DEVELOPMENT OF INDONESIAN CONSTITUTION QUESTION ANSWERING SYSTEM

Nur Aini Rakhmawati[1], Iwan Satriawan, Yordan Gunawan [2], Rochim Farul Noviyan[3]

The Constitution is the highest guideline for running the wheels of the state and government. In practice, society may not fully understand the constitution because of a lack of education and access to information. Therefore, in the case of Indonesia, it is important to increase the constitutional awareness of Indonesian citizens to effectively guard the practice of government. However, understanding and materials related to the constitution are usually only available in schools and universities. Thus, this research develops a question-answering system for a more accessible understanding of the Indonesian Constitution by using TF-IDF, Cosine Similarity, and Jaccard Similarity. Legal experts collected the dataset from Indonesian law. A user can deliver a question through a web application and an instant messaging bot (Telegram). It uses the Mean Reciprocal Rank to assess the QA system. TFIDF+Cosine similarity performance is slightly better than Jaccard similarity. Human evaluation generates worse performance than self-evaluation because users can ask different words for the same meaning question. The application is expected to accelerate the constitutional understanding of society because it is easier and more accessible for society.

Keywords: Question Answering, TF-IDF, Cosine Similarity, Jaccard Similarity, Indonesian Constitution

1. Introduction

Article 1, Paragraph (2) of the 1945 Constitution states that Indonesia is based on the Constitution. In other words, the Constitution is the highest guideline for running the wheels of the state and making decisions by authorities. However, some people are often dissatisfied with the implementation of the government by the executive, legislature, and judiciary[1]. For this reason, all Indonesian citizens need to build awareness of the Constitution to guard the running of the Indonesian state [2]

[1]Lecturer, Information Systems Department, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia, e-mail: nur.aini@is.its.ac.id

[2]Lecturer, Faculty of Law, Universitas Muhammadiyah Yogyakarta, Indonesia

[3]Student, Information Systems Department, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia

The challenge in building constitutional awareness is that knowledge of constitutional awareness can only be obtained in schools [3]. Hence, there is a need for other efforts to raise awareness of the constitution. One of the efforts is to facilitate access to the Constitution. Thus, this research develops a question answering system regarding the Indonesian Constitution using a term-weighting algorithm. Question answering (QA) is a form of information retrieval limited by using a set of questions or natural language statements as part or all of the expressions used to retrieve information. For example, who is the architect of the national monument building in Jakarta? What is the weight of gold in national monument buildings? On one hand, questions in a natural language must define the information to be searched. Still, on the other hand, these questions are not written as a list of search terms but in a form that can represent syntactic and semantic relationships between search terms [4].

The concept of question answering has been around for decades, stemming from research on natural language interfaces for accessing data or knowledge bases. This old technology is experiencing a resurgence because of the need to search for information from multimedia data, such as images, video, text, or audio, using natural language expressions. The relevance of concluding content in these media is mainly in the text [4].

ChatGPT is the famous QA system for answering multiple topics. The ChatGPT is trained by a large dataset and labelled by common people[5, 6]. The legal domain are required to be examined carefully by a legal expert. To the best of our knowledge, there is currently no known platform that specifically focuses on the Indonesian Constitution for a question-answering system, particularly targeting the general population. The study in question is related to three studies [7, 8, 9]. In one of these studies, Dhanani et al. discussed the significance of establishing a Legal Document Recommendation System (LDRS) for judges in India, which involves the extraction of decision documents from past cases. The system utilizes doc2vec for document retrieval. The problem is that doc2vec uses large resources and processing time, so the researchers optimized doc2vec by using the Generalized English and Indian Legal Dictionary (GEILD) to reduce noise in the document [7]. Weiyi Huang, Jiahao Jiang et al. [8] raised the weakness of the use of QA in a specific domain. Therefore, the researchers applied a question answering system for the legal domain. Chalkidis et al. [9] highlighted the poor use of BERT [10] in a specific domain. They formulated an approach that can be taken to use BERT for the legal domain. Term-weighting is a technique for weighting terms or words that are often used to retrieve information. One of the most widely used term-weighting techniques is Term Frequency - Inverse Document Frequency (TFIDF). TFIDF provides word weights by calculating the word's frequency of occurrence [11].

To this end, we make the following contributions:

- We develop a question-answering system for the Indonesian Constitution by using TF-IDF, Cosine Similarity, and Jaccard Similarity
- We create a dataset for the QA Indonesian constitution.
- We provide two interfaces: a web application and an instant messaging bot.

The remainder of this paper is organized as follows: Section 2 presents the background of the Constitution, TFIDF, Cosine Similarity, and Jaccard Similarity. Section 3 details the methodology and evaluation. In the fourth section, we describe the obtained results of the QA system. In the last section, we conclude the QA system and suggest improvements for future research.

2. Background Theory

2.1. Constitution

The Constitution is the highest fundamental law and uses as a guiding principle in a country [12]. The Constitution can be in the form of written basic laws such as basic laws or unwritten ones known as conventions [13]. The Constitution of the Republic of Indonesia in the form of fundamental law is the 1945 Constitution [14]. The 1945 Constitution, commonly abbreviated as UUD 1945, was amended four times from 1999 to 2002. The amendments made were one of the efforts to promote constitutionalism in Indonesia. Indonesia has adopted the principles of modern constitutionalism through the division of state power, specifically the Legislative, Executive, and Judiciary (Divisions of Powers). This adoption ensures mutual oversight and balance each other (checks and balances) [3]. In an institutional context, the Constitution refers to the highest proclamation that sets a holder of supreme sovereignty, state structure, form of state, form of government, legislative power, power judiciary, various state institutions, and rights of the people [15]. The 1945 Constitution positioned itself as the foundation and pillar of the state, so all state administrations must be following the Constitution [16]. Understanding the norms of the Constitution is a democratic and constitutional step in the process of controlling power. The government runs the country on behalf of Indonesian citizens. This is also the way to strengthen constitutionalism, where the people directly understand the rights and obligations of citizens and can instantly give control to the government when the government carries out its duties do not comply with the existing laws [17]. By understanding a constitution, people have set limits and used their rights constitutionally so that the state is not arbitrary in carrying out its duties. Constitutional awareness is conceptually defined as a personal quality that radiates insight, attitude, and behaviors that contain the ideals and noble commitments of the Indonesian nation and state. Constitutional awareness is a form of citizens' awareness of the importance of implementing constitutional values [1].

2.2. Term Frequency - Inverse Document Frequency (TFIDF)

Term Frequency - Inverse Document Frequency (TFIDF) is the most popular unsupervised term weighting method. This approach consists of two components: term frequency (TF) and inverse document frequency (IDF). TF refers to how often a term appears and is found in a document, whereas IDF refers to how often a term appears in the entire document being examined. The The TF has various formula variants, such as $\log(tf)$, $\log(tf+1)$, and $\log(tf)+1$, where tf is the frequency with which the term appears in the document. IDF is formulated as $\log(n/n_j)$, $\log(n/n_j)+1$, and $\log(n/n_j-1)$, where n_j is the set of documents containing the jth term. The TFIDF value of a document must be higher when the search term appears frequently but on the condition that the term rarely appears in the entire document being examined [9].

2.3. Cosine Similarity

Cosine similarity measures the similarity between two documents. It uses a vector space similarity measure [16]. For the QA system, the first document is a list of keywords in question q, whereas the second document is a list of keywords in answer a. The cosine similarity between q and a is given by Equation 1.

$$sim(q,a) = \frac{a \cdot q}{|a||q|} \tag{1}$$

2.4. Jaccard Similarity

Jaccard distance or similarity is one of simple similarity calculation for an unordered set. The similarity of the QA system was calculated using Equation 2 [17].

$$sim(q,a) = \frac{|a \cap q|}{|a \cup q|} \tag{2}$$

3. Methods

We develop our Question Anwer (QA) system (Figure 1), which consists of three steps: 1) Creating a dataset, 2) building the QA system, and 3) Evaluation.

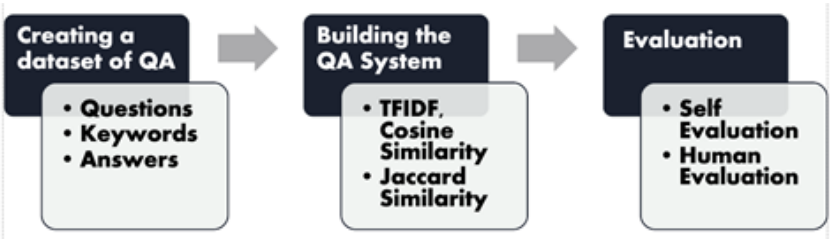

Figure 1. Methodology

Question	Keywords	Answer
Siapa yang mempunyai wewenang mengubah Konstitusi?	berwenang mengubah konstitusi	Yang berwewenang untuk mengubah Konstitusi di Indonesia (UUD 1945) adalah MPR (yang beranggotakan DPR dan DPD). Ketentuan ini sesuai dengan Pasal 3 (1) UUD 1945 yang menyebutkan bahwa Majelis Permusyawaratan Rakyat berwenang mengubah dan menetapkan Undang-Undang Dasar.
Apakah UUD 1945 atau Konstitusi Indonesia pernah mengalami perubahan?	perubahan Konstitusi di Indonesia	Ya UUD 1945 pernah mengalami perubahan sebanyak 4x. Pertama periode 18 Agustus 1945-27 Desember 1949 Penetapan Undang-Undang Dasar 1945. Kedua, periode 27 Desember 1949-17 Agustus 1950 Penetapan Undang-Undang Dasar Republik Indonesia Serikat (RIS). Ketiga,periode 17 Agustus 1950-5 Juli 1959 Penetapan Undang-Undang Dasar Sementara. Terakhir, periode 5 Juli 1959-sekarang Penetapan berlakunya kembali Undang-Undang Dasar 1945

TABLE 1. Example of the QA dataset (in Bahasa)

3.1. Creating a dataset of QA

We create a QA dataset consisting of 104 rows. Each row has three components: question, keyword, and answer. This dataset is generated based on journals and books. In addition, legal experts in the domain select the questions and their answers. An example of our dataset is shown in Table 1.

3.2. Building the QA system

The dataset from the preparation phase was used to develop the QA system. The QA system architecture developed at this stage is shown in Figure 2. The QA system provides two interfaces: Web application (`https://tanyahalal.herokuapp.com`) and Telegram (`https://t.me/tanyahukumbot`).

As shown in Figure 3, the question received is tokenized into a list of words. Then,unnecessary characters, such as punctuation marks and notations, are removed. In addition, the terms included in the stop words are omitted. Subsequently, we calculated the TF-IDF + Cosine similarity and Jaccard similarity and ranked their scores. The web application displays the

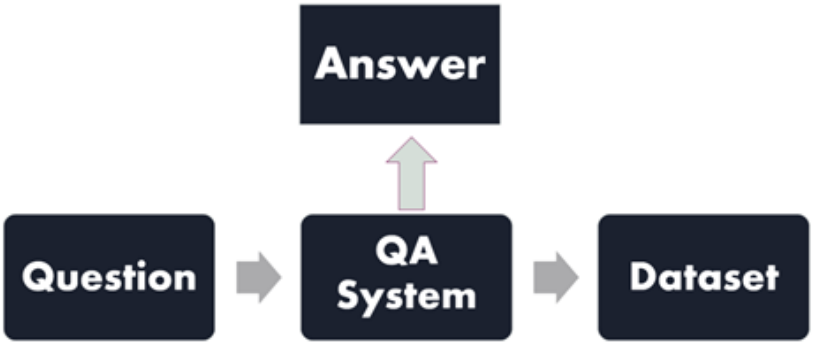

FIGURE 2. The architecture of the QA system

top five ranks for each algorithm. However, the Telegram app only returns the answer with the highest similarity score.

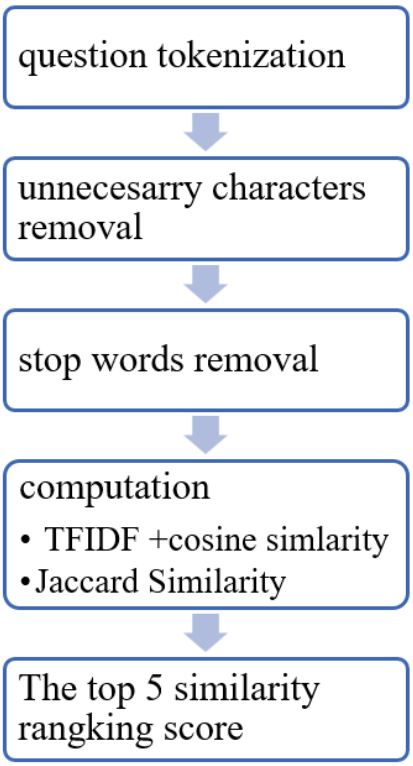

FIGURE 3. The QA process

During user interaction with the system, the Telegram server actively forwards each user's message to the backend of the QA system (Figure 4). The backend of the QA system consists of two services: bot services and QA services. The bot service connects to the Telegram server to receive messages sent by users. If the message received by the bot contains a question, it is

forwarded to the QA service via the HTTP protocol. If the message received by the bot is not a question, the bot will return an answer based on the prepared condition. The bot also acts as a question filter for QA services. The QA service receives a request for a question in an HTTP request and then processes the inquiry using the TFIDF algorithm. The answer generated by the algorithm is then returned as an HTTP response.

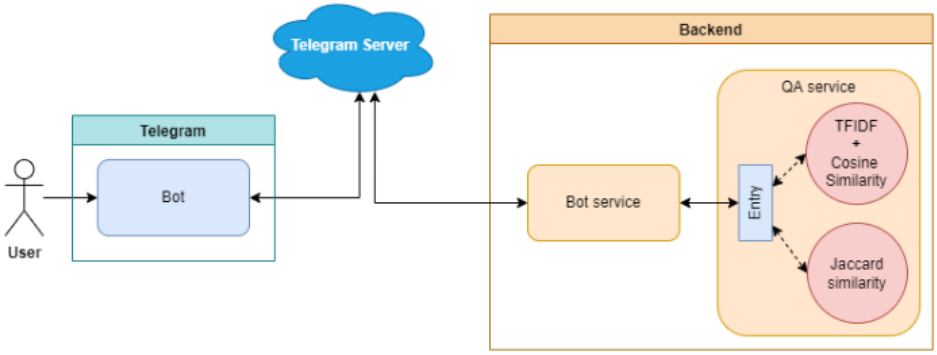

FIGURE 4. The Instant Messaging bot architecture

3.3. Evaluation Stage

The performance of the QA system is evaluated in two parts: self-evaluation and human evaluation. Self-evaluation uses a list of questions provided in the dataset, while human evaluation involves three people studying a constitution course. These evaluators create their own questions based on the answers provided. We make two combinations for the document: Question+Keyword and Question+Answer. Thus, there are four variations in the assessments, which can be explained as follows:

- Question+Keyword TF-IDF+Cosine similarity (QKTC)
- Question+Answer TF-IDF+Cosine similarity (QATC)
- Question+Keyword Jaccard similarity (QKJ)
- Question+Answer TF-IDF+Cosine similarity (QAJ)

These assessments are assessed by Mean Reciprocal Rank(MRR). MRR is useful when we want our system to return the most relevant item and want that item to be at a higher position[18]. MRR can be formulated as follows:

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i} \tag{3}$$

where:

- $|Q|$ refers to the number of questions
- $rank$ refers to the position of the first relevant result.

Algorithm	Self Evaluation	Human Evaluation
QKTC	97,1%	68,96%
QATC	59.6%	35%
QKJ	96,15%	80%
QAJ	35.5%	34%

TABLE 2. The Evaluation of QA System

The MRR only assesses the web application because it can display the top five results of the answer. The web application code can be accessed in [21], while the telegram bot application can be found in [22].

Furthermore, we also evaluate the ChatGPT by using the same list of questions. For this evaluation, we only calculate the accuracy of the ChatGPT answer. The accuracy is based on the human expert evaluation.

4. **Results and Discussions**

As shown in Table 2, the TF-IDF+cosine similarity algorithm with a dataset of a combination of questions and keywords (QKTC) produces the highest accuracy value compared to other algorithms for self-evaluation. The second-best algorithm for self-evaluation is the Jaccard similarity with the combination question and keyword (QKJ). The combination of question and answer performs worst because there is too much noise in the answer field. For instance, the "president" word could occur frequently in the legislative answer since the legislative cooperate with the president term.

The human evaluation produces worse performance than the self-evaluation because the evaluator sometimes picks different words for the question that do not appear in the question and keywords. Moreover, the QKJ performance surpassed the QKTC performance in human evaluation. We found that the evaluator added words in the question that did not appear in the correct answer. Consequently, the correct answer was not ranked first. Furthermore, given the same list of questions, the human expert stated that only 39% of the ChatGPT answers were correct. Our system is better than ChatGPT because it is based on several Indonesian constitutional books.

The web application is shown in Figure 5. A user can type any question related to the Indonesian Constitution. The QA system return the top five answers for each combination of algorithm (QKTC, QATC, QKJ, and QAJ). The web application interface uses the Streamlit(https://streamlit.io/). Users must type before asking a question (Figure 6). The bot returns only a single answer with a high similarity score.

5. **Conclusion**

The QA system for Indonesian constitution awareness has been presented, incorporating TF-IDF, Cosine similarity, and Jaccard similarity to

retrieve answers. Two interfaces are provided: a web application and a Telegram application. The web application is designed to return five answers with the highest similarity. According to our evaluation, the current system has certain limitations, including limited keywords and synonyms, because users may pose similar questions using different wording. The combination of questions and keywords showed better performance compared to keywords and answers, as the answer field contains words indirectly related to the question.

In the subsequent phase, two dictionaries will be implemented: an Indonesian word dictionary and a dictionary of terms pertaining to the constitution. The Indonesian word dictionary will be utilized alongside the Difflib module to identify non-standard words or spelling errors and make appropriate corrections. The dictionary of terms related to the Constitution restricts questions to the constitutional context. Furthermore, word embeddings such as word2vec and doc2vec will be explored as alternatives to TF-IDF.

Acknowledgement

The authors gratefully acknowledge financial support from the Institut Teknologi Sepuluh Nopember for this work under the Publication Writing and IPR Incentive Program (PPHKI) 2023 project scheme

REFERENCES

[1] *D. Sukriono*, Membangun Kesadaran Berkonstitusi Terhadap Hak-Hak Konstitusional Warga Negara Sebagai Upaya Menegakkan Hukum Konstitusi J. Legis. Indones., **vol. 13**, no. 3, pp. 273–283, May 2018, doi: 10.54629/JLI.V13I3.155.

[2] *A. Wahid, S. Sunardi, and D. A. Kurniawati*, Membumikan Konstitusi Indonesia Sebagai Upaya Menjaga Hak Kebhinekaan, Yurispruden J. Fak. Huk. Univ. Islam Malang, **vol. 2**, no. 2, pp. 180–188, Jun. 2019, doi: 10.33474/YUR.V2I2.2787.

[3] *F. Setiawan*, Peran Guru dalam Menumbuhkembangkan Kesadaran Konstitusional di Daerah Tertinggal melalui Penghayatan Nilai-Nilai Luhur Pancasila, JUPIIS J. Pendidik. ILMU-ILMU Sos., **vol. 12**, no. 1, pp. 178–185, Apr. 2020, doi: 10.24114/JUPIIS.V12I1.16311.G13239.

[4] *O. Kolomiyets and M. F. Moens*, A survey on question answering technology from an information retrieval perspective, Inf. Sci. (Ny)., **vol. 181**, no. 24, pp. 5412–5434, Dec. 2011, doi: 10.1016/J.INS.2011.07.047.

[5] *Aydın, Ömer and Karaarslan, Enis*, OpenAI ChatGPT generated literature review: Digital twin in healthcare. Emerging Computer Technologies **vol .2** pp. 22-31, 2022, doi: 10.2139/ssrn.4308687

[6] *Kocoń, Jan and Cichecki, Igor and Kaszyca, Oliwier and Kochanek, Mateusz and Szydło, Dominika and Baran, Joanna and Bielaniewicz, Julita and Gruza, Marcin and Janz, Arkadiusz and Kanclerz, Kamil and others*, ChatGPT: Jack of all trades, master of none, Information Fusion, **vol. 99**, pp. 101861, 2023, doi: 10.1016/j.inffus.2023.101861

[7] *J. Dhanani, R. Mehta, and D. P. Rana*, Legal document recommendation system: a dictionary based approach, Int. J. Web Inf. Syst., **vol. 17**, no. 3, pp. 187–203, 2021, doi: 10.1108/IJWIS-02-2021-0015.

[8] *W. Huang, J. Jiang, Q. Qu, and M. Yang*, AILA: A question answering system in the legal domain, in IJCAI International Joint Conference on Artificial Intelligence, 2020, **vol. 2021**-January, doi: 10.24963/ijcai.2020/762.

[9] *I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras, and I. Androutsopoulos*, LEGAL-BERT: The Muppets straight out of Law School, pp. 2898–2904, Oct. 2020, doi: 10.18653/v1/2020.findings-emnlp.261.

[10] *I. Tenney, D. Das, and E. Pavlick*, BERT rediscovers the classical NLP pipeline, 2020, doi: 10.18653/v1/p19-1452.

[11] *C.-H. Chen*, Improved TFIDF in big news retrieval: An empirical study, Pattern Recognit. Lett., **vol. 93**, pp. 113–122, Jul. 2017, doi: 10.1016/j.patrec.2016.11.004.

[12] A. Triningsih and O. V. Agustine, Community Participation as a Constitutional Awareness of Democracy Development in Digital Era, pp. 255–259, Jul. 2018, doi: 10.2991/ICILS-18.2018.48.

[13] *D. Sukriono*, Membangun Kesadaran Berkonstitusi Terhadap Hak-hak Konstitusional Warga Negara Sebagai Upaya Menegakkan Hukum Konstitusi, J. Legis. Indones., no. **Vol 13**, No 3 (2016): Jurnal Legislasi Indonesia-September 2016, pp. 273–283, 2016.

[14] *J. Asshiddiqie*, Konstitusi & Konstitusionalisme Indonesia. Jakarta: Sinar Grafika, 2021.

[15] *M. Laica and M.* Pendahuluan, Konstitusi dan Konstitusionalisme, J. Konstitusi, **vol. 7**, no. 4, pp. 001–008, May 2016, doi: 10.31078/JK.

[16] *T. A. Prihandono, S. K. Kusriyah, and W. Widayati*, Awareness on Constitutional Rights of Citizens and Form of Protection of Constitutional Rights of Citizens in Indonesia, J. Daulat Huk., **vol. 1**, no. 4, pp. 1003–1012, Dec. 2018, doi: 10.30659/JDH.V1I4.4145.

[17] *L. Marzuki*, Kesadaran Berkonstitusi dalam Kaitan Konstitusionalisme, **vol. 6**, no. 3, pp. 20–21, 2009, Accessed: Feb. 04, 2022. [Online]. Available: https://onesearch.id/Record/IOS1662.article-498.

[18] *N. A. Rakhmawati, A. A. Firmansyah, P. M. Effendi, R. Abdillah, and T. A. Cahyono*, Auto Halal detection products based on euclidian distance and cosine similarity, Int. J. Adv. Sci. Eng. Inf. Technol., 2018.

[19] *N. A. Rakhmawati and M. Jannah*, Food Ingredients Similarity Based on Conceptual and Textual Similarity, Halal Res. J., **vol. 1,** no. 2, pp. 87–95, 2021.

[20] *L. C. Chen and C. J. Luh*, Web page prediction from metasearch results, Internet Research, **vol. 15**, no. 4. pp. 421–446, 2005, doi: 10.1108/10662240510615182.

[21] *N. A. Rakhmawati*, nurainir/qa_hukum: QA Law with TFIDF,Cosine,Jaccard. Zenodo, Jan. 2022, doi: 10.5281/zenodo.5844259.

[22] *N. A. Rakhmawati and R. F. Noviyan*, rochimfn/tanyahukum-bot: Release v0.0.3, Jan. 2022, doi: 10.5281/ZENODO.5839674.