Int. J. Nonlinear Anal. Appl.
Volume 12, Special Issue, Winter and Spring 2021, 981-996
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/IJNAA.2021.5545

# Visualizing sharia law: an information retrieval approach based on key extraction algorithm

Khyrina Airin Fariza Abu Samah[a,*], Raseeda Hamzah[b], Nur Nabilah Abu Mangshor[a], Lala Septem Riza[c], Norah Md Noor[d]

[a]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Melaka Kampus Jasin, Melaka, Malaysia.*
[b]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia.*
[c]*Department of Computer Science Education, Universitas Pendidikan Indonesia, Bandung, Indonesia.*
[d]*Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, Skudai, Johor Malaysia.*

*(Communicated by Madjid Eshaghi Gordji)*

**Abstract**

Islamic jurisprudence as it relates to every element of Muslim life that founds in Sharia. Many interpretations exist for Sharia, all based on differing Islamic schools of thought and multiple laws concerning Sharia. In Malaysia, there are 135 sections of Sharia law related to family and marriage in Act 303. The survey conducted proves that Malaysians struggled when looking for sharia law and took a toll of time. Therefore, we propose a web-based Sharia Law Finder (SLF) system that solves the sharia law's information retrieval issues in Malaysia. The SLF adapted the key extraction (KE) algorithm, which is Term Frequency-Inverse Document Frequency (TF-IDF), and Rapid Automatic Keyword Extraction (RAKE). The testing for implementation of both algorithm effectiveness is using the functional and usability test. The solution enhances using visualization tools which are bubble charts and word cloud. Bubble charts visualize the related Sharia law based on users' queries and word cloud to visualize keywords used by the users. A total of 30 respondents have tested the functionality and usability of SLF. As a result, the system successfully works as specified functionality, 96.58% for the System Usability Scale, indicating the proposed solution's acceptance.

*Keywords:* Family and marriage act, RAKE algorithm, Sharia law, TF-IDF, Visualization

*Corresponding author
*Email addresses:* khyrina783@uitm.edu.my (Khyrina Airin Fariza Abu Samah ), raseeda@uitm.edu.my (Raseeda Hamzah), nurnabilah@uitm.edu.my (Nur Nabilah Abu Mangshor), lala.s.riza@upi.edu (Lala Septem Riza), norah@utm.my (Norah Md Noor)

*Received:* June 2021 *Accepted:* September 2021

## 1. Introduction

Islam grew drastically worldwide, and in the eighth century, Imam Syafie established the Syafie Madhab of Islamic law. Muslims are dominant in Syria, Lebanon, Pakistan, Indonesia, and even Malaysia [16]. Then sharia law was introduced in Malaysia to cover everything Islamic-based. According to Islamic Sharia, the whole of God's commandments and exhortations is called Sharia. It applies to all elements of human behavior in Islamic society to control and direct people on the way to eternal salvation [29].

Sharia is a collection of rules revealed to man by God for guidance in communication with God, each other, the world, and everyday life. Islamic law has features that distinguish it from other legal systems worldwide [35]. Whenever issues related to Sharia (or Islamic law), the first image that most people usually get is a violent death penalty. Despite Sharia being more than that, and like water in a desert, Sharia is more a source of redemption in temporal life and Hereafter [17]. There is no isolation on law and morality in Islam due to intermixed and automatically defines good deeds once it firmly believes the religion. Islam is a perfect combination of faith and practice. Both elements reflect law and morality, as specified in and approved by God in sharia law [35, 12].

The Quran is the backbone of sharia law. It is elaborated firstly by the Prophet's sunnah and secondly by the ijtihad. The role of sharia law has become a contested topic around the world. As an Islamic country, Malaysia applies sharia law in its legal law system to guide Muslims. Following the death of Prophet Mohammed in 632 CE, Sharia, or Islamic law, began to be employed as the Islamic empire expanded in both directions, towards the Atlantic and the Pacific Oceans. Muslims see the Prophet Mohammed's life and saying as a template for their behavior, and what is called hadith by experts is growing [21]. It has its position in Malaysia as Malaysia has a specific sharia law institution to deal with sharia-related issues. Sharia law plays an essential part in Malaysia's Islamic law matters [23].

For many years, the headline of news often capturing the sharia law's implementation [33]. The vast majority believed Islam of Islamic scholars to be a religion and a state, meaning that Islam should control government and public life while serving as a religion. However, some voices started to argue that Islam only serves as faith and should not govern; these ideas have been unparalleled in Islamic history [4]. Nevertheless, Malaysia has a strong base of Islamic institutions that made the application of sharia law in the country doable and practicable by the Malaysian Muslims.

Many people are researching sharia law but lack specific information relating to their cases. They lack information, and another problem arises as overloaded information around the internet. Thus, they are questioning the most relevant laws are for the cases they are currently working on. The consequence of these problems is that they have to struggle to look for the right sharia laws, which took a toll on them. A survey was conducted randomly on 41 respondents to support the evidence that 90.2% of them struggle when searching for any sharia law online. Another input stated that 84.55% of respondents spend much time searching or finding sharia law offline through law books or other resources when struggling to use online resources. Das and Jadab [10] claimed that law students face difficulties accessing online legal information in electronic environments. This problem led to insufficient sharia law that they can correlate to any cases due to time limitations.

This study uses the information retrieval (IR) method to retrieve any terms searched by the users. The information gathered is a problem-solving discipline concerned with the effective and efficient flow of needed information between the human generator and the user [5]. One of the techniques is Term Frequency-Inverse Document Frequency (TF-IDF), where the user's inputs match the most frequent words in the document. The most frequent words are then become the output. In comparison, Rapid Automatic Keyword Extraction (RAKE) implements by extracting and ranking the keywords out

of a document without any other context except for the document itself [7]. The result from the query of the user will then be displayed. This research also focuses on the family and marriage acts, consisting of 135 acts as the research study.

## 2. Related works

In this section, we briefly describe the key extraction algorithm and information retrieval.

### *2.1. Information retrieval*

For centuries the increase in text data such as books and papers in libraries has placed efficient mechanisms for finding them [13]. Searching for and retrieving information is a critical aspect of today's technological world. Recovery of information is extracting information from a large set of records important to the user's need. It is the work of defining models and systems that permit access to a collection of electronic documents (corpus) so that the user may find the contents that meet their information demands.

The web's arrival provided a historic disjunct to IR science and technology. Pre-Web, true IR was isolated, expensive, and restricted to certain, curated repositories. The document information's retrieving process shows in Figure 1.

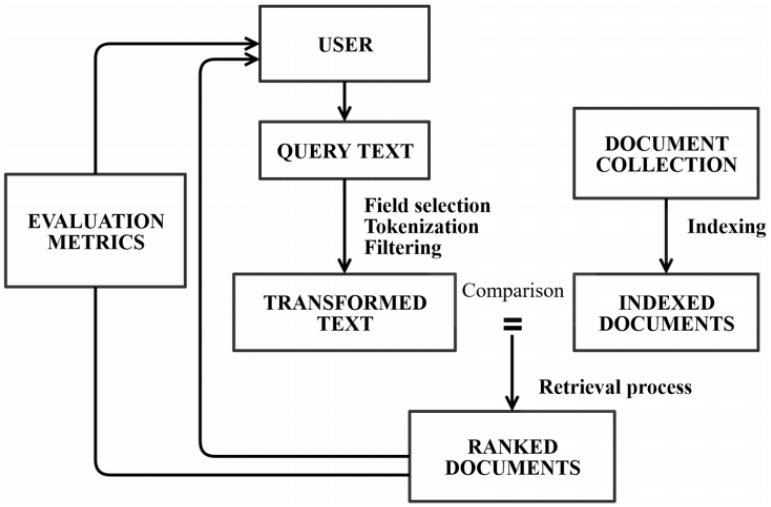



Figure 1: Process of retrieving information (Source: Juan & Luc´ıa, 2018)

Like search engines, the system retrieval for full-text enables the document's access information beyond document titles, authors, and other high-level metadata. In unstructured documents, such as those lacking relational database tables' characteristics give a defined structure, searching for information is more complex and time-consuming [30]. There are several techniques to retrieve information, such as Text-based, Multimedia-based and, Semantic web information retrieval systems.

#### *2.1.1. Text-based*

A text-based recovery program can assist in the collection's retrieval of text-based information. Some media rely on analog technologies that complicate storage in the media storage accessible to the public, the project's intended beneficiaries. According to [1], question answering text-based is considered a difficult task, and using deep learning techniques to overcome the problem. It also provides a versatile text-based IR framework for connecting and executing single IR components.

It visually describes complicated IR processes and displaying the retrieved information using user-defined styles [2].

*2.1.2. Semantic web*

It is a Web extension where information has a predetermined format definition by expressing the data formally and clearly in a standard manner. Information representation has become more formal with the advent of Semantic Web technologies [36]. Ontology is one of the Semantic Web technologies methods that many researchers have made to use ontology (knowledge base) to improve the summary process. Most online documents are related to the same subject are debated [28]. Ontology is s formalized group of conceptualization. It gives the growing definition for the specific word and its relationship to the other terms [20]. The user will annotate documents that contain the key terms to help the search engine [8].

*2.1.3. Sentence-based*

Retrieving sentences relevant to a given user's need is a discussed question from different perspectives in the literature. Moreover, there is still an appropriate solution need to discover. Sentence retrieval requires a set of documents and a textual query from the user, providing a graded collection of sentences that satisfy the user's needs [1]. The role of sentence-based retrieval is to locate appropriate sentences in response to a question from a text-based. The recovery of sentences identifies novelties, answer questions, sum up, and mining opinions [11].

*2.1.4. Comparison of information retrieval techniques*

This section reviews the differences between the mentioned techniques. Table 1 summarizes the techniques comparisons.

Table 1: The IR techniques comparison

| **Algorithm** | **Advantage(s)** | **Disadvantage(s)** |
|---|---|---|
| Text-Based | • It is widely used in various network applications, including digital libraries, e-commerce websites, and information technology management [37]. | • User-issued queries themselves will potentially compromise privacy, posing a significant threat to user privacy [37]. |
| Semantic Web | • Automate tasks that need a conceptual understanding of the items or the task itself, and automatically allow software programs to identify and combine information and resources [37]. | • Lack of expertise when searching for heterogeneous sources of information. Lack of system usability, especially at the query level, requires users to describe their information demands sophisticated languages or interfaces. Also, not having ranking algorithms for large-scale sources of information [14]. |
| Sentences-Based | • It can simplify end-user tasks of extracting the correct information from document collections in several ways [11]. | • Unable to process question answering inputted by the users [11]. |

In this study, the sentence-based choice can simplify the user's input by extracting the correct information from the documents' bundle. Moreover, users can use multiple inputs to search for their needs and provide the most relevant information based on the queries. Thus, an application platform is needed to operate the retrieving processes.

*2.2. Key extraction algorithm*

Recent advances in computing and technology have led to an ever-increasing document set. The document collection is categorizing into forms. It is expedient for decision-making to put down relevant documents together. Scientists undertaking interdisciplinary work build collections covering different topics [6]. Keyword extraction (KE) helps simplify searching related keywords and sentences in the unstructured text, such as emails and social media posts. There are different approaches to KE: Simple Statistics Approaches, Rule-Based Linguistic Approaches, and Machine Learning [34]. Hence, KE considers as a crucial method for analyzing data. The primary task of essential KE extracts a set of words or keywords highlighting the main content of documents [18]. Figure 2 shows how the KE algorithm works.

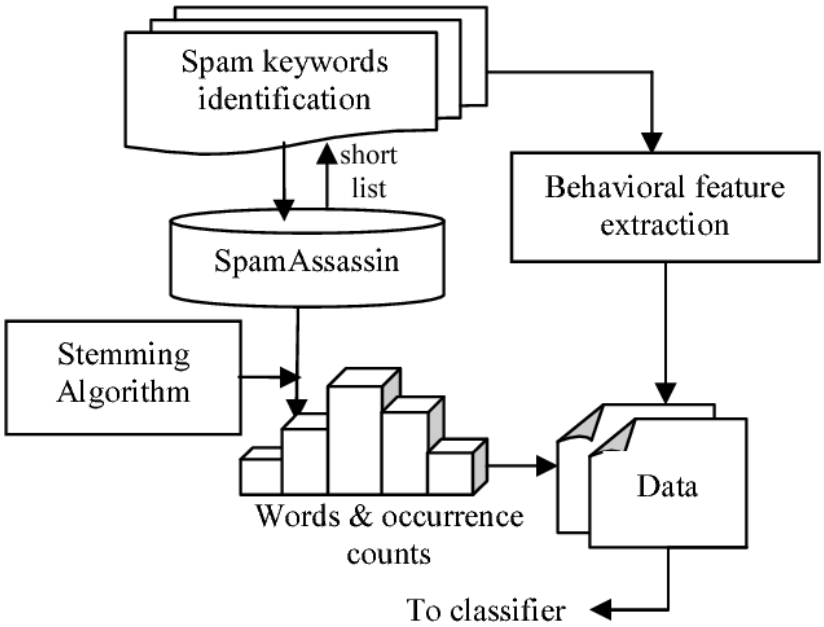



Figure 2: Illustration of keyword extraction algorithm (Source: Bhat, Malkani, Shenoy, Venugopal & Patnaik, 2011)

KE in a document provides important details about the document's content. They can aid with information gathering or document reading more efficiently. Various language processing tasks use KE: to categorize text, retrieve information [25], and cluster documents [26]. It has spawned two distinct and successful methods, which are KE unsupervised and KE supervised. It starts with extract a candidate list of keywords from heuristics-based documents in the unsupervised KE. Then, it grades the list the same as TF-IDF. Finally, prunes the graded collection to keep valid keywords free of syntactic and semantic problems [24].

On the other hand, supervised KE is primarily reliant on annotated training for corpora's text. However, most existing approaches used supervised methods like Naive Bayes to select relevant keywords [9]. There are two types of KE algorithms adapted in this study: TF-IDF and RAKE. TF-IDF is used widely in information retrieving [32]and can retrieve the sharia law that the user requested based on the related documents. RAKE is a language-independent python method.

*2.2.1. Term frequency-inverse document frequency*

Use in measuring term's total frequency appears in the whole text. The algorithm treats all keywords equally even if they have different importance while measuring a document's term frequency, regardless of incorrect stop words such as "of" [31]. Despite [22] highlights the technique's unable

to make a connection between words, [3] claimed the TF-IDF is widely using in IR, and the metric used can determine the relevant keywords. The TF-IDF calculate the term t with few conditions: 1) appears more than few times and get a higher value, 2) appears few times in a single document or various documents will get a lower value, and 3) the terms t appears in almost all documents will get a lower value.

Let $D$ represents the documents' comprehensive collection, and $t$ is a term gathered from the documents. It computes the term frequency-inverse document in Equation 1:

$$tf(w) = tf - idf(t, f, D) = tf(t, d) * idf(t, D) \qquad (1)$$

Then, the $tf(t, d)$ embodies the term's $t$ frequency in document d (the number of instances of a term in a document), conveyed by Equation 2:

$$tf(t, d) = \frac{f(t, d)}{|d|} \qquad (2)$$

The $f(t, d)$ is the $t$'s total appears frequency in document $d$, where the numerator is the dimension of the $d$, given in terms of its cardinality. Equation 3 describes the inverse document frequency $idf(t, D)$.

$$idf(t, D) = log\frac{|D|}{|\{d|t \ \epsilon \ d\}|} \qquad (3)$$

where, the denominator is the number of documents that contain the term $t$.

*2.2.2. Rapid automatic keyword extraction*

Rapid automatic keyword extraction (RAKE) collects data automatically from a collection of terms and illustrates a document overview. The term defines how the keyword defines a particular document's core information [18]. RAKE can recognize a limited collection of words, key phrases, keywords, or key parts from a textual document, even though [15] claimed the algorithm has the lowest precision. Hence, extracting keywords from documents is now considered one of the main technologies of all automated document processing [19]. Equation 4 shows the RAKE calculation.

$$word\ score = \frac{degree\ of\ wordnum}{frequency\ of\ word} \qquad (4)$$

The term "score" has the term "degree" in the numerator and the term "frequency" in the denominator. In other words, the word score is inversely related to the degree. The degree is high when a term frequently occurs, especially in long candidates. The frequency is high because a word always appears. The RAKE word score methodology favors terms that occur mostly in longer keywords and distorts those in short ones.

## 3. Research method

Figure 3 shows the research flow for Malaysia's sharia law information retrieval using TF-IDF and RAKE algorithm, consisting of four main steps: system design, data preparation, algorithm calculation, and visualization implementation.

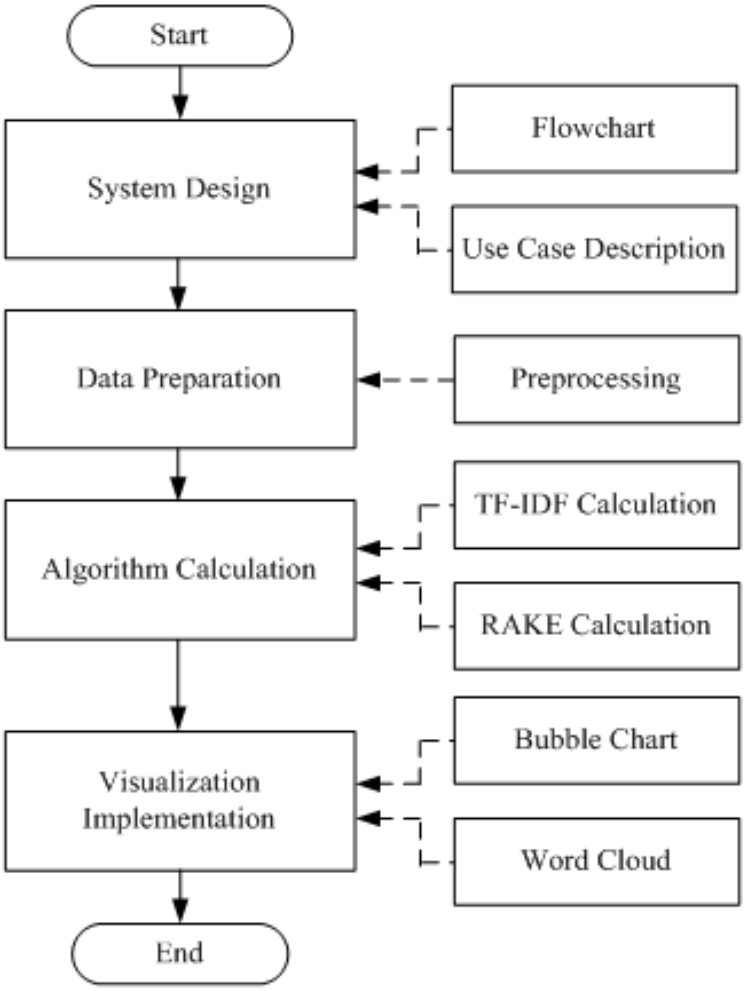



Figure 3: Research flow for Shariah Law Finder information retrieval

### *3.1. System design*

A well-design system can help in developing the system without technical errors in future development. The design phase breaks into several processes to improve the TF-IDF and RAKE algorithms' performance: the flow chart and use case description. We use the Malay language for system web development because to fulfill the main objective of this research is to overcome the difficulties of finding relevant acts.

#### *3.1.1. Flow chart of shariah law finder*

Figure 4 shows the flowchart that represents the general flow of how the SLF works. Suppose the user wants to search for any related Sharia law. In that case, the user will have to input a keyword, and the keyword will go through pre-processing to calculate TF-IDF and RAKE algorithm. The search result will be in the form of a table accordingly. The visualizations used are bubble charts for displaying related sharia law and word cloud charts for displaying every keyword used in the system. Users can search again by clicking on the search icon. If the user is not going to search again, the system will stop.

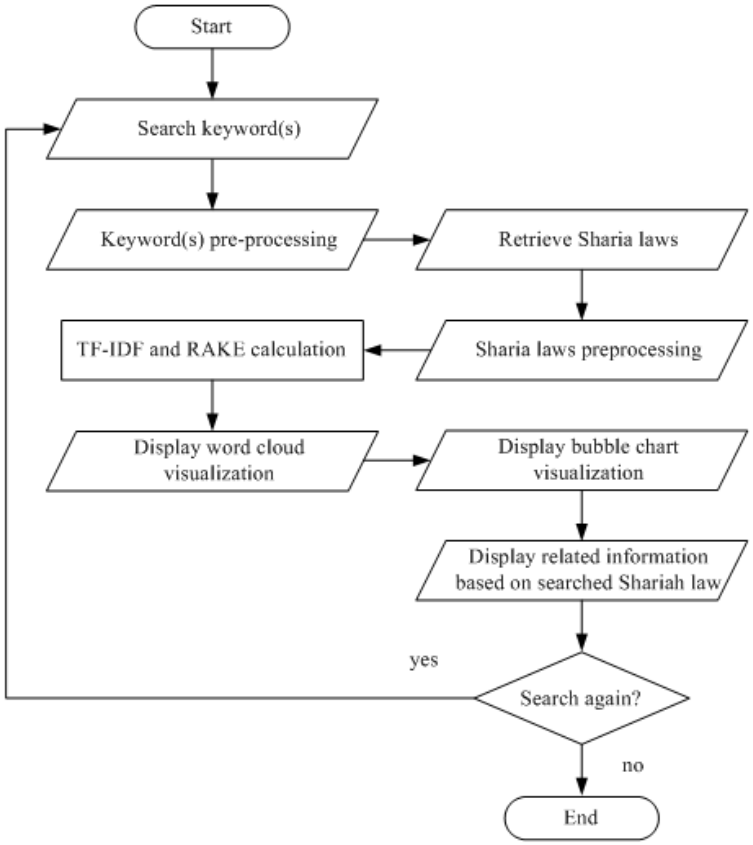



Figure 4: Flowchart for Shariah Law Finder information retrieval

*3.1.2. Use case description of shariah law finder*

Table 2 shows the description of the use case for the SLF system. The functionality test is tested according to each use case to ensure the system functioning without any error.

Table 2: Use case description for Shariah Law Finder information retrieval

| No | Use Case | Description |
|---|---|---|
| 1 | Enter keyword(s) | User enter keyword(s) to search |
| 2 | Click the "Search" button | User click the search button to find the desired keyword(s) |
| 3 | View the word cloud | User view the word cloud of the search keyword(s) |
| 4 | View the bubble chart | User view the bubble chart of the search keyword(s) using TF-IDF and RAKE and test mouseover to view the details |
| 5 | View IR of Sharia law | Related sharia law is displayed |

*3.2. Data preparation*

The data of the Sharia law is scrap from the eSyariah official website using Octoparse. The file scrapped were then checked with www.jsonlint.com to make it JSON and then stored in JSON format. Each of the data consists of its title and contain. JSON is selected since it is a lightweight file format, human and machine-readable, and loads fast in JavaScript. The system will read the

whole JSON file and store them in an array when the user starts to key in the keywords. This list will go through the previous data pre-processing and pass to the TF-IDF and RAKE function for calculation processes.

Data pre-processing or data cleaning is performing at the users' query string and the retrieved Sharia laws. There are several processes in data pre-processing, and they are quite different for each of the algorithms. In TF-IDF, the first step is the tokenization process, in which the long strings of text from sharia law are chunking into tokens of words. Next, special characters and whitespace are removed and converted to lower case. It is compulsory as it will increase the accuracy of the result later on. Then, a process of removal of stop words considers as complete. The stop words are removing because it has no relevance in this algorithm. They will not be matched up in the matching process later. For the RAKE algorithm, the only difference is no removal of stop words.

### *3.3. Algorithm calculation*

For both algorithms, every Sharia law exists considered as a document. The steps and calculation are needed for every query made by a user and describe as follow:

i. The user inputs their query in the "Key-in Keyword" field.

ii. The user clicks the "Search" button.

#### *3.3.1. TF-IDF calculation*

For this system to find related sharia law, TF-IDF calculation performs to find every word's relevance score. In a document, there exists a title and its body—both needed for TF-IDF calculation. Then the value is total up and considered as the actual TF-IDF score for the document. Using the formula in Equation 1 and followed by Equation 2 as mentioned before, the calculation starts. The final step is to match the TF-IDF score with the query to get the relevance score.

#### *3.3.2. RAKE calculation*

The calculation for the RAKE algorithm uses to find the most scored terms from the documents. The word degree is in the numerator, and the word frequency is in the denominator. In other words, the word score is inversely related to the degree. The degree is high when a term frequently occurs, especially in long candidates. The frequency is high because a word always appears. The RAKE word score methodology biases the terms that appear frequently but not in long keywords.

### *3.4. Visualization implementation*

There are two types of visualization implemented in this system: bubble chart and word cloud. Bubble chart used to visualize the related Sharia law based on users' queries. It indicates the relevance score by the bubbles' size. The higher the relevance score, the bigger the bubble size and word cloud to visualize keywords used by the users. Highcharts.js library uses to create dynamic visualization. Word cloud implemented in the system intended to visualize only the existing word searched by the user. It gives another data analysis on the most searched keyword(s) by users.

## 4. Results and discussion

This section explains the findings of Shariah Law Finder (SLF). This system is testing was tested for functionality and usability.

### *4.1. Functionality testing*

The system's functioning tests five times (Test1 through Test5) using the prototype's menu to ensure it fits the user's needs. Table 3 shows the outcome where all the functionality tests (FT) successfully run without any error.

Table 3: Functionality test result for Shariah Law Finder information retrieval

| No | Component Use Case | FT1 | FT2 | FT3 | FT4 | FT5 |
|---|---|---|---|---|---|---|
| 1 | Enter keyword(s) | Passed | Passed | Passed | Passed | Passed |
| 2 | Click the "Search" button | Passed | Passed | Passed | Passed | Passed |
| 3 | View the word cloud | Passed | Passed | Passed | Passed | Passed |
| 4 | View the bubble chart | Passed | Passed | Passed | Passed | Passed |
| 5 | View IR of Sharia law | Passed | Passed | Passed | Passed | Passed |

As in snapshot Figure 5, SLF front page demonstrates where the user can key in the desired keyword. As mentioned before, we use the Malay language for system web development to ease the searching relevant acts for Malaysians. The word "dihukum denda" or translation in English as "sentenced to a fine" was used for this functional testing. Figure 6 shows the word cloud's result for frequently searched existing words by users. As we can see, the word searched "dihukum denda" appeared in the center of the word cloud.

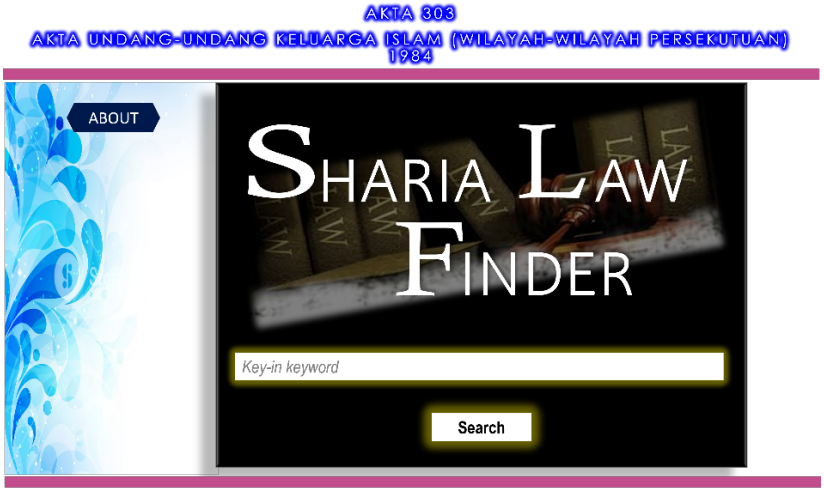



Figure 5: The front page of SLF information retriev

Next, Figure 7 shows the score of TF-IDF and RAKE, which represents using the bubble chart blue and red for the searched word. As a result, TF-IDF and RAKE algorithm found five-section related to the searched keywords with the similarity score of 264.5251 and 484.2136, respectively. The original similarity score is multiplied by 1000 for visualization purposes. The mouseover function helps to show the individual value and description of the result; the RAKE highest related section with a score value of 111.11. Based on the result, users can get the information which act number to find the information.

Users get detailed information directly by clicking the respective bullet. Table 4 shows the distribution result for the TF-IDF and RAKE score using the same keyword, "dihukum denda" or translation in English as "sentenced to a fine".


Figure 6: Word cloud for frequently searched existing words in SLF information retrieval

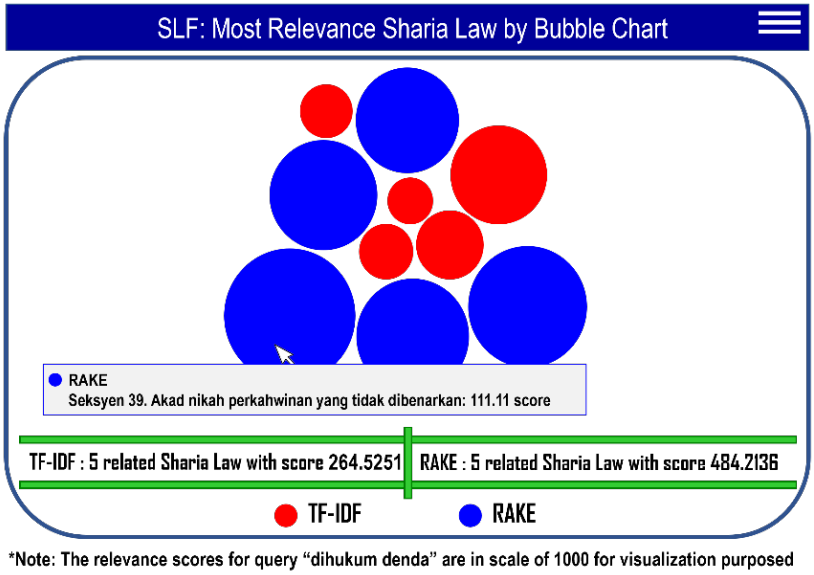



Figure 7: Bubble chart of algorithm result for keyword searched in SLF information retrieval

After completing the querying process and algorithm calculation, the SLF system retrieves the related Sharia law. Figure 8 represents the IR based on the TF-IDF algorithm, and Figure 9 represents the IR result of the RAKE algorithm. It displays in a list and organizes it from the highest relevance score to the lowest relevance score.

### *4.2. Usability testing*

For this project, the usability test performs using System Usability Scale or SUS. It is used widely as a standard questionnaire to assess perceived usability [27]. The process consists of these four steps:

1. Select 30 respondents randomly.
2. The 30 respondents will test the system using TeamViewer due to Movement Control Order during Covid-19.
3. After the test, respondents must answer the 10 SUS questionnaire using Google Form.

Table 4: TF-IDF and RAKE score and Sharia law description

| KE AgoBubble | | Score | Sharia Law |
|---|---|---|---|
| **TF-IDF** | 1 | 81.2196 | Seksyen 36.Pelanggaran terhadap seksyen 32 |
| | 2 | 56.5118 | Seksyen 129. Isteri tidak menurut perintah |
| | 3 | 45.2094 | Seksyen 131. Persetubuhan luar nikah antara orang yang bercerai |
| | 4 | 43.7098 | Seksyen 128. Tidak memberi keadilan yang sewajarnya kepada isteri |
| | 5 | 37.8745 | Seksyen 35. Tidak hadir di hadapan Pendaftar dalam masa yang ditetapkan |
| **RAKE** | 1 | 111.1100 | Seksyen 39. Akad nikah perkahwinan yang tidak dibenarkan |
| | 2 | 100.0000 | Seksyen 38. Akuan atau pernyataan palsu untuk mendapatkan perkahwinan |
| | 3 | 95.2381 | Seksyen 35. Tidak hadir di hadapan Pendaftar dalam masa yang ditetapkan |
| | 4 | 90.9091 | Seksyen 126. Meninggal langsung isteri |
| | 5 | 86.9565 | Seksyen 123. Poligami tanpa kebenaran Mahkamah |

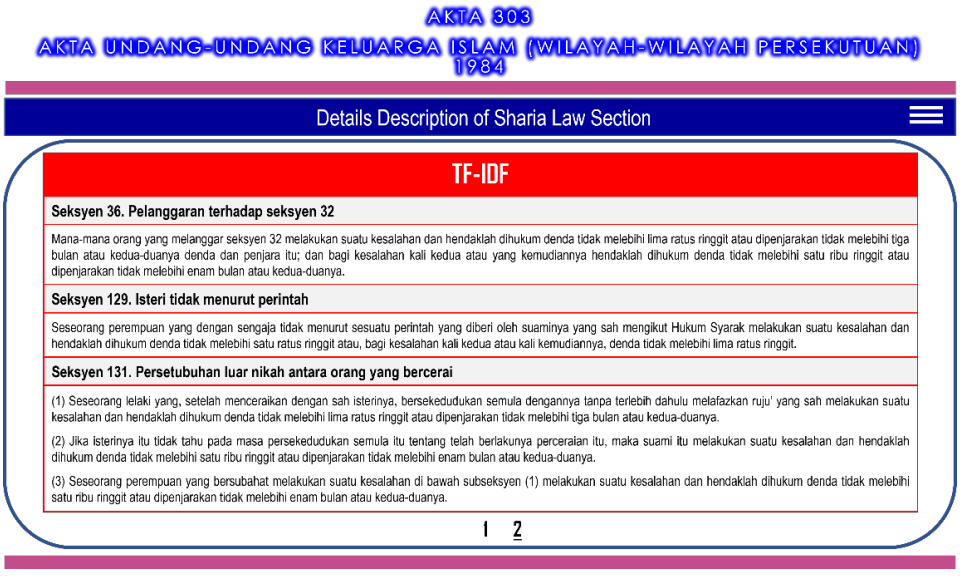



Figure 8: SLF information retrieval based on TF-IDF algorithm result

4. Analyze the SUS result.

Once completed, all the recorded responses analyze and calculates results using the computer. Figure 10 summarizes the SUS questionnaire score result on all 30 respondents for each of the multiple bar charts' questions.

The score is determined based on inputs on each component, ranging from 1 to 5, for system usability result scoring. One point deducts from the score for odd-numbered questions. Meanwhile,

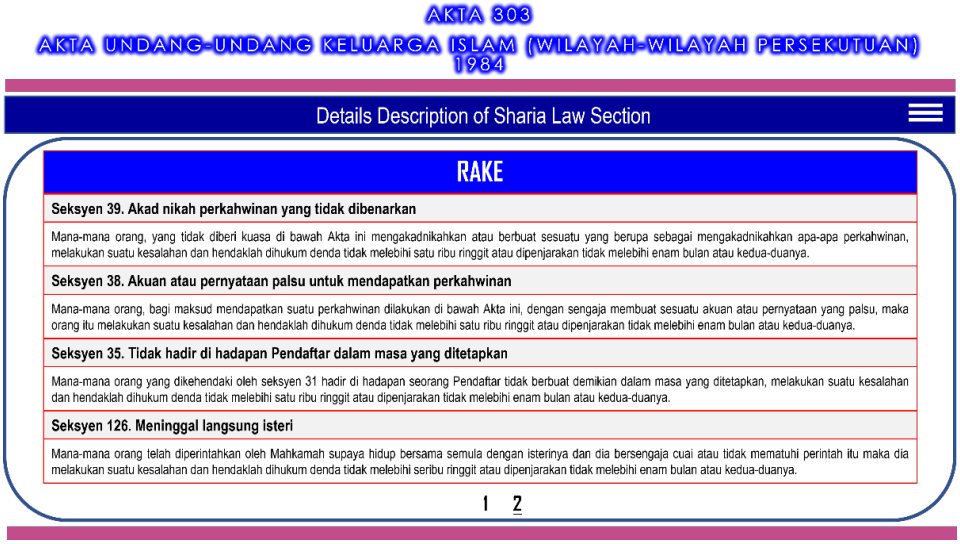
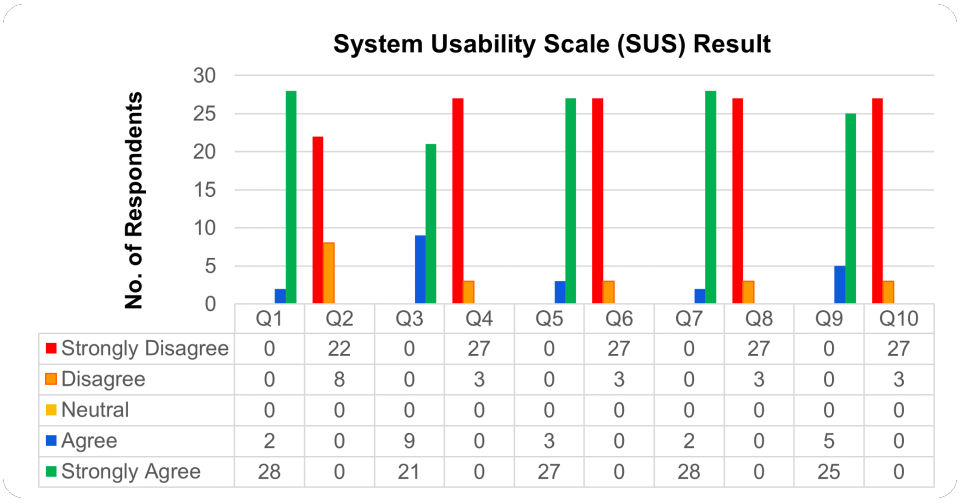



Figure 9: SLF information retrieval based on RAKE algorithm result




| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Strongly Disagree | 0 | 22 | 0 | 27 | 0 | 27 | 0 | 27 | 0 | 27 |
| Disagree | 0 | 8 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 |
| Neutral | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Agree | 2 | 0 | 9 | 0 | 3 | 0 | 2 | 0 | 5 | 0 |
| Strongly Agree | 28 | 0 | 21 | 0 | 27 | 0 | 28 | 0 | 25 | 0 |

Figure 10: SUS result by question for SLF information retrieval

subtracts the scores to 5 for even-numbered questions. It will add up all scores from each question and times with 2.5 to find the final score. The range for the result is 0 to 100. Figure 11 shows a histogram of SUS scores. The range starts at 90% and steadily increases to 92%, with a new range every 2%. The frequency for each range shows on the vertical axis.

Based on the histogram, the highest frequency is between 96% to 9%, of which seventeen respondents fall into that range. There is one respondent who scored at the lowest at 90% to 92%, two respondents gave 92% to 94% scores, and there are seven respondents who gave from 94% to 96% scores, and three respondents fell in the last range from 98% to 100% scores.

If the score of SUS reaches from 68 to 80.3, the system is rating as 'Good' according to SUS guidelines. Since this study obtained an average SUS of 96.58%, reaching and exceeding 68%, it concludes that the SLF system is very usable. It can reduce the time taken to find and view each of the Sharia law sections.

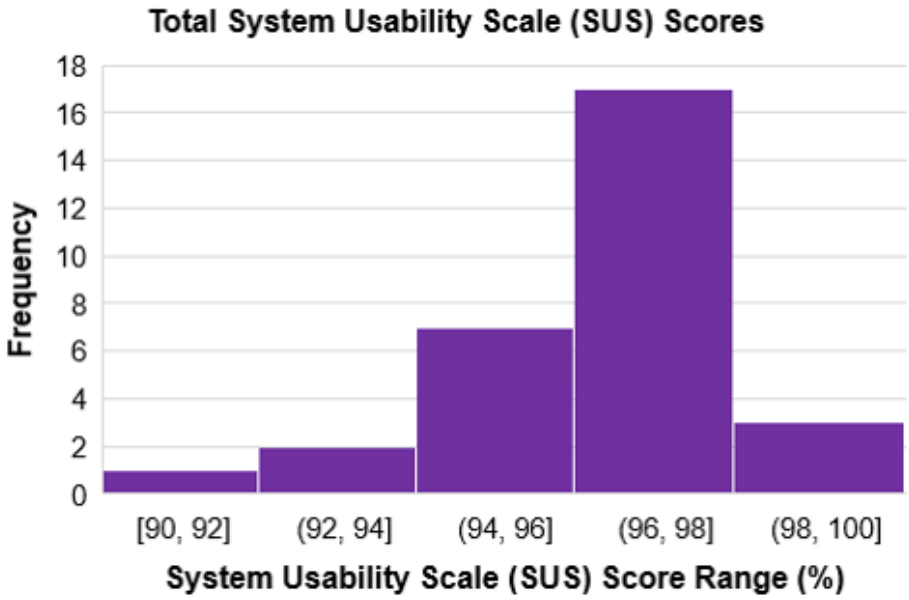



Figure 11: Histogram SUS result for SLF information retrieval

## 5. Conclusion

There are 135 sections in Act 303 under Sharia law related to family and marriage in Malaysia. Due to the time-consuming search for information related to Sharia law, this study aims to overcome the problem by developing a web-based system Sharia Law Finder (SLF). SLF is develops using the information retrieval algorithm, which is TF-IDF and RAKE. The development has three benefits: 1) users can visualize the word cloud of the most frequent keywords used, 2) the bubble chart indicates each algorithm's score, and 3) view the information retrieval on the searched keyword(s). The functionality and usability of the system successfully achieve the objective with a SUS of 96.58%. The system can cover other Acts and apply the visualization technique to ease classification or recommend the related acts for further recommendation.

## Acknowledgment

The research was sponsored by Universiti Teknologi MARA Cawangan Melaka under the TEJA Grant 2021 (GDT 2021/1-28).

## References

[1] Z. Abbasiantaeb and S. Momtazi, *Text-based question answering from information retrieval and deep neural network perspectives: A survey*, 2021.

[2] Y. Ai, R. Gerling, M. Neumann, C. Nitschke, and P. Riehmann, *TIRA: text based information retrieval architecture*, in 2nd International Workshop on Text-Based Information Retrieval (TIR), (2005) 67–75.

[3] M. T. Artese and I. Gagliardi, *What is this painting about? experiments on unsupervised keyphrases extraction algorithms*, IOP Conference Series: Materials Science and Engineering, 364 (2018) 1–8.

[4] Y. Auf, *Islam and sharia law: Historical, constitutional, and political context in Egypt*, 2016.

[5] V. Balakrishnan and L. Ethel, *Stemming and lemmatization: a comparison of retrieval performances*, Lecture Notes on Software Engineering, 2 (3) (2014) 262–267.

[6] P. Bafna, P. Dhanya, and A. Vaidya, *Document clustering: TF-IDF approach,* in International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), (2016) 61–66.

[7] M. W. Berry and J. Kogan, *Text mining: applications and theory*. Wiley, 2010.

[8] C. Biancalana, F. Gasparetti, A. Micarelli, and G. Sansonetti, *Social semantic query expansion*, in ACM Transactions on Intelligent Systems and Technology, (2013) 1–43.
[9] R. Campos, V. Mangaravite, A. Pasquali, A. M. Jorge, C. Nunes, and A. Jatowt, *Yake! collection-independent automatic keyword extractor*, in European Conference on Information Retrieval, (2018) 806–810.
[10] R. K. Das and A. Jadab, *Patterns of information seeking behaviour of law students in digital environment: a study*, Journal of Information Science Theory and Practice, 5 (1)(2017) 15–25.
[11] A. Doko, M. Štula, and L. Šerić, *Improved sentence retrieval using local context and sentence length*, Information Processing & Management, 49 (6)(2013) 1301–1312.
[12] M. Ebrahimi, *Islamic identity, ethical principles and human values*, European Journal of Multidisciplinary Studies, 2 (6) (2017) 325–336.
[13] M. Erritali, *Information retrieval: Textual indexing using an oriented object database*, Indonesian Journal of Electrical Engineering and Computer Science, 2 (1) (2016) 205–214.
[14] R. T. Fernández and D. E. Losada, *Effective sentence retrieval based on query-independent evidence*, Information Processing & Management, 48 (6)(2012) 1203–1229.
[15] S. Ganiger and K. M. M. Rajashekharaiah, *Comparative study on keyword extraction algorithms for single extractive document*, in Second International Conference on Intelligent Computing and Control Systems (ICICCS), (2018) 1284–1287.
[16] H. Gulam, *The application of shariah (islamic law) in some different countries and its implications*, Jurnal Syariah, 24 (2)(2016) 321–340.
[17] H. Gulam, *Comparing the legal rules (ahkam-i shariah) in the maliki and shafii schools with the hanafi school of thought*, Jurnal Ilmiah Islam Futura, 16 (1) (2017) 1–8.
[18] H. M. Hasan, F. Sanyal, D. Chaki, and M. H. Ali, *An empirical study of important keyword extraction techniques from documents*," in 1st International Conference on Intelligent Systems and Information Management (ICISIM), (2017) 91–94.
[19] H. M. M. Hasan, F. Sanyal, and D. Chaki, *A novel approach to extract important keywords from documents applying latent semantic analysis,* in 10th International Conference on Knowledge and Smart Technology (KST), (2018) 117–122.
[20] W. Jingjing, *Improving product-related patent information access with automated technology ontology extraction*, National University of Singapore, 2013.
[21] T. Johnson and M. A. Sergie, *Islam: governing under sharia*, 2014.
[22] R.G. Juan and M.-G. Luc´ıa, *A text mining-based approach for analyzing information retrieval in Spanish: music data collection as a case study*, in International Symposium on Distributed Computing and Artificial Intelligence, (2018) 259–266.
[23] W. Khairuldin, W. Anas, and A. H. Embong, *The binding of laws to personal opinion of muftis in Malaysia*, Sciences, Journal of Academic Research in Business and Social, 8 (11) (2018) 522–529.
[24] S. Lahiri, *Keywords at work: investigating keyword extraction in social media applications*, University of Michigan, 2018.
[25] F. Liu, D. Pennell, F. Liu, and Y. Liu, *Unsupervised approaches for automatic keyword extraction using meeting transcripts*, in Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the ACL, (2009) 620–628.
[26] E. L. Lydia, P. K. Kumar, K. Shankar, S. K. Lakshmanaprabu, R. M. Vidhyavathi, and A. Maseleno, *Charismatic document clustering through novel K-Means non-negative matrix factorization (KNMF) algorithm using key phrase extraction*, International Journal of Parallel Programming, 48 (3) (2020) 496–514.
[27] M. Mol et al., *Dimensionality of the system usability scale among professionals using internet-based interventions for depression: A confirmatory factor analysis*, BMC Psychiatry, 20 (1) (2020) 1–10.
[28] N. Moratanch and C. Gopalan, *A survey on extractive text summarization*, in International Conference on Computer, Communication and Signal Processing (ICCCSP), (2017) 1–6.
[29] M. K. Muhammed and M. M. Ali, *The shariah and law: An analysis*, International Journal of Law, Government and Communication, 2 (5) (2017) 48–64.
[30] D. L. Nkweteyim, *Data structures for information retrieval*, in IST-Africa Conference Proceedings, (2014) 1–8.
[31] S. Qaiser and R. Ali, T*ext mining: use of TF-IDF to examine the relevance of words to documents*, International Journal of Computer Applications, 181 (1) (2018) 25–29.
[32] K. A. F. A. Samah, N. F. S. Norhisam, S. F. A. Fesol, and R. Aminuddin, *TF-IDF and data visualization for syafie Madhhab hadith scriptures uuthenticity*, in 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE), (2021) 65–70.
[33] R. Sandberg and S. Thompson, *The sharia law debate: The missing family law context.*, Law and Justice, 177

(2016) 181–192.
[34] S. Siddiqi and A. Sharan, *Keyword and keyphrase extraction techniques: a literature review*, International Journal of Computer Applications, 109 (2) (2015) 18–23.
[35] P. Thalib, *Distinction of characteristics sharia and fiqh on islamic law,* Yuridika, 33 (3) (2018) 439–452.
[36] N. S. Tukiman and A. Ismail, *Semantic-Based Information Retrieval for Java Learning Management System*, Australian Journal of Basic and Applied Sciences, 8 (1) (2014) 52–57.
[37] Z. Wu, S. Shen, X. Lian, X. Su, and E. Chen, *A dummy-based user privacy protection approach for text information retrieval*, Knowledge-Based Systems, 195 (2020) 1–13.