

Integrating Ontology-based Knowledge to Improve Biomedical Multi-Document Summarization Model

Quoc-An Nguyen^{1*}, Khanh-Vinh Nguyen^{1*}, Hoang-Quynh Le¹,
Duy-Cat Can¹, Tam Doan-Thanh^{1,2}, Trung-Hieu Do³, and Mai-Vu Tran^{1**}

¹ University of Engineering and Technology
Vietnam National University Hanoi, Vietnam
{annq,22025025,1hquynh,catcd,vutm}@vnu.edu.vn

² Viettel Group, Hanoi, Vietnam
tamdt9@viettel.com.vn

³ Hanoi Medical University, Hanoi, Vietnam
dotrunghieu05220161@daihocyhanoi.edu.vn

Abstract. Most existing extractive summarization models use the original text’s internal information and calculate each sentence’s importance individually. When applied to specific domains (such as verbal text, biomedical literature, etc.), these models have some drawbacks: the variety of synonym terms, unknown words or terminologies, and the intra-document and inter-document relations between sentences or terms. In this work, we proposed an ontology-based summarization model that leverages many knowledge bases to understand the input documents. Our proposed model was built with an integrated ontology and a signal transmission-based method for extending domain knowledge such as related terms, and relationships between terms and sentences. The proposed model has been proven effective with the highest **ROUGE-2 F1** score in the test dataset of the MEDIQA 2021 MAS shared tasks.

Keywords: extractive summarization · multi-document summarization · query-based summarization · ontology construction

1 Introduction

In an age of overwhelming information, human beings have to distill a tremendous amount of data, which leads to time-consuming, costly, and heavy workloads. One effective solution to this problem is applying text summarization. The general idea of text summarization is receiving the document(s) as input, then generating a concise and informative summary [8]. Its goal is to assist users to capture the crucial information of the original documents without the need to

* Contributed equally

** Corresponding author

read the entire one. Text summarization is a challenging topic in the field of natural language processing and information retrieval. Researchers are constantly exploring new techniques to improve the quality of summaries.

Biomedical summarization is a potential area of research in text summarization. However, there are several distinctive challenges that make it become a tough task. Firstly, biomedical documents are highly scientific and domain-specific, with plenty of jargon and definitions such as drugs, diseases, and symptoms. Moreover, a term can have divergent meanings in different contexts, and different terms can be used to describe the same definition. This can lead to ambiguity and variability in the language used in biomedical literature, making it challenging to accurately summarize the content. Secondly, new biomedical terms can not be updated in pre-processing models. As the result, pre-processing models label plenty of unknown words. Thirdly, answering some questions requires biomedical relations such as chemical-disease, disease-gene, and disease-symptoms. The identification of biological relationships plays an important role in the query-based summary problem. From these challenges, in this work, we have proposed the ontology-based extractive multi-document summarization model in which ontology is used to leverage many knowledge bases to understand the input documents.

This research has two vital missions: *Ontology construction* and *Building extractive multi-answer summarization model*. Ontology construction tasks focus on building integrated ontology, which is leveraged to extend biological knowledge. WordNet also is used for enhancing common sense knowledge. After that, a summarization model is built based on scoring methods to select important sentences in original answers. Our paper contributions are:

- Proposing an ontology integration process and application scenario to augment question sense.
- Improving sentence scoring methods by injecting question keywords' weights.
- Building a summarization model which has better performance than the published model on MEDIQA 2021 MAS shared tasks dataset⁴.

2 Related work

There are two main approaches to text summarization - extractive and abstractive summarization. Extractive summarization generates synopsis by selecting remarkable sentences or expressions from the original content, while abstractive strategies reword and rebuild sentences to form the summary [1]. Because the advantage of extractive summarization is rapid, we decided to utilize it for complicated data like biomedical ones. There are various extractive text summarization methods have been researched and developed over an extended period. Frequency-based methods are the earliest one, it aims to estimate the importance of components based on the frequency of words or sentences [7]. Graph-based

⁴ <https://sites.google.com/view/mediqa2021>

methods commonly use sentence-based graphs to represent a document or cluster. LexRank/TexRank builds a graph from sentences/words and uses degree centrality as the score for ranking [9]. Besides, machine learning methods convert the summarization problem to a supervised classification problem at the sentence level. Some supervised models implemented are Naive Bayes, Decision tree, and Support Vector Machine to deep learning models [6]. These models depend on training data, they sometimes overfit these data but do not generalize well to novel ones, which leads to poor performance in real-life applications.

In some cases, models use additional user queries to extract summaries that only focus on answering these queries. Some question-driven scoring methods are used to estimate the relation between the query and answer components. Weighted-relaxed word mover’s distance (wRWMD) calculates the shortest distance between the query and answer sentences [18]. The Hierarchical Sentence Ordering (HSO) method is used for filtering important sentence which has tight relation to questions and other sentences [19]. Because the user’s queries tend to be short and have insufficient information to understand, query sense should be improved by using some external databases. Wikipedia is a sizeable unstructured database for searching keywords in Wikipedia articles and detecting related keywords [16]. ConceptNet and WordNet provide pre-existing relationships between words [19]. For an extended biological sense, the Medical Subject Headings database can be used in query expansion method [10].

Ontologies have been proven to be the most effective way for humans and machines to communicate and share information [17]. In this paper, ontology is manipulated to extend medical knowledge. There is a large number of biological ontologies nowadays such as Disease Ontology⁵, Gene Ontology⁶, MeSH⁷. However, there are no comprehensive ontologies, because actual ontologies focus on particular aspects of health. That makes it hard to apply to the summary problem because the user questions are diverse in many aspects. Ontology integration and ontology population are used to handle these obstacles [12]. Numerous ontology integration and population process are published [13, 3]. Although they have high performance and can be applied to many different data types including unstructured text, these approaches require much time and labour.

3 Proposed model

The proposed architecture is shown in **Figure 1**. It contains three main phases: ontology construction phase, pre-processing and ontology augmentation phase, and summarization phase. The ontology construction phase focus on creating an integrated ontology from the existing biomedical database. Pre-processing and ontology augmentation phase is used to generate a structured representation of the question and answers. Then, the summarization phase is used to manipulate

⁵ <https://disease-ontology.org/>

⁶ <http://geneontology.org/>

⁷ <https://www.ncbi.nlm.nih.gov/mesh/>

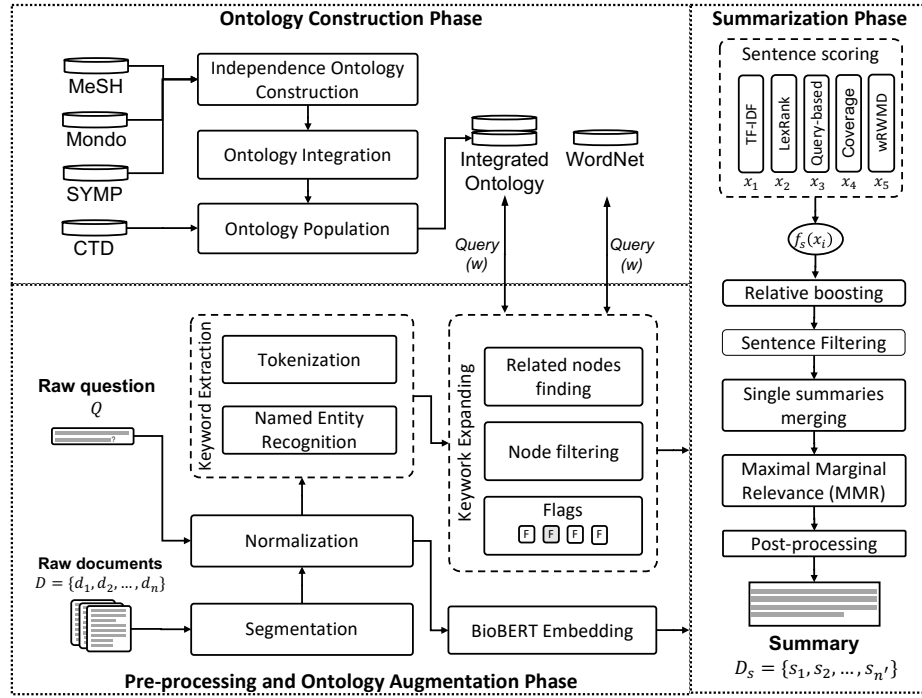


Fig. 1. Overview of proposed model

to create the extractive summary containing key information answering the given question.

3.1 Ontology Construction Phase

The biomedical text contains various term types (synonyms, keyword variation, related hypernym, related hyponym) and term relations (chemical-disease, disease-gene, disease-symptoms). The ontology can show relationships between terms about medical meaning. In a study focusing on solving the summarization problem, we propose a simple process that builds a comprehensive ontology based on existing structured databases. The pipeline for ontology construction has four main parts. Firstly, four biomedical databases are crawled:

- **Medical Subject Headings (MeSH):** MeSH⁸ is used to index book collections. In MeSH, the descriptors or subject headings follow a hierarchy, which helps build the relationship structure between terms. However, MeSH provides an incomplete amount of keywords on essential topics such as diseases, chemicals, and genes. There are two problems to solve: pruning unimportant trees to reduce noise and adding more in-depth data sources.

⁸ <https://www.ncbi.nlm.nih.gov/mesh/>

- **Mondo Disease Database (Mondo):** Mondo⁹ is a semi-automated database that merges diseases from existing sources into a consolidated database. It has a tabular structure, providing IDs from original databases. It supports adding disease knowledge to the summary model.
- **Symptom Ontology (SYMP)** Symptom Ontology¹⁰ includes a patient’s reported changes in feelings or appearance that are suggestive of disease. SYMP only provides symptoms knowledge without mapping to related diseases. Mapping SYMP to other diseases database is a great difficulty.
- **Comparative Toxicogenomics Database (CTD)** CTD¹¹ contains relations information such as chemical–gene chemical–disease, and gene–disease, which are created by human knowledge. It is an important data set in the ontology population phase.

Since databases have different structures and access methods, independent ontology is used to create a unified ontology structure from crawled data. Ontologies have a taxonomy structure, in which records are grouped or classified in hierarchical architecture. Each node will consist of the following attributes:

- **ID:** A unique sequence of characters is used to retrieve records.
- **terms:** Contains terms that refer to the same concept.
- **stem.from:** Storing name of original databases, serving for integration phase.
- **define:** Saving node’s definition, which supports ontology population phase.
- **edge:** Storing pointers to related nodes (**hypernym**, **hyponym**) and related relations. They help optimise speed access and computation.

In which, terms and definitions are normalized like raw texts. The data fields of MeSH, Mondo and SYMP are extracted and filtered to build independent ontologies respectively.

Because of overlapping information, independent ontology is entered into the ontology integration step. For indexing articles target, MeSH has a clear structure and a variety of topics, about 16 categories. MeSH is used as a backbone and Mondo ontology is integrated with MeSH based on cross references. There are two main cases:

- **Existence of cross-references between pair nodes:** Mondo’s node (A) is integrated into MeSH node (B) with two steps, (i) fusing A’s terms and B’s terms, (ii) cutting the descendant of A and converting it to the descendant nodes of B.
- **No cross-references between the pair nodes:** Mondo’s node is integrated into a shallow position

Finally, we use an ontology population process to add new semantic relations to integrated ontology. There are two ways to extract biomedical relationships. Firstly, we extract 1,048,547 chemical–disease and 1,048,576 gene–disease relations from CTD. Moreover, one method is used based on the term’s definition

⁹ <https://mondo.monarchinitiative.org>

¹⁰ <http://symptomontologywiki.igs.umaryland.edu>

¹¹ <http://ctdbase.org>

with three steps: (i) getting keywords from the term’s definition, (ii) searching keywords in ontology, (iii) if keywords exist, the relation is added to the integrated ontology [14]. As a result, chemical-disease, symptom-disease and gene-disease relations are added to the ontology.

3.2 Pre-processing and Ontology Augmentation Phase

Pre-processing The pre-processing phase takes the original documents and a raw question, as the input. Firstly, documents are segmented into sentences. After that, the normalization method is used to remove noise from the raw documents (HTML tags, duplicate spacing, etc.). Both processed documents and the raw question are put into the keyword extraction phase to extract keywords including tokens and NERs. The question’s keywords are expanding by Ontology Augmentation. Finally, BioBERT is used for creating word embedding vectors and sentence embedding vectors from the text.

Ontology Augmentation Ontology is leveraged to extend biological knowledge while WordNet is used for enhancing common sense knowledge. With ontology, when getting a keyword, the method searches key nodes which contain this keyword base on the longest common sub-sequence and sets an initial weight for them. Then, this method expands to related nodes (having relation to key nodes) and calculates their weight based on the spreading activation function. Spreading activation relies on signal transmission ideals when the strength of the signal gets weaker the farther it travels [18]. In the t^{th} expansion, the weights of the extended words are as follows:

$$weight_{w_t} = (1 - decay) * \frac{weight_{w_{t-1}}}{|Out_{w_{t-1}}|} \quad (1)$$

where $weight_{w_t}$ is weight of w_t and $|Out_{w_{t-1}}|$ is their number of outputs in $(t - 1)^{th}$ expansion. The spreading process begins to find related nodes and weight until $weight_t < threshold$. The method also provides some flags which allow setting which related term types are chosen. When stopping, this method returns keywords and their node’s weight in pair format. The spreading process in WordNet is similar to that in Ontology, in which a synset as a node.

3.3 Summarization Phase

Using information from the pre-processing and ontology augmentation state, the summarization phase generates the summary for each sample. Our summarization model tries to predict which sentences are important to the document by sentence scoring. This phase has four main parts: sentence scoring, relative boosting, removing duplicate semantic sentences and post-processing.

Sentence scoring We use five different types of scores to determine the importance of sentences. Besides, this paper proposed some customizations to inject the question keywords' weight.

TF-IDF score is a probabilistic approach that represents the importance of words in sentences [7]. The TF-IDF score of a word w located in document d of corpus D is defined as $TF-IDF(w, d, D)$. In this score, we add the weight of the maximum keywords max_w to the original formula as follows:

$$TF-IDF(w, d, D) = \max_w \times TF(w, d) \times IDF(w, D) \quad (2)$$

A sentence's TF-IDF score is the sum of the TF-IDF scores of its component words.

LexRank score is used for detecting essential sentences in the answer based on the document's graphs [5]. In LexRank, the graph's nodes present sentences, and a graph's weighted edge refers to the similarity score of node pairs. The score of sentences presenting the centrality of the sentence in the answer is calculated as follows:

$$p(e) = \frac{\alpha}{n} + (1 - \alpha) \sum_{r \in adj_e} \frac{\text{sim}(e, r)}{\sum_{z \in adj_r} \text{sim}(z, r)} p(r) \quad (3)$$

where adj_e is the set of nodes adjacent to e , n is the number of nodes, and α is the damping factor, sim is the Cosine similarity between vectors.

Coverage-based score We used the longest common sub-sequence to calculate the keyword's ratio per sentence as follows:

$$kw(s) = \frac{\left| \left\{ k : k \in q \mid \frac{\max_{i \in s} lcs(k, i)}{\text{length}(k)} \geq \text{thres} \right\} \right|}{|s|} \quad (4)$$

where q is the question's keywords and extending keywords, s is the sentence's keywords, and thres is the threshold.

Query-based score This score of sentence s focuses on calculating the similarity between the question vector v_s and sentence vector v_q by Cosine similarity as follows:

$$\text{score}(s) = \text{sim}(v_s, v_q) \quad (5)$$

Weighted-relaxed word mover's distance (wRWMD) is used to estimate the "distance" between question and sentence [18], which is as follows:

$$\text{wRWMD}(q, s) = \frac{\sum_{i \in q} w_i \times \max_{j \in s} \text{sim}(v_i, v_j)}{\sum_{i \in q} w_i} \quad (6)$$

where q is the question's keywords and extended keywords which exist in the answers, s is the sentence keywords set, w_i is normalized IDF. In the model, the question's keyword weight is added to the wRWMD method with w'_i as follows:

$$w'_i = \text{weight}_i \times w_i \quad (7)$$

Scores combination All scores are normalized by the min-max normalization method. Then, the final score is calculated in a weighted sum.

Relative boosting Because adjacent sentences often complement each other about semantic information (e.g. explaining or summarizing the previous sentences). As a result, answers frequently contain continuous sentences as a cluster in the paragraph. Relative boosting increases scores for some sentences close to high-scoring sentences. The updated score of sentence i -th is as follows:

$$score'_i = \max_{j=\max(i-n_l+1,1)}^{\min(i+n_r-1,n)} score_j \quad (8)$$

where n is the total number of sentences, n_l and n_r are the number of sentences in two directions: left and right of the current sentence i . After ranking and filtering the top highest-score sentences, we restore the sentence's position and concatenate all selected sentences to create single-answer extractive summaries.

Removing duplicate ideas The single-answer summaries are concatenated and then, passed into the maximal marginal relevance (MMR) [2]. MMR is used to remove overlap semantic and question-irrelevant sentences as follows:

$$MMR = \operatorname{argmax}_{s_i \in A \setminus R} \left[\lambda \left(\operatorname{sim}(s_i, q) - (1 - \lambda) \max_{j \neq i} (\operatorname{sim}(s_i, s_j)) \right) \right] \quad (9)$$

where R is the result sentence set, and A is the remaining sentences in the answer. s is the sentence in A , and q is the question. The output of this method also is multi-answer extractive summarization. The Cosine distance is followed to calculate the similarities.

Post-processing This method removes noises from the summary as questionable sentences, example sentences, long sentences, and duplicated information.

4 Experiments and Results

4.1 Dataset

This model uses MEDIQA-AnS Dataset as training set [20] which comprises 156 health-related questions, corresponding answers, and expert-created summaries of these answers. It includes single-answer and multi-answer summaries for extract and abstract strategies. The validation and test set are the datasets provided by the MEDIQA 2021 MAS shared tasks. The validation and test sets are summaries written by experts based on the original answers from the question-answering system CHiQA¹². The validation set includes 50 questions, as well as answers and multi-answer summaries to these questions. There are 80 questions and related answers in the test set. Because extract relates to choosing sentences and abstract rewritten sentences, **Table 1** provides statistics on given datasets to extract based on the sentence level.

¹² <https://chiqa.nlm.nih.gov>

Table 1. The statistics of dataset

Statistic aspects	Training		Validation	Test
	Article	Section		
Average				
Answers per Question	3.54	3.54	3.85	3.8
Sentences per Answer	84.93	29.07	14.50	13.03
Sentences per Single-answer Summary	6.31	6.31	-	-
Sentences per Multi-answer Summary	10.30	10.30	11.06	-
Compression ratio				
Single-answer Summary	0.12	0.49	-	-
Multi-answer Summary	0.06	0.18	0.33	-

4.2 Experiments and results

Figure 2 indicates the statistic of nodes and terms in three independent ontologies and in the integrated ontology. In three independent ontologies, MeSH has the most quantity of terms and nodes with 29,917 nodes and 245,918 terms in 16 topics. Mondo contains 24,409 nodes and 113,034 terms that focus on human diseases. SYMP has 944 nodes and 1179 terms about symptoms. In integrated ontology, there are 51,227 nodes and 346,098 terms.

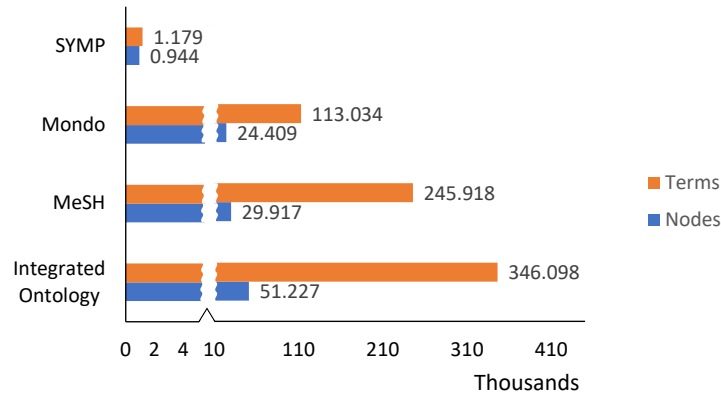
**Fig. 2.** The statistic of nodes and terms

Table 2 shows the number of relations after the population process. Extracting the CTD database allows adding 115,632 relations while extracting the term's definition allows adding 42,152 relations. As a result, a total of 157,784 semantic relations were extracted in the ontology population part.

ROUGE scores are used to evaluate automatic summarization and machine translation technologies [11]. The metric compares the generated summary to a human reference(s), the higher scores show the closer relations between them.

Table 2. The statistic of extracted relations in ontology population

Quantity	CTD database	Extracting term's definition
Chemical-disease relations	67,456	24,726
Symptom-disease relations	-	17,425
Gene-disease relations	48,176	-

Following MEDIQA 2021 MAS shared tasks, some ROUGE scores such as ROUGE-1, ROUGE-2, and ROUGE-L are used to evaluate the automatic summarization model in this paper, where F1 scores are main evaluation scores.

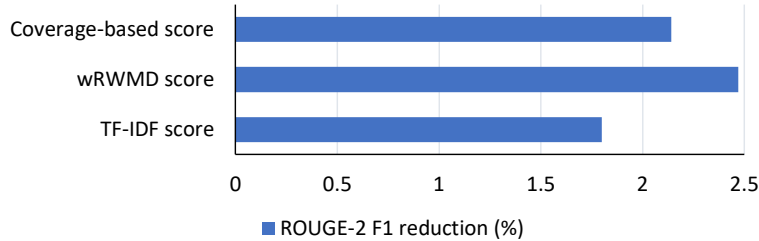
Table 3 shows the performances of the proposed model with some published models in this dataset [22, 4, 15, 21]. The ontology-based summarization model has the best performance on all three F1 scores. The proposed model has the best performance score when cutting integrated ontology where nodes about the drug, disease, symptoms, drug-disease relations and symptom-disease relations are focused on, and other related topics and gene-disease relations are disabled.

Table 3. Comparison model's results

Model name	ROUGE-1			ROUGE-2			ROUGE-L
	P	R	F1	P	R	F1	F1
Fine-tuning RoBERTa	0.475	0.878	0.585	0.407	0.767	0.508	0.435
BART	0.616	0.672	0.607	0.473	0.531	0.472	0.429
Fine-tuning T5	0.420	0.899	0.547	0.358	0.774	0.468	0.328
Prosper-thy-neighbor	0.528	0.814	0.611	0.432	0.680	0.504	0.441
Ontology-based model	0.530	0.821	0.616	0.437	0.687	0.511	0.446

The highest results in each column are highlighted in bold.

Figure 3 shows the reduction of ROUGE-2 *F1* per each scoring method when disabling the Ontology Augmentation phase. Keyword-based score and wRWMD have significant gains when using ontology.

**Fig. 3.** Reduction of ROUGE-2 *F1* per each scoring method

5 Conclusions

This paper presents the ontology-based extractive multi-document summarization model in which ontology is used to leverage many knowledge bases to understand the input documents. Because biomedical ontologies focus on particular aspects of health, we proposed a process to build a comprehensive ontology from some existing databases. After that, question keywords are through an ontology augmentation phase to find related keywords, and calculate their weight. In the summarization phase, we combined and optimized several scoring criteria such as TF-IDF, LexRank, TextRank, Coverage-based, Query-based and MMR scores. We propose some customizations to inject the question keywords' weight into sentence scoring. As a result, an integrated ontology includes 16 aspects and 5 biomedical relations. The summarization model performs better with 0.511 ROUGE-2 F1 than other published models in MEDIQUA 2021 MAS.

In the future, this model can be extended to build an ontology in more biomedical topics, using text data sources to infer related terms and biomedical relations. Moreover, this model can be implemented in Vietnamese to meet the needs for medical knowledge in my country.

Acknowledgments

This research has been done under the research project QG “Research and Development of Vietnamese Multi-document Summarization Based on Advanced Language Models” of Vietnam National University, Hanoi (Code: QG.22.61). Quoc-An Nguyen was funded by the Master, PhD Scholarship Programme of Vingroup Innovation Foundation (VINIF), code VINIF.2022.ThS.001.

References

1. Abualigah, L., Bashabsheh, M.Q., Alabool, H., Shehab, M.: Text summarization: a brief review. *Recent Advances in NLP: the case of Arabic language* pp. 1–15 (2020)
2. Bennani-Smires, K., Musat, C., Hossmann, A., Baeriswyl, M., Jaggi, M.: Simple unsupervised keyphrase extraction using sentence embeddings. In: *Proceedings of the 22nd Conference on Computational Natural Language Learning*. pp. 221–229 (2018)
3. Blomqvist, E.: Ontocase-automatic ontology enrichment based on ontology design patterns. In: *International Semantic Web Conference*. pp. 65–80. Springer (2009)
4. Can, D.C., Nguyen, Q.A., Duong, Q.H., Nguyen, M.Q., Nguyen, H.S., Ngoc, L.N.T., Ha, Q.T., Tran, M.V.: Uetrice at mediqa 2021: A prosper-thy-neighbour extractive multi-document summarization model. In: *Proceedings of the 20th Workshop on Biomedical Language Processing*. pp. 311–319 (2021)
5. Erkan, G., Radev, D.R.: Lexrank: Graph-based lexical centrality as salience in text summarization. *Journal of artificial intelligence research* **22**, 457–479 (2004)
6. Gambhir, M., Gupta, V.: Recent automatic text summarization techniques: a survey. *Artificial Intelligence Review* **47**(1), 1–66 (2017)

7. Hovy, E., Lin, C.Y., et al.: Automated text summarization in summarist. *Advances in automatic text summarization* **14**, 81–94 (1999)
8. Ježek, K., Steinberger, J.: Automatic text summarization (the state of the art 2007 and new challenges). In: *Proceedings of Znalosti*. pp. 1–12. Citeseer (2008)
9. Kaynar, O., Görmez, Y., Işık, Y.E., Demirkoparan, F.: Comparison of graph-based document summarization method. In: *2017 International Conference on Computer Science and Engineering (UBMK)*. pp. 598–603. IEEE (2017)
10. Kogilavani, A., Balasubramanie, P.: Ontology enhanced clustering based summarization of medical documents. *International Journal of Recent Trends in Engineering* **1**(1), 546 (2009)
11. Lin, C.Y.: Rouge: A package for automatic evaluation of summaries. In: *Text summarization branches out*. pp. 74–81 (2004)
12. Lubani, M., Noah, S.A.M., Mahmud, R.: Ontology population: Approaches and design aspects. *Journal of Information Science* **45**(4), 502–515 (2019)
13. Mitra, P., Noy, N.F., Jaiswal, A.R.: Omen: A probabilistic ontology mapping tool. In: *International Semantic Web Conference*. pp. 537–547. Springer (2005)
14. Mohammed, O., Benlamri, R., Fong, S.: Building a diseases symptoms ontology for medical diagnosis: an integrative approach. In: *The First International Conference on Future Generation Communication Technologies*. pp. 104–108. IEEE (2012)
15. Mrini, K., Dernoncourt, F., Yoon, S., Bui, T., Chang, W., Farcas, E., Nakashole, N.: Ucsd-adobe at mediqa 2021: Transfer learning and answer sentence selection for medical summarization. In: *Proceedings of the 20th Workshop on Biomedical Language Processing*. pp. 257–262 (2021)
16. Nastase, V.: Topic-driven multi-document summarization with encyclopedic knowledge and spreading activation. In: *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*. pp. 763–772 (2008)
17. Osman, I., Yahia, S.B., Diallo, G.: Ontology integration: approaches and challenging issues. *Information Fusion* **71**, 38–63 (2021)
18. Ozyurt, I.B., Bandrowski, A., Grethe, J.S.: Bio-answerfinder: a system to find answers to questions from biomedical texts. *Database* **2020** (2020)
19. Rahman, N., Borah, B.: Improvement of query-based text summarization using word sense disambiguation. *Complex & Intelligent Systems* **6**(1), 75–85 (2020)
20. Savery, M., Abacha, A.B., Gayen, S., Demner-Fushman, D.: Question-driven summarization of answers to consumer health questions. *Scientific Data* **7**(1), 1–9 (2020)
21. Yadav, S., Sarrouiti, M., Gupta, D.: Nlm at mediqa 2021: Transfer learning-based approaches for consumer question and multi-answer summarization. In: *Proceedings of the 20th Workshop on Biomedical Language Processing*. pp. 291–301 (2021)
22. Zhu, W., He, Y., Chai, L., Fan, Y., Ni, Y., Xie, G., Wang, X.: paht_nlp@ mediqa 2021: Multi-grained query focused multi-answer summarization. In: *Proceedings of the 20th Workshop on Biomedical Language Processing*. pp. 96–102 (2021)