a slightly different order: *Miscellaneous* (45 percent), *Character, Morals (Akhlaq)* (10 percent), and *Marriage (Nikah)* (9 percent). It can be noted that although there was a stated decision in 2008 not to answer questions on sexual issues anymore, there has been no significant drop after that time in these two topics.

Topic 10 indicates that *askimam.org* is often consulted in dream interpretation, a traditional task for religious scholars, not least within Sufism (Kinberg n.d.). Incidentally, the prominent place for dream interpretation, which is not represented by an internal category, corresponds with findings in previous research on fatwas (Kinberg 2012). The fatwas placed under this topic mainly belong to the two internal categories of *Miscellaneous* (57 percent) and *Character, Morals (Akhlaq)* (22 percent).

In sum, comparing the internal categorization and the results of the BERTopic modelling shows that the latter produces a result that in parts corresponds to

the former, but also provides a more fine-grained image. It also reveals patterns in the questions that are not evident from the way in which the administrators have chosen to categorize the fatwas. Given the prominent place of the topics of dream interpretation, the meaning of names and sexuality, and their distribution over several internal categories, it is noteworthy that these have not received their own internal categories. There may be several reasons for this. One is that the internal categories depend on the site administrators' views on what traditional issues in *fiqh* that *fatwas* should cover, which then does not fully match the questioners' concerns. One could here also notice less prominent topics such as those apparently relating to listening to music (38), plastic surgery (43) or university studies (33) which all are predominantly placed in the internal category of *Miscellaneous* (at a level of 77, 61, and 86 percent respectively). Another reason may simply be that these topics have not been spotted in the manual internal categorization process.

As noted above, most previous research addressing the content of e-fatwas (and fatwas in general) approach the material with questions on how particular issues are discussed, using the site's search engine or internal categorization to extract relevant data (for a prime example of this, see Kutscher 2011a). The method suggested here allows for more exploratory research with an empirically founded, "bottom-up" approach asking questions of what issues are actually addressed, and then continuing with an in-depth study of those issues.

However, also research with purposive sampling of data on a particular issue can benefit from the suggested method. The identified topics, due to the embedding procedure, can help extract fatwas in the collection relevant to a particular issue even if those fatwas do not contain certain keywords associated with it. Once the topic modelling has been performed, BERTopic allows for a semantic search of topics. This means that the user may enter a search phrase or search word that *does not appear in the texts* and still extract the topics most relevant to the issue. Entering the search word "sexology," for example, which is not present in the fatwas, will return a list of topics ordered by relevance, with topic 9 in the top position. From this, anyone interested in the sexology of *askimam.org* can proceed with an in-depth study of the fatwas placed under those topics.

These suggestions on how topic modelling may be used both for a purposive sampling of fatwas and for discovering otherwise hidden, prominent themes (e.g., sexuality, naming practices, and dream interpretation) in a large collection of fatwas do not exhaust the possibilities. In conclusion, I will use the results in this limited case study to illustrate how the process of topic modelling with BERTopic can facilitate the pursuit of other potentially interesting research questions.

A case of comparative Islamic studies: Regional differences

Above, I mentioned how I used ChatGPT to create a list of countries in “the West.” I did the same with the more clearly geographical regions of Africa, the Middle East, and South Asia.

For each topic, I then assessed whether the number of fatwas on that topic in a particular region was lower or higher than the expected number given the overall relative frequency. In this way, I could discover whether a particular topic was overrepresented or underrepresented in questions from a particular region. Using a binomial test, I also tested whether such over- or underrepresentation was statistically significant. The result is displayed in Table 3. For “the West,” “South Asia,” and “Middle East” I decided on a threshold of $p < 0.1$, that is, the noted over- or underrepresentation had a less than ten percent probability of being the result of chance. For Africa, which appeared to be the region that deviated the most from the others, this threshold could be set to $p < 0.05$.

Instead of providing a list of characteristic words for the topic, as is the default option for topic representation in BERTopic, I here opt for more condensed topic labels consisting of a word or a phrase, but still keeping the original number of the topic. The new representation is in some cases the result of my interpretation of the original representation (i.e., the list of words). However, in other cases, I asked ChatGPT4 for help with constructing a suitable label, based on the original word lists as well as the texts used in the topic modelling. The latter, ChatGPT-suggested labels, are marked with an asterisk in Table 4.

Region	Overrepresented	Underrepresented
South Asia (p<0.1)	36. Permissibility and implications of marrying relatives (p=0.02)* 40. Jinns and black magic (p=0.08)*	10. Dream interpretations (p=0.00)* 35. Menstruation (p=0.02)* 43. Cosmetic procedures and body modifications (p=0.02)*
West (p<0.1)	10. Dream interpretations (p=0.00)* 24. Hair maintenance and alterations (p=0.02)* 29. Hijab and niqab (p=0.07)* 43. Cosmetic procedures and body modifications (p=0.06)*	0. Halal and haram (p=0.00) 2. Zakat (p=0.00) 7. Hajj and 'Umra (p=0.00)* 13. Divorce (p=0.04) 14. Inheritance (p=0.04)*
Africa (p<0.05)	0. Halal and haram (p=0.00) 2. Zakat (p=0.00) 14. Inheritance (p=0.00) 18. Talaq (p=0.00)	5. General history and theology(p=0.01)* 8. Qur'an (p=0.02) 9. Sex (p=0.05) 12. Islam and contemporary political and social issues (p=0.00)* 13. Divorce (p=0.00) 24. Hair maintenance and alterations (p=0.04)* 29. Hijab and niqab (p=0.01)* 30. Mental and emotional health concerns (p=0.01)* 42. The concepts of shirk and kufi (p=0.00)* 51. Nikah (p=0.04)
Middle East (p<0.1)	2. Zakat (p=0.03) 7. Hajj and 'umra (p=0.00)* 26. Dress codes (p=0.06)*	3. The meaning of names (p=0.01)* 4. Fasting (p=0.03) 24. Hair maintenance and alterations (p=0.01)*

Table 4. Overrepresented and underrepresented topics in different regions

Some of the possible regional differences merit comment. Africa, which to a large extent (88 percent) means South Africa holds the most topics, especially underrepresented ones, even if the threshold is set to $p=0.05$. The character of the overrepresented topics is interesting, particularly considering the overlap with underrepresented topics in “the West.” At least two of the overrepresented–underrepresented topics (*Talaq* and *Inheritance*) relate to the application of *Shari’a* in the field of personal law. One reason for this could be that in South Africa, Islamic family law and laws of succession apply in the legal system, which is not generally the case in for example Europe and the United States (McDonald 2014; Ismail 2015, 20–22). It is conceivable that to at least a portion of South Africa’s Muslims, *askimam.org* fulfills a traditional role of a fatwa-producing body, that is, as a consultative instance in relation to the practical administration of Islamic law. One can in this context note the quite long list of underrepresented topics, which houses several topics of more general character (concerning religious beliefs and contemporary issues). That *Divorce* is underrepresented, and *Talaq* overrepresented, in the case of Africa, may be attributable to a terminological bias in the underlying model (not recognizing that the words are synonyms) as well as regional differences in the use of “Islamic English.”

The list of overrepresented topics in the West is also interesting. The overrepresentation of the overall relatively popular topic of *Dream interpretation* could indicate a lack of a local dream interpreting infrastructure, leading to a tendency to turn to online services. The overrepresentation of the topics *Cosmetic procedures and body modifications*, and *Hair maintenance and alterations* could indicate a certain individualist orientation. It is already known that the issue of female dress, mainly veiling, is a contentious issue in Europe and North America, so the overrepresentation of the topic *Hijab and niqab* comes as no surprise.

Turning to South Asia, the overrepresentation of the topic *Permissibility and implications of marrying relatives* does rhyme quite well with estimations that Pakistan and India have the globally highest rates of consanguineous marriages (Iqbal et al. 2022). Perhaps the above-suggested infrastructure for dream interpretations is more established in South Asia, which would account for its underrepresentation.

In fatwas originating in the Middle East, one finds an overrepresented topic with the largest deviation from the overall distribution: *Hajj and 'umra*. Most of the fatwas from this region come from Saudi Arabia and the United Arab Emirates. Saudi Arabia is the site of both forms of pilgrimage and the UAE is neighboring Saudi Arabia. The opportunities for Muslims residing there to participate in these ritual events are probably higher than for Muslims residing elsewhere. One can here also note the underrepresentation of the topic of the meaning of personal names. Given that it is in the Middle East region that Arabic is mainly spoken and that Muslim names often have Arabic origin, it is possible that this topic may not be of prime interest to questioners residing here.

Of course, these interpretations of the relative distribution of topics over the different geographical areas require further investigation, through accessing the content of both questions and answers in the fatwas. This short application is meant only to exemplify how one can take the results of the topic modelling further and generate hypotheses that could be pursued further, concerning for example differences between different time periods

Discussion and Concluding Remarks

The main purpose of this article has been to provide an example of how LLMs can augment Islamic studies and how LLM-based methods can be a valuable addition to the researcher's toolbox. The case is highly limited and barely scratches the surface of the potential of the new technology. It was chosen because of its relative transparency, and because it was possible to connect to an established research area within Islamic studies.

A fatwa collection like the one constructed and used for this case is suitable for this kind of topic modelling, especially when, as was the case, the titles provide condensed information on the themes of the individual fatwas. Other kinds of text material may not be that well-structured, but as long as it is possible to divide the material into somewhat semantically coherent text chunks (sentences, paragraphs) similar topic modelling can be done on such material.

Topic modelling is a computer-augmented technique for doing what is often done manually in much traditional text analysis, that is, reading texts and rearranging parts of these texts into semantically coherent themes. The difference lies in the speed and effort required. As mentioned above, the topic modelling of 20 thousand text units by BERTopic took two minutes. Just as a comparison: to manually read through the texts (without any attempt at organizing and extracting topics), assuming an average reading pace of 300 words a minute, would have claimed 20 hours. I have not pursued further the question of whether or not the quality of the categorization into topics using BERTopic is comparable to a manual procedure performed by an expert in the field. However, judging from the topic representations, and based on the authors' domain-specific knowledge, the BERTopic categorization makes sense. The topics discovered could be expected in the context of religious counselling. Moreover, substantial overlaps between internal categories of *askimam.org* and topics discovered by BERTopic point in the same direction. Still, a more detailed assessment may be necessary. It must be kept in mind that the data used to train the embedding model is general data, mainly collected from the internet. Albeit e-fatwas may have constituted part of that data, their share cannot be assumed to have been substantial. Embedding models are, however, often open source and may be fine-tuned for particular tasks, such as categorizing fatwas. To my knowledge, no such fine-tuned models for fatwas exist at the time of writing, and it goes beyond the purpose of the current article, and the current competence of the author, to construct one.

Hence, the training data may have resulted in certain biases in the underlying model, having effects on the end result in terms of categorization. Bias in LLMs, in relation to for example gender, culture, and ethnicity is the focus of ongoing research. However, bias is not a phenomenon limited to LLMs. The topic modelling performed here indicates that the *human* manual categorization performed by *askimam.org*'s administrators also may have a bias. Certain large topics identified by BERTopic do not have corresponding internal categories: most notably dream interpretations, naming practices, and sex. These major topics were either not detected, or consciously ignored, by the administrators.

It is important to note that the case used in this article is limited to 20 thousand fatwas from one single website. It is furthermore a website with a clear ideological orientation and a certain geographical bias towards South Africa. Thus, the tentative results cannot be generalized to e-fatwas at large. Whether the patterns indicated here are specific to *askimam.org* or general, remains to be seen in future research. The number of fatwas already available for computer-augmented analysis most probably counts in millions. The unified structure of question and answer, consistent over time and space, makes fatwas into ideal material for comparative analysis, and what has been done here on a relatively small collection of fatwas, can easily be scaled up. Different collections of fatwas can be compared with one another and analyzed based on, for example, the affiliations of the fatwa issuing body, or (if given) the geographical origin of the questioners. In this respect, the capacities of LLMs for multilingual embeddings allow for the incorporation of linguistically diverse material in the comparison.

The fatwas of *askimam.org* already exist in digital format, as e-fatwas. But as mentioned above, fatwa issuing, and fatwa collection are activities with a long history. As more fatwa collections from different historical periods are turned into e-fatwas, and hence become available in a form that can become objects of computer-augmented analysis, the potential for diachronic comparative studies using computer-assisted methods such as the one showcased here becomes larger, revealing perhaps new patterns in Islamic legal reasoning over the ages (considering the responses from *muftis*) but also patterns of change and continuity in the concerns of Muslims (the questions posed).

From a more general perspective, the introduction of LLMs into research, also at this early stage in the development of the technology, does hold great promises for the future. LLMs can already now fast and effectively translate and summarize text, provide comprehensive overviews of previous research, identify research gaps and suggest new areas for research, extract relevant information from texts, categorize and provide detailed descriptions of images, decipher manuscripts and transcribe speech. A lot of the tedious and repetitive work associated with humanistic research can be outsourced to systems using Artificial Intelligence. At the same time, however, and in some areas, these

new systems also challenge some of the special competencies that have been part of forming humanistic scholars' identities for centuries. Some of these competencies will inevitably become redundant as technology progresses. As the current article hopefully illustrates, at least in a rudimentary manner, the new technology also opens for research that would previously have been practically impossible to perform.

References

Agrama, Hussein Ali. 2010. "Ethics, Tradition, Authority: Toward an Anthropology of the Fatwa." *American Ethnologist* 37 (1): 2–18.

Al-Faruqi, Isma'il R. 1986. *Toward Islamic English*. Hemdon: International Institute of Islamic Thought (IIIT).

Al-Marakeby, Muhammad. 2022. "Rethinking Modern Fatwa Typology: An Ethnographic Study on al-Azhar Fatwa Council." *Islamic Studies Review* 1 (2): 197–216.

Ask Imam. 2008. "Unanswered [Fatwa #16599]," archived January 20, 2021, at the Wayback Machine, https://web.archive.org/web/20210120190537/http://www.askimam.org/public/question_detail/16599.

Awass, Omer. 2023. *Fatwa and the Making and Renewal of Islamic Law: From the Classical Period to the Present*. Cambridge: Cambridge University Press.

Banh, Leonardo, and Gero Strobel. 2023. "Generative Artificial Intelligence." *Electronic Markets* 33 (1): article 63.

Bunt, Gary R. 2003. *Islam in the Digital Age: E-jihad, Online Fatwas and Cyber Islamic Environments*. London: Pluto Press.

Bunt, Gary R. 2024. *Islamic Algorithms: Online Influence in the Muslim Metaverse*. London: Bloomsbury Academic.

Caeiro, Alexandre. 2006. “The Shifting Moral Universes of the Islamic Tradition of Ifṭā’: A Diachronic Study of Four Adab al-Fatwā Manuals.” *The Muslim World* 96 (4): 661–685.

Caeiro, Alexandre. 2013. “Transnational Ulama, European Fatwas, and Islamic Authority: A case study of the European Council for Fatwa and Research.” In *Producing Islamic knowledge*, edited by Martin van Bruinessen, and Stefano Allievi, 121–141. London: Routledge.

Copinger-Symes, Tom R. 2017. “Is Osama bin Laden’s ‘Fatwa Urging Jihad Against Americans’ Dated 23 February 1998 Justified by Islamic Law?” In *International Law and Islamic Law*, edited by Mashood A. Baderin, 213–234. Routledge.

De Angelo, Carlo. 2017. “Muslim Emigration to the West: The Jurisprudence of the Saudi Permanent Committee for Scholarly Research and Fatwas.” *Journal of Arabic and Islamic Studies* 17: 204–222.

Egger, Roman, and Joanne Yu. 2022. “A Topic Modeling Comparison Between lda, nmf, top2vec, and bertopic to Demystify Twitter Posts.” *Frontiers in sociology* 7: 886498.

Eickelman, Dale F., and James P. Piscatori. 2004. *Muslim Politics*. Princeton, NJ: Princeton University Press.

Golub, Koraljka. 2020. “Automatic Identification of Topics: Applications and Challenges.” In *Doing Digital Humanities: Concepts, Approaches, Cases*, edited by Joacim Hansson, and Jonas Svensson, 5–26. Växjö: Linnaeus University Press.

Gräf, Bettina, (ed.). 2009. *Global Mufti: The Phenomenon of Yūsuf al-Qaradāwī*. Columbia University Press.

Grootendorst, Maarten. 2022. “BERTopic: Neural Topic Modeling With a Class-Based TF-IDF Procedure.” *arXiv preprint arXiv:2203.05794*.

Hamzah, Dyala. 2005. “Is There an Arab Public Sphere? The Palestinian Intifada, a Saudi Fatwa and the Egyptian Press.” In *Religion, Social Practice, and Contested Hegemonies: Reconstructing the Public Sphere in Muslim Majority Societies*, edited by Armando Salvatore, and Mark LeVine, 181–206. Springer.

Hugging Face. n.d. “sentence-transformers all-MiniLM-L6-v2,” archived September 10, 2024, at the Wayback Machine, <https://web.archive.org/web/20240910143344/https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2>.

Ibahrine, Mohammed. 2020. “Digitally Researching Islam.” In *Second International Handbook of Internet Research*, edited by Jeremy Hunsinger, and Lisbeth Klastrup, 785–798. Amsterdam: Springer

Illar, Derek John. 2010. “Cyber Fatwās and Classical Islamic Jurisprudence.” *UIC John Marshall Journal of Information Technology and Privacy Law* 27 (4): 577–582.

Ingram, Brannon D. 2018. *Revival from Below: The Deoband Movement and Global Islam*. Oakland: University of California Press.

Iqbal, Sarosh, Rubeena Zakar, Florian Fischer, and Muhammad Zakria Zakar. 2022. “Consanguineous Marriages and Their Association With Women’s Reproductive Health and Fertility Behavior in Pakistan: Secondary Data Analysis From Demographic and Health Surveys, 1990–2018.” *BMC Women’s Health* 22 (1): 118.

Ismail, Farhana. 2015. “An Analysis of the Discursive Representations of Women’s Sexual Agency in Online Fatwas: A Case Study of Askimam.org.” Masters degree thesis, University of KwaZulu-Natal.

- Kemper, Michael. 2022. "Ijtihād in Putin's Russia? Signature Fatwas from Moscow and Kazan." *Journal of the Economic and Social History of the Orient* 65 (7): 935–960.
- Kinberg, Leah. 2012. "Dreams Online: Contemporary Appearances of the Prophet in Dreams." In *Dreams and visions in Islamic societies*, edited by Felek Özgen, and Alexander D. Knysh, 139–157. New York: SUNY.
- Kinberg, Leah. n.d. "Dreams." In *Encyclopaedia of Islam, THREE* [Online version]. Leiden: Brill.
- Kutscher, Jens. 2009. "The Politics of Virtual Fatwa Counseling in the 21st Century." *Masaryk University Journal of Law and Technology* 3 (1): 33–49.
- Kutscher, Jens. 2011a. "Islamic Shura, Democracy, and Online Fatwas." *CyberOrient* 5 (2): 50–72.
- Kutscher, Jens. 2011b. "Towards a Solution Concerning Female Genital Mutilation? An Approach From Within According to Islamic Legal Opinions." *Scripta Instituti Donneriani Aboensis* 23: 216–236.
- Larsen, Lena. 2016. "Islamic Jurisprudence and Transnational Flows: Exploring the European Council for Fatwa and Research." In *From Transnational Relations to Transnational Laws*, edited by Anne Hellum, Shaheen Sardar Ali, and Anne Griffiths, 139–163. London: Routledge.
- Larsen, Lena. 2018. *How Muftis Think: Islamic Legal Thought and Muslim Women in Western Europe*. Leiden: Brill.
- Larsson, Göran. 2011. "Islam and Tattooing: An Old Question, a New Research Topic." *Scripta Instituti Donneriani Aboensis* 23: 237–256.
- Manning, Christopher D. 2022. "Human Language Understanding and Reasoning." *Daedalus* 151 (2): 127–138.

Maravia, Usman, Zhazira Bekzhanova, Mansur Ali, and Rakan Alibri. 2021. "British Muslims Caught Amidst Fogs: A Discourse Analysis of Religious Advice and Authority." *Religions* 12 (2): 140.

Marcotte, Roxanne D. 2016. "Fatwa Online: Novel Patterns of Production and Consumption." In *Political Islam and Global Media*, edited by Noha Mellor, and Khalil Rinnawi, 231–245. London: Routledge.

Masud, Muhammad Khalid, Brinkley Morris Messick, and David Stephan Powers, (eds.). 1996a. *Islamic Legal Interpretation: Muftis and Their Fatwas, Harvard Studies in Islamic Law*. Cambridge: Harvard University Press.

Masud, Muhammad Khalid, Brinkley Morris Messick, and David Stephan Powers, (eds.). 1996b. "Muftis, Fatwas, and Islamic Legal Interpretation." In *Islamic Legal Interpretation: Muftis and Their Fatwas*, edited by Muhammad Khalid Masud, Brinkley Morris Messick, and David Stephan Powers, 3–32. Cambridge: Harvard University Press.

McDonald, Zahraa. 2014. "Religious Freedom Within a Liberal Constitution: An Overview of Muslim Personal Law in South Africa." *Politikon* 41 (2): 209–226.

Metcalf, Barbara Daly. 1982. *Islamic Revival in British India : Deoband, 1860–1900*. Princeton: Princeton University Press.

Meyer, Jesse G., Ryan J. Urbanowicz, Patrick C. N. Martin, Karen O'Connor, Ruowang Li, Pei-Chen Peng, Tiffani J Bright, Nicholas Tatonetti, Kyoung Jae Won, and Graciela Gonzalez-Hernandez. 2023. "ChatGPT and Large Language Models in Academia: Opportunities and Challenges." *BioData Mining* 16 (1): 20.

Mishra, Smeeta, and Gaby Semaan. 2010. "Islam in Cyberspace: South Asian Muslims in America Log In." *Journal of Broadcasting and Electronic Media* 54 (1): 87–101.

Moretti, Franco. 2013. *Distant Reading*. London: Verso Books.

Munshi, Amr Abdullah, Wesam Hasan AlSabban, Abdullah Tarek Farag, Omar Essam Rakha, Ahmad Al Sallab, and Majid Alotaibi. 2022. “Automated Islamic Jurisprudential Legal Opinions Generation Using Artificial Intelligence.” *Pertanika Journal of Science and Technology* 30 (2): 1135–1156.

Naveed, Humza, Asad Ullah Khan, Shi Qiu, Muhammad Saqib, Saeed Anwar, Muhammad Usman, Nick Barnes, and Ajmal Mian. 2023. “A Comprehensive Overview of Large Language Models.” *arXiv preprint arXiv:2307.06435*.

Noralla, Nora. 2023. “Gender Trouble in the Land of the Nile: Transgender Identities, the Judiciary and Islam in Egypt.” *Yearbook of Islamic and Middle Eastern Law Online* 1 (aop): 1–37.

Olsson, Susanne. 2017. “Shia as Internal Others: A Salafi Rejection of the ‘Rejecters.’” *Islam and Christian–Muslim Relations* 28 (4): 409–430.

Pasuni, Afif. 2018. “Negotiating Statist Islam: Fatwa and State Policy in Singapore.” *Journal of Current Southeast Asian Affairs* 37 (1): 57–88.

Potter, Jonathan, and Chloe Shaw. 2018. “The Virtues of Naturalistic Data.” In *The SAGE handbook of qualitative data collection*, edited by Uwe Flick, 182–199. London: SAGE.

Ragab, Ahmed. 2023. “‘And God Knows Best’: Knowledge, Expertise, and Trust in the Postcolonial Web-Sphere.” In *Critical Approaches to Science and Religion*, edited by Myrna Perez Sheldon, Ahmed Ragab, and Terence Keel, 181–202. New York: Columbia University Press.

Rusli, Rusli, and Nurdin Nurdin. 2022. “Understanding Indonesia Millennia Ulama Online Knowledge Acquisition and Use in Daily Fatwa Making

Habits.” *Education and Information Technologies* 27 (3): 4117–4140.

Schussman, Aviva. 1998. “The Legitimacy and Nature of Mawlid al-Nabi:(Analysis of a Fatwa).” *Islamic Law and Society* 5 (2): 214–234.

Singer, Julia. 2021. “Fatwas from Islamweb. Net on Robotics and Artificial Intelligence.” In *Artificial Intelligence in the Gulf: Challenges and Opportunities*, edited by Elie Azar, and Anthony N. Haddad, 279–301. Singapore: Palgrave Macmillan.

Šisler, Vit. 2009. “European Courts’ Authority Contested? The Case of Marriage and Divorce Fatwas On-line.” *Masaryk University Journal of Law and Technology* 3 (1): 51–78.

Šisler, Vit. 2011a. “Cyber Counsellors: Online Fatwas, Arbitration Tribunals and the Construction of Muslim Identity in the UK.” *Information, Communication and Society* 14 (8): 1136–1159.

Šisler, Vit. 2011b. “The Internet, New Media, and Islam: Production of Islamic knowledge and Construction of Muslim Identity in the Digital Age.” PhD Thesis, Univerzita Karlova v Praze.

Skovgaard-Petersen, Jakob. 1997. *Defining Islam for the Egyptian State: Muftis and Fatwas of the Dār al-iftā*. Leiden: Brill.

Svensson, Jonas. 2021. “Minding the Pandemic : A CSR Perspective on Patterns in Muslim Religious Responses to COVID-19.” *Approaching Religion* 11 (2): 23–44.

Usha, Sanyal. n.d. “Barelwīs.” In *Encyclopaedia of Islam, THREE* [Online edition]. Leiden: Brill.

Van den Branden, Stef, and Bert Broeckaert. 2011. “Living in the Hands of God. English Sunni E-Fatwas on (Non-) Voluntary Euthanasia and Assisted Suicide.” *Medicine, Health Care and Philosophy* 14: 29–41.

Weimann, Gabriel. 2011. "Cyber-Fatwas and Terrorism." *Studies in Conflict and Terrorism* 34 (10): 765–781.

Weiss, Bernard G. 1996. "Ibn Taymiyya on Leadership in the Ritual Prayer." In *Islamic Legal Interpretation: Muftis and Their Fatwas*, edited by Muhammad Khalid Masud, Brinkley Morris Messick, and David Stephan Powers, 63–71. Cambridge, MA: Harvard University Press.

Whyte, Shaheen. 2022. "Are Fatwas Dispensable? Examining the Contemporary Relevance and Authority of Fatwas in Australia." *Oxford Journal of Law and Religion* 11 (2–3): 314–342.

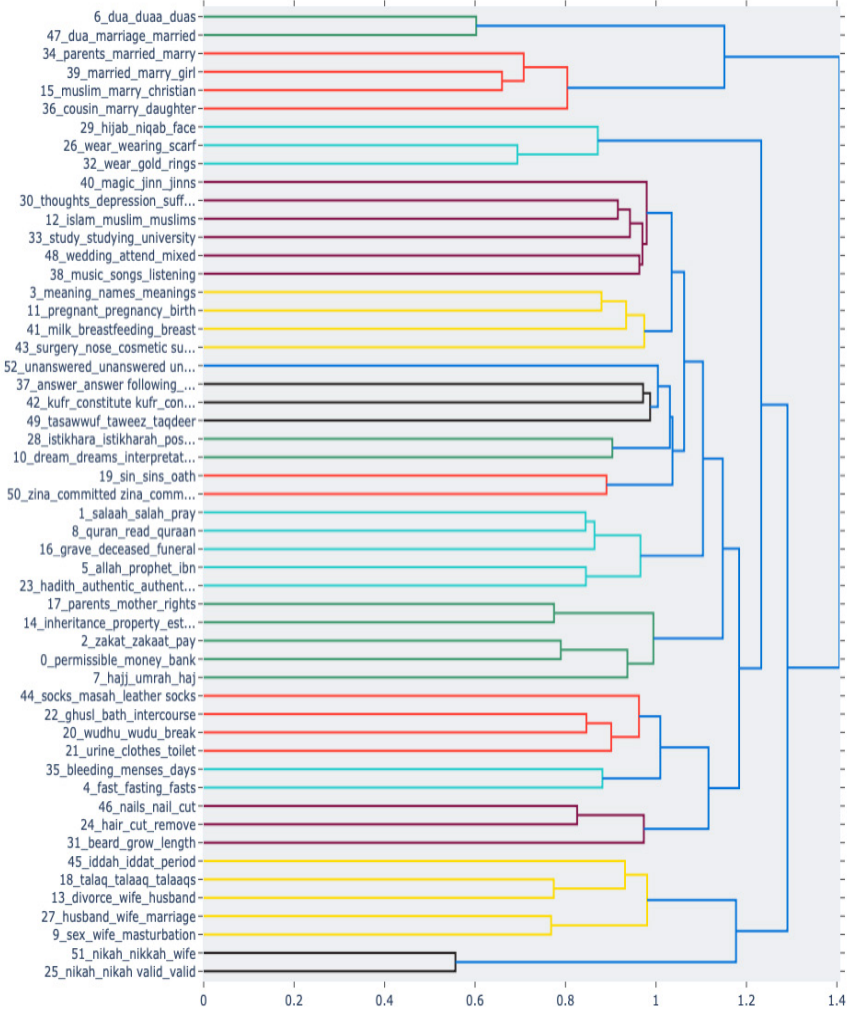
World Bank Group. 2024. "Individuals using the Internet (% of population)," archived September 11, 2024, at the Wayback Machine, <https://web.archive.org/web/20240911063458/https://data.worldbank.org/indicator/IT.NET.USER.ZS>.

Wynn, L. L., and Angel M. Foster. 2018. "Muftis in the Matrix: Comparing Online English-and Arabic-Language Fatwas about Emergency Contraception." *Journal of Middle East Women's Studies* 14 (3): 314–332.

Yakar, Emine Enise. 2022. "Women's Political Leadership: One Question and Two Divergent fatwās." *Journal of Law and Religion* 37 (2): 334–364.

Appendix 1: Dendrogram of topics produced by BERTopic

Hierarchical Clustering



Appendix 2: Table of mergers

Original internal categories	Mergers
Misc. Fiqh Jurisprudence and Rulings (<i>Fiqh</i>) Miscellaneous Other Daily Unknown	Miscellaneous
Beliefs and Practices (<i>Aqeedah</i>) Basic Tenets of Faith	Beliefs and Practices (<i>Aqeedah</i>)
Prayer (<i>Salaat</i>) Prayer	Prayer
Character and Morals Character, Morals	Character, Morals (<i>Akhlaaq</i>)
Divorce (<i>Talaaq</i>) Talaq	Divorce (<i>Talaaq</i>)
Ramadan Ramadhan	Ramadan
Hadith Hadith and Quraan (selected)	Hadith
Quraan Hadith and Quraan (selected)	Quraan
Women Men	Gender
Purity (<i>Tahaarah</i>) Purity Wudu	Purity (<i>Tahaarah</i>)
History History and Biography	History and Biography
Funeral Funerals (<i>Janaaiz</i>)	Funerals (<i>Janaaiz</i>)
Fasting (<i>Sawm</i>) Fasting	Fasting (<i>Sawm</i>)
Masjid and Waqf I'tikaf Tarawih	Masjid