Journal of Education and e-Learning Research
Vol. 9, No. 4, 224-232, 2022
ISSN(E) 2410-9991 / ISSN(P) 2518-0169
DOI: 10.20448/jeelr.v9i4.4201

Developing a Semantic Question Answering System for E-learning Environments using Linguistic Resources

Maram Almotairi[1]
Fethi Fkih[1,2]

[1]*Department of Computer Science, College of Computer, Qassim University, Buraydah, Saudi Arabia.*
Email: ma.almotairi@qu.edu.sa
[2]*MARS Research Laboratory, University of Sousse, Sousse, Tunisia.*
Email: f.fki@qu.edu.sa

(Corresponding Author)

Abstract

The Question answering (QA) system plays a basic role in the acquisition of information and the e-learning environment is considered to be the field that is most in need of the question-answering system to help learners ask questions in natural language and get answers in short periods of time. The main problem in this context is how to understand the questions without any doubts in meaning and how to provide the most relevant answers to the questions. In this study, a question-answering system for specific courses has been developed to support the learning environment. The research outcomes indicate that the proposed method helps to solve the problem of ambiguities in meaning through the integration of natural language processing tools and semantic resources that can help to overcome several problems related to the natural language structure. This method also helps improve the capability to understand students' needs and, consequently, to retrieve the most suitable answers.

Keywords: E-learning environment, Information retrieval, Linguistic resources, Natural language processing, Question Answering system, Wordnet.

Citation | Maram, A., & Fethi, F. (2022). Developing a semantic question answering system for E-learning environments using linguistic resources. *Journal of Education and e-Learning Research*, 9(4), 224-232. 10.20448/jeelr.v9i4.4201.
History:
Received: 26 July 2022
Revised: 12 September 2022
Accepted: 30 September 2022
Published: 10 October 2022

Publisher: Asian Online Journal Publishing Group

Funding: This study received no specific financial support.
Authors' Contributions: All authors contributed equally to the conception and design of the study.
Competing Interests: The authors declare that they have no conflict of interest.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained.
Ethical: This study followed all ethical practices during writing.

Contents

Contribution of this paper to the literature
This study shows a significant step in developing the QA system of an e-learning environment that is able to answer a user's questions, written in a natural language, in an acceptable response time and with high accuracy. Furthermore, the e-learning environment is an important field and the QA system helps to improve the quality of this field and helps learners get the appropriate answers.

1. Introduction

In recent years, the growth of information has been steadily increasing. This has led to the problem of overloading of information. At present, when a student enters his question, the search engine collects a huge amount of information together with much irrelevant data, resulting in massive redundancy. In this case, the user may need to spend much time at finding the appropriate answer to his question. Users often prefer to ask questions without any rules of the query formation, to get short and accurate answers. However, the search engines are often not able to satisfy the information needs of the learners (Mohemad, Noor, Ali, & Li, 2017). The Question Answering (QA) system can automatically generate answers to the user's questions by using the Natural Language Processing (NLP) technology. It allows users to ask questions using natural language and receive accurate answers. Moreover, the QA system plays an important role in improving the learning environment by saving the learner's time and giving accurate answers (Tribhuwan & Mahender, 2015; Urvey, 2018). When a user asks his/her question using natural language, the computer understands the question by using NLP tools (De Silva, 2021; Popovski, Seljak, & Eftimov, 2020; Rajosoa, Hantach, Abbes, Ben, & Calvez, 2019). So, the QA system reduces the overload of learners' information through the use of NLP technology in educational contexts (Gurevych, Bernhard, Ignatova, & Toprak, 2009).

In QA systems, the answers are generated with specific information from a specific application domain, or with general information, depending on the topic. So, there are two domains to implement the QA systems: Open and Closed Domain systems. The open domain does not require a specific dictionary, it is generally searched for answers within a collection of large documents (Calijorne & Parreiras, 2020; Ramprasath & Hariharan, 2012). The closed domain however requires searching within specific document collections, and the quality of the answers is better. Also, the user needs to know the keywords of the specific domain to formulate the questions. There are various closed-domain QA systems developed in many environments, like e-learning environments (Mishra & Jain, 2016). So, the QA system plays an important role in improving the e-learning environment through saving the time of the learners and getting correct answers to their questions. Moreover, the system supports an e-learning environment with several advantages. In response to a learner query expressed in natural language, the system is supposed to return some accurate answers rather than a long list of ranked documents (Khushhal, Majid, Abbas, Nadeem, & Shah, 2020).

The main challenge is how to understand the user query without any ambiguity in meaning and how to retrieve the most relevant answer to this query? So, this study aims to integrate linguistic and semantic resources in the QA system to overcome problems related to the natural language structure, such as polysemy, synonymy, semantic ambiguity etc. This method will help improve the capability of the system to well understand the learner's needs and, consequently, to retrieve the most suitable answer. This study attempts to develop and implement an efficient QA system that can contribute to improving the e-learning environment. In order to do so, NLP tools will be used for analyzing and retrieving relevant information from textual data. In addition, a well-known ontology tool, Wordnet, will be integrated into the QA system to overcome semantic problems. The following are the objectives of the study:

- Creating a dataset for the selected course by compiling a dataset corpus composed of correct answers.
- Using specific techniques and tools to develop an efficient Question- Answering system.
- Retrieving the relevant answers from the dataset corpus.

The rest of the paper is organized as follows: Section 2 presents a brief review of the QA system. Section 3 describes the method and specific techniques that are used to develop the proposed QA system. Section 4 contains the experiences and evaluation results and section 5 provides the conclusions to this work.

2. Related Work

The primary task of the Question- Answering systems is to bring perfect answers to users' questions which are asked in natural language, instead of bringing complete documents (Palaniappan, Sridevi, & Subburaj, 2018; Sweta & Raghuwanshi, 2016). The performance of a question answering system lies in its focus on formalizing users' requirements and having good source corpus (Mishra & Jain, 2016). Abdi, Idris, and Ahmad (2018) presented a QA system that can be applied to the Al-Hadith domain, using linguistic knowledge. Their proposed method was able to improve performance and provide specific answers to users' queries. However, they encountered some common failures in the area of evaluation.

Agarwal, Sachdeva, Yadav, Udandarao, Mittal, Gupta and Mathur (2019) proposed a conceptual network model that applies intelligent indexing algorithms to improve answer generation. This model helps classroom learning by building interactive conversations. . Based on the approach described in Su, He, Fan, Zhang, and Guizani (2019), the authors proposed a method for Knowledge Base Question Answering systems. This method was used for getting an answer based on a dynamic memory network and reasoning judgment according to given questions. On the other hand, Zafar, Dubey, Lehmann, and Demidova (2020) presented an interactive approach to semantic QA systems. This approach facilitated the intuitive interaction of the user to identify the intended semantic query effectively. Therefore, the effectiveness of the semantic QA system increased for complex questions.

Based on their proposed model Ali and Yadav (2019) developed QA systems through reformulating the user's query before providing a relevant answer. They used the semantic web-based framework for QA systems as the environment and explained the performance of the model on the "eLearning" dataset. In this study, the results showed that the model's performance was more effective than the existing question- answering systems. LOD

(Linked Open Data) were used to develop the QA system and map the user's question to a standard query language (Al-Smadi, Al-Dalabih, Jararweh, & Juola, 2019). However, this approach has its limitations in the LOD representation of Arabic resources.

Albarghothi, Khater, and Shaalan (2017), used a semantic Web search which attempted to increase and enhance conventional search results by using ontologies that play a major role in this sort of search. Ontology is defined as the model of regular understanding of relations, interest and entities while Fontology is considered the backbone layer that is utilized to describe the semantics of web data. The Protege tool is used to develop and edit ontology. The tool however does not have the flexibility to update the names of classes or properties, which affect the accuracy of the data, as it requires deleting the categories and re-entering them to update the names. So, it still needs to improve the accuracy of results by adding more semantic rules.

Based on the approach described by Mohemad et al. (2017) the ontology model and parsed queries are assisted to detect matched answers depending on the process of conceptual matching. For ontology building, many entities and elements are identified, such as relations, individuals, classes, restrictions, rules, attributes, function terms, axioms and events. Abdi, Hasan, Arshi, Shamsuddin, and Idris (2020) proposed a QA system based on ontology; it has five main components: the ontology component, the user interface component, the question processing component, the answer extraction component and the users' query database component. The ontology used the lexical database to process the user's query.

3. Methodology

This section presents the general structure of the study's proposed methodology, as shown in Figure 1. The methodology includes two main steps:

1. First is the creation of a new dataset for a specific course; this involves the steps that help to create the datasets, like data collection and data preparation. Also, all the datasets have the same attributes.
2. The second step is the implementation of the semantic QA system which describes the steps and requirements.

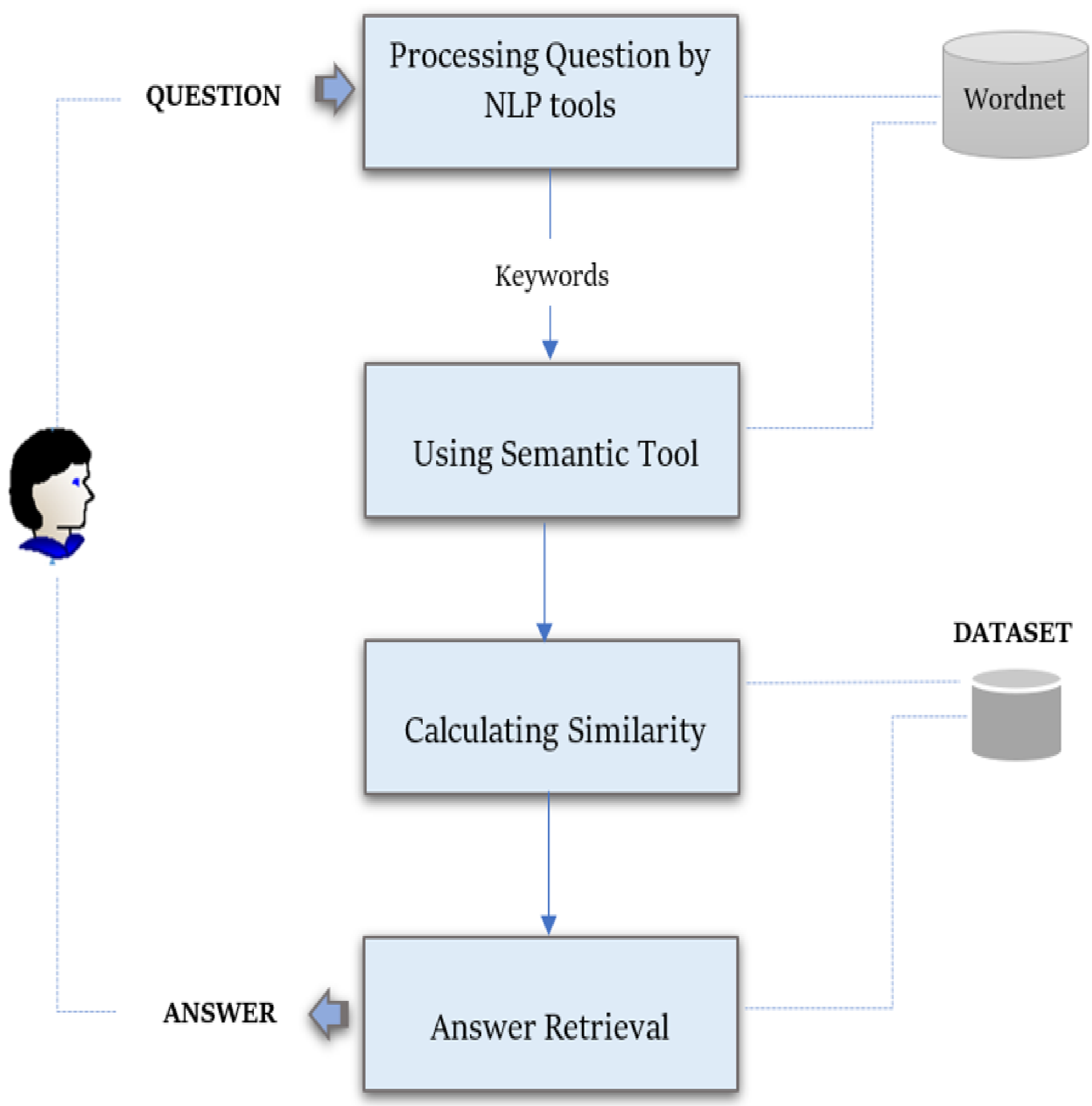

Figure 1. The structure of the proposed QA system.

In Figure 2, the proposed method includes the following steps shown in detail:

1. Processing the question: the system extracts the keywords from the question by using tokenization, removing the stop word and applying lemmatization to extract the keywords.
2. Using the semantic tool to find the synonyms (closely related meanings) of the extracted keywords from the question based on wordnet.
3. Calculating the similarity (by using cosine) between the extracted words and the dataset which contains the question topic.
4. Retrieving the exact answer, that has a higher value of cosine with other related answers. The related answer is another answer that depends on the descending cosine value.

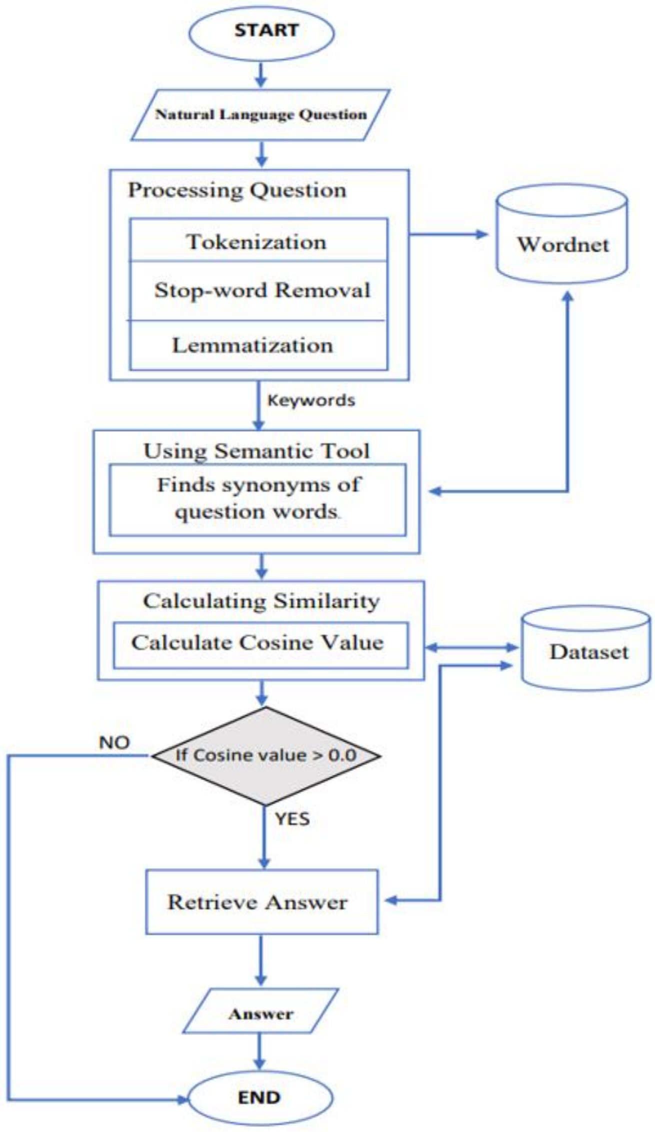

Figure 2. Flow chart of the proposed QA system.

In the present study, the process was started by preparing the dataset, selecting the specific course and integrating the data from different sources. Data were collected from three books: Information and Computer Technology, Computers and Communication Technology and Introduction to Computers. Furthermore, an Excel file of the dataset was created by using the manual method "Copy/Paste" for selected data and merging them into a standard spreadsheet in a worksheet format. Data were cleaned by removing incorrect or incomplete data and changing some data from uppercase to lowercase. Based on the aims of the study, some attributes in the dataset that did not relate to these aims were removed. The final attributes were the answer, question topic, page name, chapter name and a chapter number; the dataset contained500 rows.

3.1. Processing Questions using NLP Tools

When performing specific NLP-based tasks, it is important to understand the meaning of words in different contexts and pay attention to the similarities between the words as well. So, Wordnet is used to help solve the linguistic problems of NLP models. Wordnet comprises lexical resources that provide the structure of a hierarchy based on a set of one or more synonyms and semantic features of individual words. It is a large lexical database of English. It is grouped into nouns, verbs, adjectives and adverbs in sets of cognitive synonyms, each expressing a distinct concept. The structure of Wordnet makes it a useful tool for computational linguistics and natural language processing (Raja, John, Chakravarthi, Arcan, & Mccrae, 2018). This step includes mainly four operations to process the natural language question:

- Shifting to lowercase: When the user enters the question, the system converts the whole question from uppercase to lowercase. In python language, we can use the function 'lower' to convert the input questions to lowercase. For example, the question "What is Data?" Becomes "what is data?" after this operation.
- Removing punctuations: For more clarity and comprehension, punctuation marks are used to separate sentences and their elements (eg. the comma, question mark, semicolon, dash, etc). Based on the purpose of the data, the punctuation marks can be removed. When the punctuation is removed, it affects the processing before tokenizing the user's question. For example, the question will be "what is data" after removing the punctuation.
- Tokenization: It is the process of segmenting a text into linguistic units such as words, numbers, alphanumerics, symbols and phrases. It is used as input for processing. In python language, Natural Language Toolkit (NLTK) provides several ways to tokenize text. The NLTK library is a toolkit that can be used for natural language processing (Kanaan, Hammouri, Al-Shalabi, & Swalha, 2009; Wang & Hu, 2021). It was developed by Steven Bird and Edward Loper from the Department of Computer and Information Science at the University of Pennsylvania (Yao, 2019). Word tokenize is used to detect words

in a text. Based on the previous example, the question will be "['what', 'is', 'data']" after applying this operation.

- Removing the stop words: Stop words are the most common words in any language such as the, a, between, what, how.... etc. They help to build ideas but do not provide any meaning (Bouziane, Bouchiha, Doumi, & Malki, 2018). These stop words appear in many questions, so they are removed. In this context, all stop words are removed from the user question to minimize their effect. Moreover, it improves the performance of the system. Based on a previous example, the question will be "['data']" after removing the stop words.
- Lemmatization: A lot of words have different inflected forms. So, lemmatization works to group these forms together into a single item. It is similar to stemming, but it brings context to the words. For example: the word ['computers'], t will be "['computer']" after applying lemmatization.

3.2. Using Semantic Tools

The second step in the system is to find the synonyms for the words in the question. This means there may be two or more words or expressions in the same language with the same meaning. For example, the synonyms of the word (process) are: advance, advancement, furtherance, going, headway, passage, procession, progress and progression. In this step, it is applied for each word in the question. The user may enter the word that has a lot of words sharing the same meaning, so the system chooses the word that is found in the dataset, depending on the list of synonyms. Sometimes, the word has no synonyms, so the system puts in the original word.

3.3. Calculating Similarity

In this step, a Cosine similarity is used to calculate the similarity between the extracted keywords and the dataset which contains the question topic. It is a similarity measurement method between two different texts or documents by measuring the cosine of the angle between the document representation vectors (Fauzi, Utomo, Pramukantoro, & Setiawan, 2017). Figure 3 shows the code of cosine similarity and how it is measured. In the proposed method in the present study, 'A' means the word that is extracted from the question and 'B' is the word in the dataset.

```
Cosine_Value = 1- pairwise_distances( A, B, metric = 'cosine')
```

Figure 3. A python code of Cosine.

The answer which is retrieved has the higher cosine value between the extracted word from the question and the word in the attribute of Question Topic. The following section describes how the answer is retrieved by calculating the similarities.

3.4. Answer Retrieval

The last step in the question answering system is Answer Retrieval. It is considered the main task of QA systems to provide the best answer. The process of retrieving answers depends on extracting the main words of the question which match the question topic attributes. These words are then used to find the complete line which contains the answer. Next, the system retrieves two parts for the user: the exact answer and a related answer - which relies on the cosine value. In addition, the system retrieves the answer that has a higher cosine value. The system retrieves the other answer related to this question based on a descending cosine value. . On the other hand, the system retrieves no answer when it does not find the answer.

4. Experimental Results

The main objectives of this study can be categorized under two aims: (1) compiling the answer dataset corpus of selected courses and (2) developing an efficient Question- Answering system by using specific techniques and tools. Consequently, the user, who has a question and knows something about the course, can enter the question about course concepts. The user doesn't require learning any skills of programming language, but s/he has to have some idea of course content. When the user enters the question, the answer appears together with the related answer to the question. This step relies on the relationship between the question keywords and an additional answer, as mentioned in the previous section. This section describes the results of the input question. The main steps of the study's proposed method are NLP tools, synonyms, calculating similarity and extracting the answer together with the related answer. Figure 4, shows the steps from first entering the question to getting an exact answer from the system.

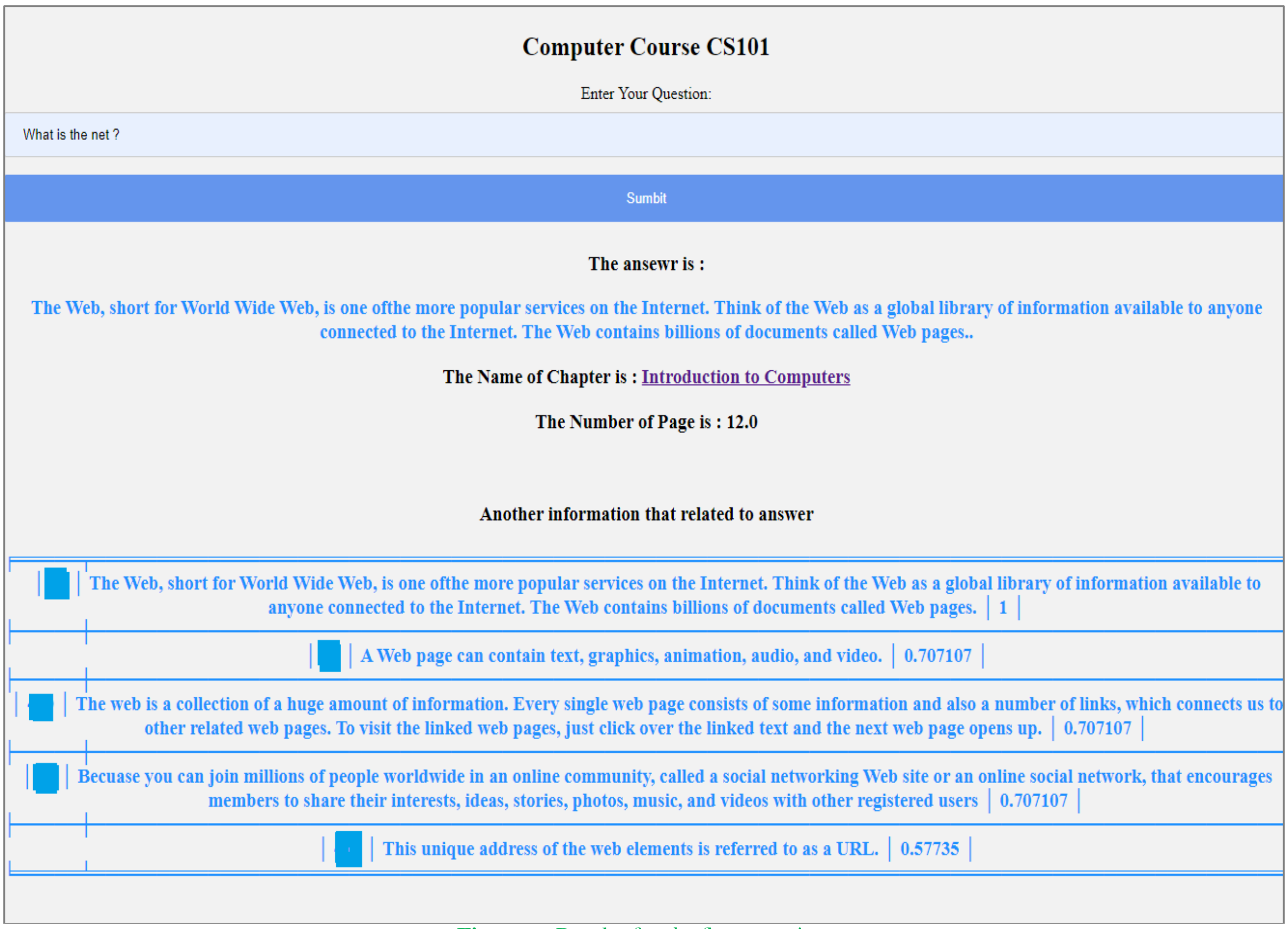

Figure 4. Results for the first question

In this example, the word "net" is not found in the dataset corpus. So, the system retrieves the answer by relying on the synonyms of the word. When the user enters his/her question, the system retrieves the exact answer that has the higher value, the page number that has this answer, the chapter name and the additional answer, beginning with the higher cosine value going down to a lower value. The second example in Figure 5 shows the steps from entering the question to getting an exact answer from the system.

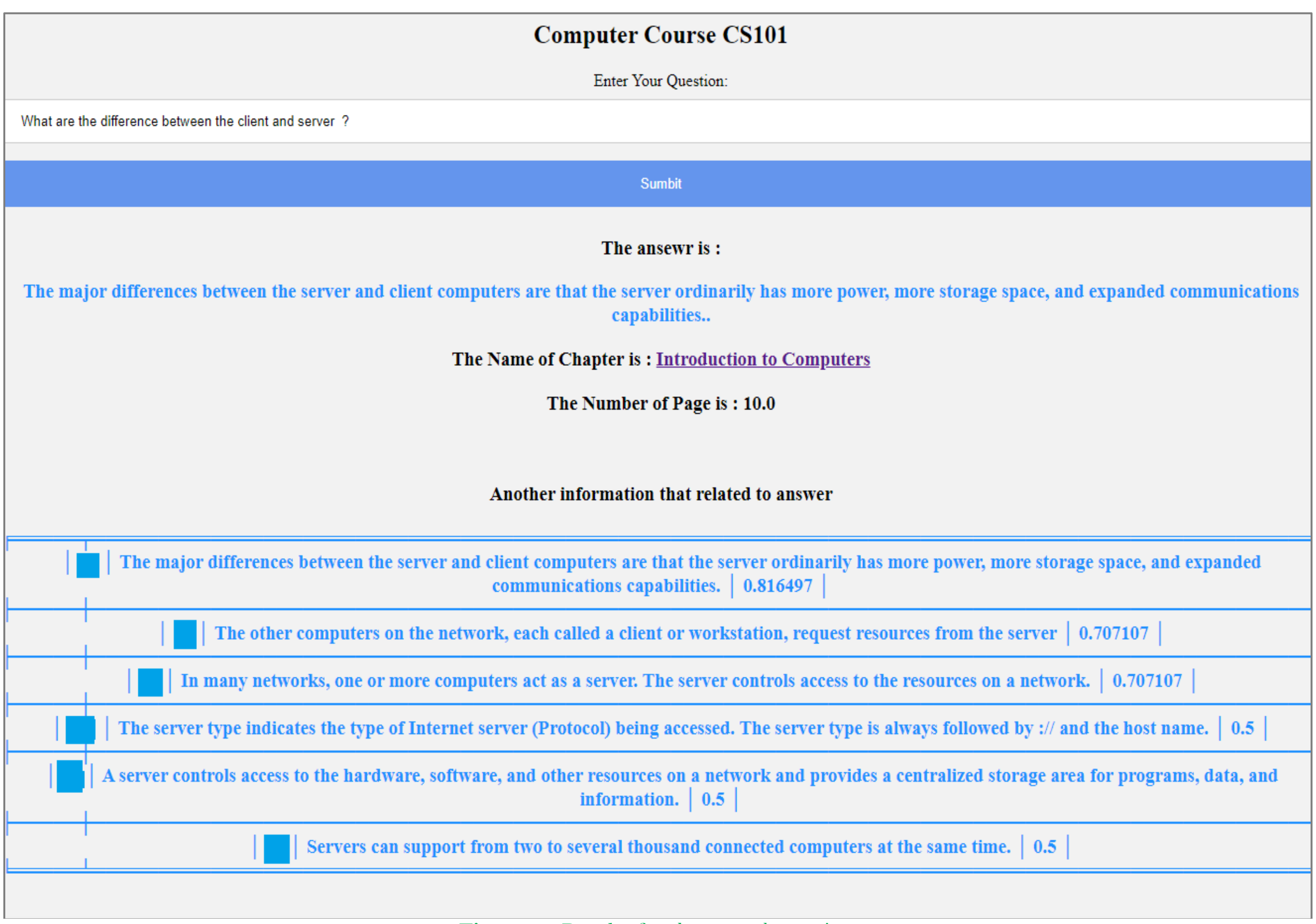

Figure 5. Results for the second question

By using NLP tools, the researcher is able to extract the keywords which facilitate the processes of the remaining steps. Here, the synonyms step is an important part of the proposed system as it helps greatly in improving the performance of the system and retrieving the exact answer for the user. After the process of extracting the keywords from the question, the synonyms step works on these words and finds words that have similar meanings. When the user enters the question with a word that is not found in the dataset, the system uses the synonyms step on this word to perform a search to find the appropriate answer. In addition, calculating similarity by using cosine is useful to extract the exact answer.

4.1. Evaluation Metrics

The main purpose of this evaluation is to gauge the performance of the proposed system. Generally, the performance of the proposed method can be measured by computing three standard measures: Precision, Recall, and F- measure on a test set. This can be defined as follows:

Equation 1 computes the Precision. It identifies what portion of the answers identified by the system is correct.

$$\text{Precision(P)} = \frac{\text{Number of Correct Answers}}{\text{Total Number of Answers}} \quad (1)$$

Equation 2 computes the Recall. It identifies what portion of the correct answers is identified by the system.

$$\text{Recall(R)} = \frac{\text{Number of Correct Answers}}{\text{Total Number of Question s}} \quad (2)$$

Equation 3 presents the F_measure. It combines both Precision and Recall.

$$\text{F} - \text{measure} = 2 \times \frac{(\text{P x R})}{(\text{P + R})} \quad (3)$$

The process of creating the users' questions in this study contains two main steps: (1) generating a set of questions according to the course dataset. (2) After deleting all same and duplicated questions, 200 questions were selected. To that purpose, the proposed system was tested with 200 different questions. In addition, the results of evaluation metrics were classified depending on the number of questions. The evaluation results are shown in Figure 6 and Table 1.

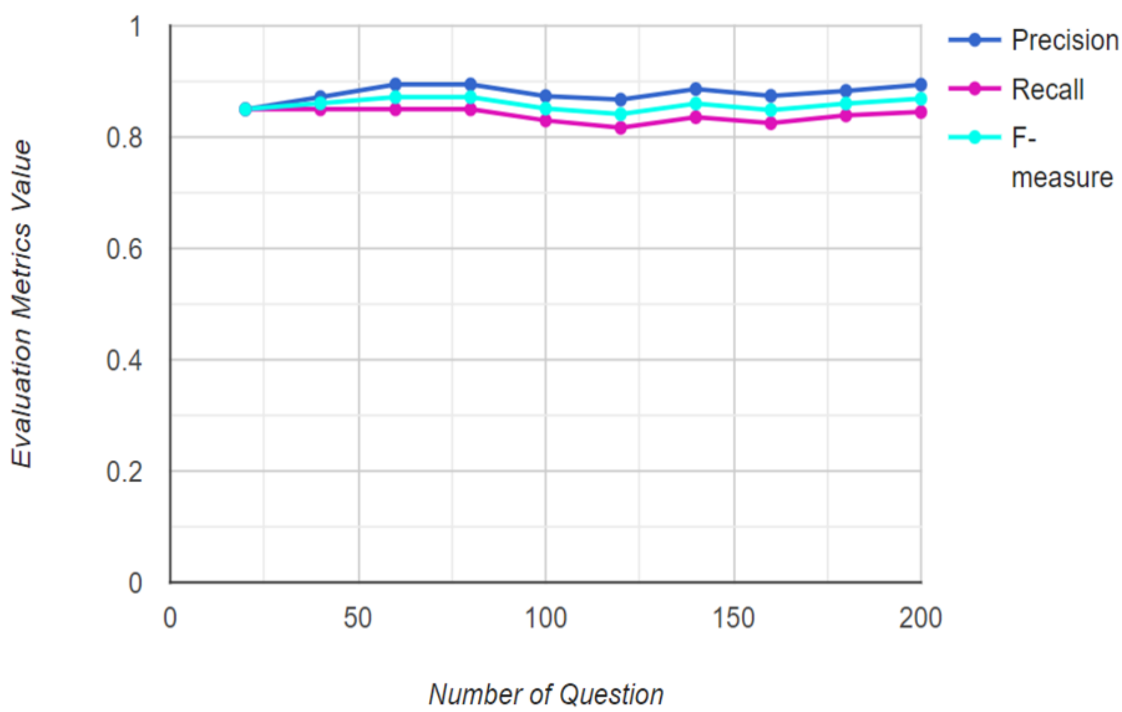

Figure 6. Performance of the proposed method.

Table 1. Performance evaluation of the proposed method.

No. of Questions	Precision	Recall	F-measure
20 Questions	0.85	0.85	0.85
40 Questions	0.8718	0.85	0.8608
60 Questions	0.8947	0.85	0.8718
80 Questions	0.8947	0.85	0.8718
100 Questions	0.8737	0.83	0.8513
120 Questions	0.8673	0.8167	0.8412
140 Questions	0.8864	0.8357	0.8603
160 Questions	0.8742	0.825	0.8487
180 Questions	0.8831	0.8389	0.8604
200 Questions	0.8942	0.845	0.8689

According to the results presented in Table 1, the proposed method has achieved a 0.8942 level of success in precision, 0.845 in Recall and 0.8689 in F-measure. Also, it shows that in 20 out of 200 questions presented, the system returned wrong answers. Furthermore, 11 of the 200 questions could not be answered because there was no corresponding pattern in the query database for the question word keypad in by the user.

4.2. Comparison of Proposed System with other Systems

The systems proposed in the present study are A Web Question Answering System (AWBQAS) and Quranic Arabic/English Question Answering System (QAEQAS) (Ahmed & Anto, 2017; Hamoud & Atwell, 2016). AWBQAS proposed a Web QA system to improve the e-learning environment. This structure includes two main modules: the question analysis module and the answer extraction module. This system aims to develop the technique to detect the type of question. At the same time, NLP tools were used in the question processing modules, such as question tokenization and stop-word removal. In addition, question classification and question focus identification were used. The n-gram similarity metric was used to calculate the question n-grams. The system also used the World Wide Web as a source of knowledge which means that the open domain corpus was used. So, if a student submits a question in natural language the system generates an exact answer from the web and returns it to the student. Otherwise, our proposed use of the closed domain was used, based on a specific course corpus that helps students get the exact answers that serve their course. QAEQAS proposed the Quranic Arabic/English Question Answering System. Here, the NLP tools were used to process the question, through question tokenization and stop-word removal. Specific functions were used, such as re.match and re.findall to find matching answers in the dataset. In addition, set and result functions were used to score and rank the results. The QAEQAS was then able to return the correct answer to the user.

As shown in Figure 7, the system attained the best results, which means it shows good performance. QAEQAS achieved 0.8880 precision, 0.595 on recall and a 0.7125 f-measure. On the other hand, AWBQAS achieved 0.6010 precision, 0.58 on recall and a 0.5903 f-measure. It achieved a lower value compared with other systems, which means it has lower performance.

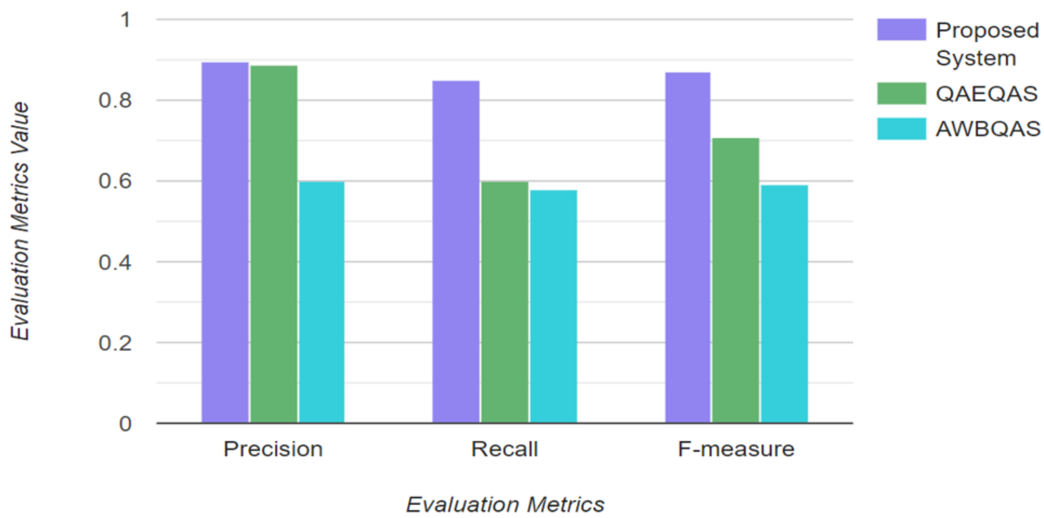

Figure 7. Performance between proposed system and other systems.

5. Conclusion

Many fields need systems to help users ask questions and get short and exact answers at the same time; e-learning is one such system. The problem lies in understanding the user's question without any ambiguity and retrieving the most relevant answer to this question. Therefore, the QA system is considered intelligent in understanding the natural language and generating the answers autonomously. The focus of this study was on developing an effective system that serves the field of e-learning.

In this paper, the semantic QA system in the e-learning environment was improved to be more effective. . A specific course dataset was created to achieve the aims of the study. In particular, this method used NLP tools: tokenization, stop word removal and lemmatization to improve the performance of the system. Moreover, to retrieve the most suitable answer, the synonym was used to overcome problems related to the natural language structure. In addition, the similarity was calculated by using cosine between the extracted keywords and the dataset which contains the question topic. This method improved the capability of the system to understand the students' needs in the selected course and, consequently, to retrieve the most suitable answers. The proposed system was evaluated through a 200 question course dataset (testing dataset) to assess the performance. Success was achieved with 89% Precision, 85% Recall and 87% F-measure. The main contribution is that a new dataset was created for the selected course, and it demonstrated the effectiveness of the technique used to improve the performance of the system. The proposed method was compared with other systems; the results showed that the method is good as the system can be adapted to a large number of questions. So, this method shows a significant step in developing the QA system of an e-learning environment. Consequently, the research aims were achieved successfully. This paper presents the following suggestions for future work. The proposed method is not able to cover all types of questions. Furthermore, it cannot provide answers to complex questions where the query of the user combines two or more question. During the evaluation, the experiment showed that there were some common failures encountered. These errors will be taken into consideration to improve the performance of the proposed system.

References

Abdi, A., Idris, N., & Ahmad, Z. (2018). QAPD: An ontology-based question answering system in the physics domain. *Soft Computing, 22*(1), 213-230.

Abdi, A., Hasan, S., Arshi, M., Shamsuddin, S. M., & Idris, N. (2020). A question answering system in hadith using linguistic knowledge. *Computer Speech & Language, 60*, 101023.

Agarwal, A., Sachdeva, N., Yadav, R. K., Udandarao, V., Mittal, V., Gupta, A., & Mathur, A. (2019). *EDUQA: Educational domain question answering system using conceptual network mapping*. Paper presented at the ICASSP, IEEE International Conference on Acoustics, Speech Signal Process.

Ahmed, W., & Anto, B. (2017). An automatic web-based question answering system for e-learning. *Information Technologies and Teaching Aids, 58*(2), 1-10.

Al-Smadi, M., Al-Dalabih, I., Jararweh, Y., & Juola, P. (2019). Leveraging linked open data to automatically answer Arabic questions. *IEEE Access, 7*, 177122-177136.Available at: https://doi.org/10.1109/access.2019.2956233.

Albarghothi, A., Khater, F., & Shaalan, K. (2017). Arabic question answering using ontology. *Procedia Computer Science, 117*, 183-191.Available at: https://doi.org/10.1016/j.procs.2017.10.108.

Ali, I., & Yadav, D. (2019). Question reformulation based question answering environment model. *International Journal of Information Technology, 13*(1), 59-67.Available at: https://doi.org/10.1007/s41870-019-00332-8.

Bouziane, A., Bouchiha, D., Doumi, N., & Malki, M. (2018). Toward an Arabic question answering system over linked data. *Jordanian Journal of Computers and Information Technology, 04*, 1-14.Available at: https://doi.org/10.5455/jjcit.71-1514749838.

Calijorne, S. M. A., & Parreiras, F. S. (2020). A literature review on question answering techniques, paradigms and systems. *Journal of King Saud University-Computer and Information Sciences, 32*(6), 635-646.
De Silva, N. (2021). Survey on publicly available sinhala natural language processing tools and research. *arXiv preprint arXiv:1906.02358*, 1–18.
Fauzi, M. A., Utomo, D. C., Pramukantoro, E. S., & Setiawan, B. D. (2017). *Automatic essay scoring system using N-GRAM and cosine similarity for gamification based e-learning*. Paper presented at the The ACM International Conference Proceeding Series.
Gurevych, I., Bernhard, D., Ignatova, K., & Toprak, C. (2009). Educational question answering based on social media content. *Frontiers in Artificial Intelligence and Applications, 200*(1), 133–140.
Hamoud, B., & Atwell, E. (2016). Using an islamic question and answer knowledge base to answer questions about the Holy Quran. *International Journal on Islamic Applications in Computer Science Technology, 4*(4), 20–29.
Kanaan, G., Hammouri, A., Al-Shalabi, R., & Swalha, M. (2009). A new question answering system for the Arabic language. *American Journal of Applied Sciences, 6*(4), 797–805.
Khushhal, S., Majid, A., Abbas, S. A., Nadeem, M. S. A., & Shah, S. A. (2020). Question retrieval using combined queries in community question answering. *Journal of Intelligent Information Systems, 55*(2), 307-327.
Mishra, A., & Jain, S. K. (2016). A survey on question answering systems with classification. *Journal of King Saud University-Computer and Information Sciences, 28*(3), 345-361.
Mohemad, R., Noor, N. M. M., Ali, N. H., & Li, E. Y. (2017). Ontology-based question answering system in restricted domain. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 9*(3-4), 29-33.
Palaniappan, S., Sridevi, U., & Subburaj, J. (2018). Ontology based question answering system using JSON-LD for Closed Domain. *International Journal of Pure and Applied Mathematics, 119*(12), 1969-1980.
Popovski, G., Seljak, B. K., & Eftimov, T. (2020). A survey of named-entity recognition methods for food information extraction. *IEEE Access, 8*, 31586-31594.
Raja, B., John, P., Chakravarthi, B. R., Arcan, M., & Mccrae, J. P. (2018). *Improving wordnets for under-resourced languages using machine translation.* Paper presented at the GWC 2018, The 9th Global WordNet Conference, Nanyang Technological University (NTU),Singapore, 8 – 12 January.
Rajosoa, M., Hantach, R., Abbes, S., Ben, & Calvez, P. (2019). *Hybrid question answering system based on natural language processing and SPARQL query*. Paper presented at the In Proceeding 3rd International Workshop Applyling Knowledge Represent. Semantic Technol. Robotics.
Ramprasath, M., & Hariharan, S. (2012). An overview of question answering system. *IUP Journal of Information Technology, 8*(2), 58–66.
Su, L., He, T., Fan, Z., Zhang, Y., & Guizani, M. (2019). Answer acquisition for knowledge base question answering systems based on dynamic memory network. *IEEE Access, 7*, 161329-161339.Available at: https://doi.org/10.1109/access.2019.2949993.
Sweta, L., & Raghuwanshi, M. (2016). Closed domain question answering system using nlp techniques. *IJESRT, 5*(4), 632–639.
Tribhuwan, S., & Mahender, C. (2015). A review: Analysis of questions for development of automated inference engine of QA system. *Advances in Computational Research, 7*(1), 288–290.
Urvey, S. (2018). Question answering systems: Analysis and survey. *International Journal of Computer Science and Engineering Survey, 9*(6), 1–13.
Wang, M., & Hu, F. (2021). The application of NLTK library for python natural language processing in corpus. *Theory and Practice in Language Studies, 11*(9), 1041-1049.Available at: https://doi.org/10.17507/tpls.1109.09.
Yao, J. (2019). *Automated sentiment analysis of text data with NLTK*. Paper presented at the Journal of Physics: Conference Series, IOP Publishing.
Zafar, H., Dubey, M., Lehmann, J., & Demidova, E. (2020). IQA: Interactive query construction in semantic question answering systems. *Journal of Web Semantics, 64*, 100586.Available at: https://doi.org/10.1016/j.websem.2020.100586.