

A Survey of Extractive Arabic Text Summarization Approaches

Samira Lagrini¹(✉), Mohammed Redjimi², and Nabih Azizi¹

¹ Labged Laboratory, Computer Science Department,
Badji Mokhtar University, PO BOX 12, Annaba 23000, Algeria
samiraboite@yahoo.fr, nabihal11@yahoo.fr

² Université 20 Août 1955, Skikda 21000, Algeria
redjimimed@yahoo.fr

Abstract. Automatic text summarization is an important research area originating from the late 50's but not losing its celebrity until now. Over the past half a century, automatic text summarization has seen a great interest especially in English language. However, in Arabic language, few works have been done in this field. This paper intends to survey the most relevant approaches in Arabic text summarization, giving special emphasis to extractive techniques. The limitation of these approaches and the main difficulties faced when dealing with such application are also discussed. Special attention is devoted to the peculiarities of Arabic language, which have posed challenges to the task of summarization.

Keywords: Arabic text summarization · Clustering
Rhetorical Structure Theory (RST) · Machine learning
Graph theory · Text entailment · Extractive approaches
Summary evaluation

1 Introduction

Text summarization is an essential application of Natural Language Processing (NLP). It is an imperative and timely tool for understanding text information. The objective of automatic text summarization is abridging texts into briefer version, conserving its overall meaning [1]. This allows the reader to decide whether a document contains required information with minimum effort. There is no doubt about the importance of such application. For example, it could be used as an informative tool in search engine web pages to find the pertinent and required information [2].

According to [3], a summary is “a text produced from one or more texts, that conveys important information of the original texts and that is no longer than half of the original text(s) and usually significantly less than that”.

Summarization systems can be categorized according to several characteristics: language, input, method output, generality...etc. (see Fig. 1). This enables summaries to be characterized by various properties [4]. For example, according to the degree of generality, a summary can be classified into generic or query driven summaries. Generic summary attempts to represent all relevant topics in the input document while

query driven summary depends on the user information need. We can also distinguish between single document summarization and multi-document summarization depending on the number of input documents to be summarized. Regarding the output, a summary can be either indicative or informative. Indicative summary is used to specify what topics are tackled in the input document. This will allow users to get an overall and a brief idea of the source text. Informative summary is intended to cover all topics addressed in the source text with further details.

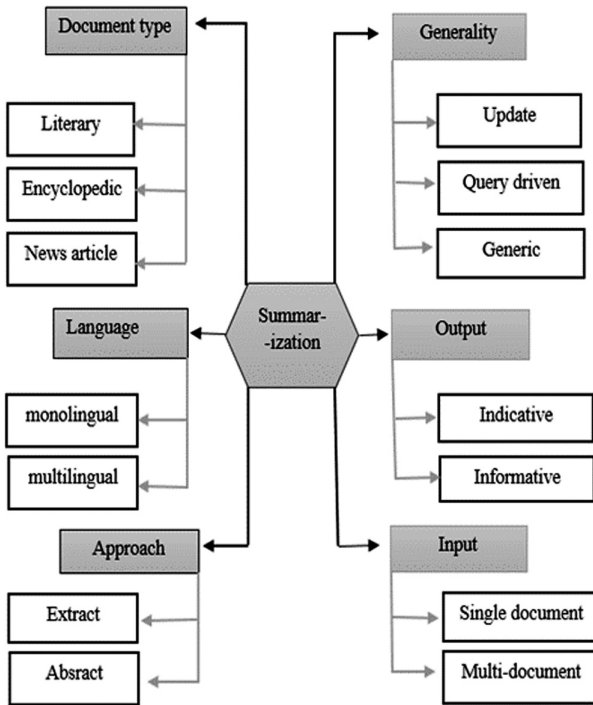


Fig. 1. Summarization taxonomy

Furthermore, we can talk about monolingual and multilingual summarizer. Monolingual summarization systems are designed to work with only one language and have the input document and the output summary in the same language, unlike multilingual summarizers, which cover more than one language.

Moreover, a broad difference is made between extract and abstract depending on the adopted approach. An extractive approach consists in selecting key sentences from the source document based on statistical and linguistic features, and concatenating them into a briefer form [1]. Abstractive approach differs mainly from extractive approach by providing summaries having some degree of inference about background knowledge not necessary presented in the original document [5]. In other words abstractive summarization means that, a new text is generated using the lexical, syntactical, semantic and rhetoric ingredients of the original text.

The goal of this paper is to survey the most salient extractive Arabic text summarization approaches.

The rest of this article is organized as follows: Sect. 2 gives an overview of some salient peculiarities of Arabic language. Section 3 focusses on summary evaluation issues. Then, Sect. 4 resumes the main proposed approaches for extractive Arabic text summarization. Section 5 explains the limitation of these approaches and the major challenges faced when dealing with such application. Finally, Sect. 6 concludes this paper.

2 Arabic Language Particularities

Arabic is the first language of more than 200 million persons through the world, and the official language of 21 countries [6]. Arabic language possesses specific peculiarities that make it distinctive, but at the same time, they pose several challenges to various Arabic natural language processing (ANLP) tasks, such as automatic summarization, sentence segmentation, and even word stemming. Some of these challenges include its complex morphology, the ambiguity, and its inflectional and derivational nature.

Regarding morphology, Arabic language is very rich and very complicated. Indeed, several words (sometimes more than ten) in Arabic can be formed using one single root, some patterns and some affixes. Affixed letters are very similar to root letters, which leads to several ambiguity. Thus, one single word could have diverse morphological features, as well as different POS. For instance, the word “فهم” can be tagged as a conjunction “ف” followed by the pronoun “هم” (they), or as a verb (to understand), or as a noun “فهم” (understanding).

Many reasons lead to this ambiguity. One salient reason is the lack of vowels that are only used in the holy Quran, and which are completely omitted in Modern Standard Arabic (MSA) written texts. Taking as an example, the word (علم) that can be read as (عَلَّمَ/Ellama/he teaches), or as (عَلِمَ/Elima/he knew), or as (عِلْمٌ/Elm/a science) or as (عَلَمٌ/Elam/a banner)

Another possible reason is the omission of writing marks, like Hamza (ء) and dots on letters. Therefore, dissimilar words can be written in the same way. For instance omitting the two dots on (ة) in the word (معلمة/a teacher), makes it exactly similar to (معلمه/his teacher). Similarly, omitting Hamza in the words (لأن/because) makes it identical to the verb (لأن/it softened). This type of ambiguity causes serious problems in many ANLP tasks including word sense disambiguation, machine translation and even word stemming.

Furthermore, Arabic does not have capital letters [7], which affect the recognition of named entities in the annotation process. For example, the word “وفاء” can be annotated as a named entity, or as an Accusative of purpose of the verb “وفى” which means ‘to honor’.

Finally, Arabic is a highly derivational and inflectional language [8]. Arabic words are generally composed of several morphemes. Thus, we can easily find one single word that can be represented by a complete statement. For instance, the word (أَنْتَزِمُكُمْ هَا) represents a statement that means ‘Shall we compel you to accept it’ (see Fig. 2.)

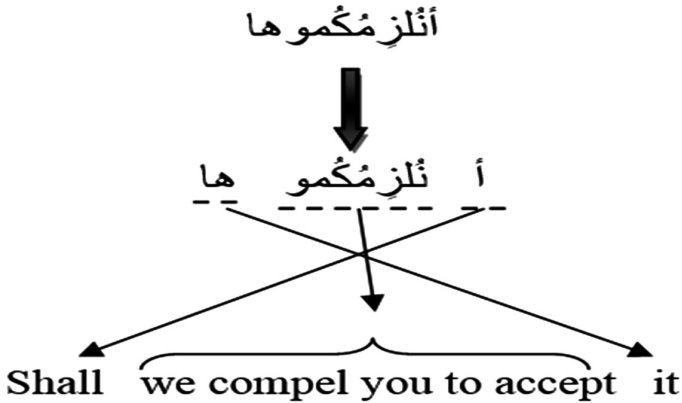


Fig. 2. Example of Arabic inflection

In fine, it is to be pointed out here, that the aforementioned challenges make difficult ANLP tasks, which can probably explain the lack of publically available tools and resources for Arabic language.

3 Summary Evaluation

Assessing the quality of summary is a challenging task in the field of automatic texts summarization. Indeed, there is no sole “perfect” summary. Summaries written by different people can be different at the content level. Writing this type of documents requires a deep understanding of the text in order to identify ideas, style and arguments, which each person does differently.

Another factor behind this challenging task is the fact that evaluating summary requires a comparison with reference summaries [9]. This implies the existence of benchmark corpora that contains documents to be summarized and their reference summaries. Creating such benchmark corpora is an expensive and time-consuming task [10]. Moreover, several summaries can be appropriate for the same document, and even the same person can summarize the same document in different way over time [11].

Moreover, evaluation process itself is a great problem. Person evaluation is time-consuming [12], and provides unsteady evaluation score. To overcome these problems, automatic methods such as ROUGE [13] and AUTOSUMENG [14] have been introduced.

According to [15], Evaluation methods are classified into intrinsic and extrinsic methods.

In extrinsic methods, summaries should be evaluated based on their utility and ability to perform certain tasks, such as classifying documents, or using summaries instead of original documents in question/answer systems. A summary is then considered effective if it allows its reader to answer the questionnaire as well as other readers who have read the source text. Intrinsic methods evaluate summaries based on their properties and content. Intrinsic methods consist of comparing machine

summaries with expected output data, such as one or more reference summaries, or relevant sentences chosen by human subjects to be included in the summary.

ROUGE (Recall-Oriented Understudy for Gisting Evaluation) [13] is a prominent measure that involves the differences between words distribution. It consists of a package that includes various ROUGE measures, like ROUGE-L (Longest Common Subsequence), ROUGE-N (N-gram Co-Occurrence Statistics), ROUGE-W (Weighted Longest Common Subsequence), etc.

ROUGE is highly used by DUC (Document Understanding Conferences) ever since 2004. This measure is considered as a standard by the community, because of its strong correlation with manual notations.

Although, based solely on the content of summary, ROUGE suffers from numerous drawbacks related to its dependence on the units (N-grams) used for the calculation of the scores. The multi-word units as “United States of America” and relatively unimportant words such as “the”, “but”, etc. biased the number of co-occurrences. In addition, many preprocessing steps that rely on language dependent resources are required previously [9].

Other automatic methods are also used. AutoSummENG (Automatic Summary Evaluation based on N-gram Graphs) [14] has been introduced as a language independent evaluation method. The basic idea behind this method is to create at first an n-gram graph for the candidate summary as well as for reference summaries. Then, the average of the similarities between the candidate summary and each reference summary is calculated in order to evaluate the system. As a variation of AutoSummENG, the MeMoG (MergedModel Graph) [14] relies on one merged graph representing the references summaries to calculate its similarity with the candidate summary rather than using all graphs of reference summaries.

At the ACL 2013 Multi-Ling Workshop, NPower (N-gram graph Powered Evaluation via Regression) [16] was added to the automatic evaluation methods. The authors used linear regressions to Combine AutoSummENG and MeMoG methods, and the evaluation process is formulated as a machine-learning problem. For more details about this method, see [16].

4 Arabic Text Summarization Approaches

This section describes the principal approaches proposed in the field of Arabic text summarization.

4.1 Discourse Theories

Rhetorical Structure Theory. The Rhetorical Structure Theory (RST) [17] is perhaps the most popular theory of discourse. In the RST framework, texts are represented by labeled trees, whose leaves correspond to atomic text segments, called elementary discourse units (EDUs), and internal nodes correspond to the rhetorical relations. Adjacent nodes in the tree structure are linked by rhetorical relations (causal, joint, manner, etc.) forming a discourse sub tree, which can then be subject to this linking. For more details about this theory, one should see [17].

The first employment of this theory on Arabic text summarization has been addressed by [18]. The authors suggested different techniques, algorithms and design patterns to be considered when developing Arabic summarizers based on RST.

Then, in [19] the rhetorical structure theory was also used for classifying Arabic security documents. The authors propose a technique that parses each paragraph in the document, build the rhetorical tree that represents its structure, and then determines the importance of each paragraph by examining the tree root. If the importance of the paragraph conforms to the user instruction, the classifier labels it with the required classification.

In [20], the authors propose a two-pass algorithm. In the first pass, RST is used to generate a primary summary. Therefore, a rhetorical analysis of a text is performed in order to generate all possible RS trees, upon which the primary summary is generated. In the second pass, each sentence within the primary summary is awarded a score based on word frequency and overlap with title keyword. To produce the final summary, sentences having the highest score are selected tacking into account the user compression ratio.

Other approaches provide a hybrid model like in [21]. The proposed model combines RST and vector space model (VSM). The model discovers at first the most important paragraphs based on semantic criteria, and then uses the VSM to rank these paragraphs based on the cosine similarity feature. Results revealed that combining VSM with RST improves the precision of the summary over employing RST only.

Segmented Discourse Representation Theory (SDRT). The Segmented Discourse Representation Theory (SDRT) proposed by [22] is a theory of discourse interpretation that seeks to combine two paradigms: discourse analysis and dynamic semantics. According to SDRT, a text is segmented into text units linked to each other via rhetorical relations, resulting into directed graphs called SDRS graphs. Unlike RST, in SDRT multiple discourse relations can link one discourse unit to adjacent or non-adjacent units. That is to say, several discourse relations can simultaneously link two text units in SDRT.

For the best of our knowledge, [23] addressed the first employment of this theory on Arabic text summarization. The authors tackle discourse analysis of Arabic documents following the SDRT framework. They explore how discourse structure can be exploited to produce indicative summaries. To this end, they design several algorithms that take as input the document discourse structure and generate as output a set of elementary discourse units, which will be used to produce the summary. To check the effect of discourse structure on producing indicative summaries, a comparison was made between the produced summaries and reference summaries, manually generated from two discourse annotated corpora following RST and SDRT framework. Results revealed that all discourse structure (graphs vs. trees) are very useful and can highly improve the results of automatic Arabic text summarization.

4.2 Cluster Based Approach

Many Arabic text summarization systems use clustering to generate a summary. For instance, [24] proposed an Arabic single and multi-document summarization approach

based on automatic sentence clustering and an adapted discriminate analysis method. Their system uses a clustering algorithm to group similar sentences into clusters. The proposed approach takes advantage of term's discriminate power to score sentences.

In the same context, [25] proposed a model based on document clustering and key phrase extraction. The model used a hybrid clustering (partitioning and k-means) to group Arabic documents into several clusters, then it extracts important key-phrases from each cluster. The model reached good results for single and multi-document summarization but no comparison with other systems was achieved.

Unlike the previous presented systems, [26] uses clustering to group words with the same root in the same cluster. The number of words in that cluster determines the weight of each word in the cluster. Then the score of each sentence is calculated based on several features. Sentences having the highest score are selected to be included in the final summary.

Finally, in [27], the authors investigate the use of clustering in Arabic multi-document summarization and for redundancy elimination. To this end, the authors conducted two experiments. In the first one, K-means algorithm is used to cluster sentences. More precisely, a number of sentences are selected randomly as the initial centroids, and then all sentences are assigned to the closest cluster based on their cosine similarity measure. To produce the summary, two selection methods are used: In the first method, the first sentence of each cluster is selected, while in the second one, all sentences in the biggest cluster are selected. For the second experiment, sentences selection is carried out before the clustering, and only the first sentence from each document and the most similar sentence are selected. Then, all the subsequent steps are similar to the first experiment. For evaluation, DUC-2002 dataset and an Arabic parallel translation version are used. Evaluation results are compared with the best five systems in the DUC 2002 competition. The proposed summarizers achieved the best scores when comparing ROUGE-1 results.

4.3 Machine Learning Based Approach

In the machine learning based approach, the summarization process is formulated as a binary classification problem. A set of training documents and their references summaries are required. Sentences are classified based on statistical features as summary or non-summary sentences.

Several Arabic summarization systems have been adopting machine learning and statistical techniques. For instance, in [28] the authors integrate Bayesian and genetic programming (GP) classification methods to generate summaries, using a reduced set of features. The system uses manually labelled corpora for training. Experiments show that Bayesian classifier tends to have high recall unlike GP classifier, which has a high precision. When combining both classifiers, the authors found that the recall and the summary size are increased, but when using the intersection of the two classifiers, the precision is increased and the summary size is decreased.

Later, in [29], the authors investigate the use of several classification methods including: probabilistic neural network (PNN), genetic algorithm (GA), Gaussian mixture model (GMM), feed forward neural network (FFNN), and mathematical regression (MR) for automatic text summarization task. The authors proposed a

trainable summarizer that use ten features such as sentence centrality, position, keywords, sentence similarity to the title, etc. The authors investigate, at first, the contribution of each feature on the summarization process. Then all features are used to train the previously mentioned methods on a manually created corpus, in order to obtain features weights. After that, the models are used to rank sentences in the testing corpus. Highest-ranking sentences are selected to produce summaries. Numerous experiments were also performed using DUC 2001 corpus. The obtained results indicated that GMM model is the best.

In the same context, [30] use support vector machine (SVM) algorithm in their system to produce summaries. The authors use eight statistical features among which sentence position, title keyword, indicative expression, TF-IDF score, etc. Only 60 Arabic documents are used in their experiments. The preliminary results published are encouraging (F-measure = 0.991). However, the authors could have extended their evaluation on a larger corpus to prove the effectiveness of their approach.

Recently, [31] proposed a supervised approach using AdaBoost to produce Arabic extracts. The authors use a set of statistic features such as overlap with word title, sentence position, sentence length, etc.

After building the AdaBoost learning model, all features are extracted from each sentence in the input document (to be summarized). Then, the features vectors are passed to the AdaBoost classifier, which decides whether their corresponding sentences should be included in the summary. The authors use a manually created corpus. The performance evaluation in term of F-measure is compared to those obtained using j48 decision trees as well as multilayer perceptron (MLP). The obtained results indicate that the proposed model outperforms multilayer perceptron and j48 decision trees.

4.4 Graph Based Approach

In the graph-based approach, the document is represented in the form of undirected graph. For every sentence, there is a node. An edge between two nodes is drawn if there is a relation between these two nodes. A relation can be a cosine similarity above a threshold, sharing a common word, or any other type of relationships. After drawing a graph, it is possible to view the sub-graphs of connected nodes as a cluster of distinct topics covered in the document.

Recently [32] proposes a graph-based approach for Arabic document summarization. In this approach, each document is represented by a weighted directed graph, whose nodes correspond to document sentences, and edges weights correspond to similarity between sentences. This similarity is determined by ranking the sentences according to some statistical features. The summary is extracted by finding the shortest path between the first and the last nodes in the graph considering the user compression ratio. Evaluation is done using EASC corpus, and intrinsic methods.

4.5 Textual Entailment Based Approach

Textual entailment has been introduced as a general framework for modelling semantic variability in several NLP tasks. An entailment relation consists in determining whether the meaning of one sentence can be inferred by another one [4]. The summary obtained

by using entailment inferences only includes sentences that are not entailed by any of the sentences in the previously accumulated summary.

Very little research has been done to combine Arabic text summarization and text entailment to produce extracts. In a single case [33], the authors tackle the problem of developing Arabic text summarization system (LCEAS), that produces extracts without redundancy. Lexical cohesion is applied to distinguish the important sentences from the unimportant ones in the text. As a result, poor information is removed from the text before applying the text entailment algorithm. In the next stage, cosine directional similarity method is applied to decide which sentences are not redundant. The text entailment algorithm suggested in [34] is enhanced to make it suitable for Arabic language. Performances evaluation of LCEAS are compared with previous Arabic text summarization systems. Results indicate that LCEAS outperforms the previous Arabic text summarization systems.

4.6 Ontology Based Approach

Arabic WordNet is a lexical database for Arabic. It clusters words into sets of synonyms called synsets, together with short general definitions called gloses, and determines the different semantic relations between these synonym sets.

Some researchers tend to use this lexical database in their systems. For instance, [12] presented a new query based Arabic text summarization system (OSSAD) using Arabic WordNet and an extracted knowledge base. Both Arabic WordNet and the domain specific knowledge base are used to expand the user's query. For summarization, the authors use decision tree algorithm. When comparing OSSAD generated summary against other Arabic summarization systems tested on the same data, the results show that OSSAD reach the best performances.

We end this section by Table 1, which presents a summary of the surveyed studies in chronological order.

Finally, it should be noted that, we can't compare the results obtained by these studies, because these systems are not evaluated using the same corpus and the same evaluation methods. As we can see in Table 1. In the majority of the surveyed researches, authors used their own corpus to evaluate their systems. This is due to the lack of publically available Arabic gold-standard summaries for several years.

5 Limitations of Extractive Approaches and Main Challenges

As we over mentioned, all the summarization approaches described in this paper are extractive. This means that sentences are selected from the input document to produce a summary. Unless a background repository is being used, the system is limited only to the words explicitly mentioned in the input text [5]. In machine learning based approach such as [28–30] other limitations appear. One limitation is ignoring relevant words that appear in abundance in the testing document but not in the training document, so the system lacks the ability to analyze such words, and it will treat them as unimportant words.

Table 1. A summary of the surveyed studies in chronological order

Research work	Year	Approach	Input	Evaluation method	Used corpus
[18]	2005	RST	Single document	Precision	Author's corpus
[28]	2006	Bayesian and Genetic programming	Single document	Recall, precision, F-measure	Authors' corpus
[29]	2009	GA, MR, FFNN, PNN, GMM	Single document	ROUGE-1, recall, precision	Authors' corpus
[19]	2009	RST	Single document	Not mentioned	Authors' corpus
[30]	2010	SVM	Single document	Precision, recall, F-measure	Authors' corpus
[27]	2011	K-means algorithm	Multi-document	ROUGE-1, recall, precision	DUC 2002 corpus and Arabic translated version
[26]	2012	Clustering	Single document	Recall and Precision	Authors' corpus
[20]	2012	RST	Single document	Precision, recall, F-measure, All ROUGE measures	Authors' corpus
[21]	2013	RST and Vector space model	Single document	Recall, precision, F-measure	Authors' corpus
[12]	2013	Ontology + decision tree algorithm (C4.5)	Single document	ROUGE-L	Authors' corpus and EASC corpus
[24]	2014	Clustering + mRMR	Single and multi-document	ROUGE-1 ROUGE-2	EASC, TAC2011 MultiLing Pilot corpus
[25]	2014	clustering	Single and multi-document	ROUGE measures	EASC corpus
[32]	2014	Graph	Single document	Precision, recall, F-measure	EASC corpus
[31]	2015	AdaBoost	Single document	Precision, recall, F-measure	Authors' corpus
[23]	2015	SDRT	Single document	Precision, recall, F-measure	Authors' corpus, Arabic Treebank (ATB v3.2 part3)
[33]	2015	Text entailment and Lexical cohesion	Single and Multi-document	ROUGE-2, ROUGE-L, ROUGE-W, ROUGE-S, AutoSummEng	Authors' corpus and EASC corpus

Another limitation is the lack of detection for the implicit relationships between words in the input document. The ability to detect such relationships requires an external knowledge and an analysis module. Most of the Arabic text summarization approaches are affected by a similar limitation in the detection of concepts and the relatedness between them. We think that, this is due to the shortage of linguistic resources for Arabic language.

For discourse theories based approach, other challenges appear. For instance, identifying discourse units boundaries in Arabic texts is not an easy task. One possible reason is the irregular use of punctuation marks in Arabic texts.

Furthermore, Arabic discourse connectives are highly ambiguous. Indeed, we can easily find an Arabic discourse connective that can signal more than one discourse relation and in some cases has no discourse usage. This leads to several problems in discourse segmentation and even relations labeling. Taking as an example the connectives “و”. According to [35] this connective has six meaning, which can be classified into two classes called “fasl” and “wasl”. The first class includes the states where the connective is a good indicator to begin a segments (it has a discourse usage). This class contains: (1) “واو القسم” that means testimony, (2) “ورب” that means few or someone and (3) “واو الاستئناف” that simply joins two unrelated sentences. The second class includes the different states where the connector has no discourse usage, and it has no effect on the segmentation. This class contains: (1) “واو الحال” that introduces a state, (2) “واو المعية”, which means the accompaniment and (3) “واو العطف” that relates words or sentences.

Finally, we can say that determining the effective features that extract the main ideas from the input document and that cover all important themes is a greater challenge in extractive text summarization especially for Arabic language.

6 Conclusion

This survey paper is focusing on extractive Arabic text summarization approaches. We presented the most recent progresses and researches raised in this field.

At first, we described some basic notions related to automatic text summarization, and some salient characteristics of Arabic language, and then we presented the main approaches proposed in this field to generate Arabic extracts. Finally, we discussed the limitations of these approaches and the major challenges faced when dealing with such application.

As a conclusion, we can say that Arabic text summarization is still in its initial stage compared to works done in English and other languages. This is partially, due to the shortage of ANLP tools and the complex morphology of Arabic language.

References

1. Gupta, V., Lehal, G.S.: A survey of text summarization extractive techniques. *Emerg. Technol. Web Intell.* **2**(3), 258–268 (2010)
2. El-Shishtawy, T., El-Ghannam, F.: keyphrase based Arabic summarizer (KPAS). In: 8th International Conference on Informatics and Systems (INFOS), p. NLP-7. IEEE (2012)
3. Radev, D.R., Hovy, E., McKeown, K.: Introduction to the special issue on summarization. *Comput. Linguist.* **28**(4), 399–408 (2002)
4. Lloret, E., Ferrández, O., Munoz, R., Palomar, M.: A Text summarization approach under the influence of textual entailment. In: NLPCS, pp. 22–31 (2008)
5. Bawakid, A.: Automatic documents summarization using ontology based methodologies. Doctoral dissertation. University of Birmingham (2011)
6. Hasanuzzaman, H.: Arabic language: characteristics and importance. *The Echo. J. Humanit. Soc. Sci.* **1**(3), 11–16 (2013)
7. Keskes, I., Zitoune, F.B., Belguith, L.H.: Splitting Arabic texts into elementary discourse units. *ACM Trans. Asian Lang. Inf. Process. (TALIP)* **13**(2), (2014)
8. Shaheen, M., Ezzeldin, A.M.: Arabic question answering: systems, resources, tools, and future trends. *Arabian J. Sci. Eng.* **39**(6), 4541–4564 (2014)
9. Lloret, E., Palomar, M.: Text summarization in progress: a literature review. *Artif. Intell. Rev.* **37**(1), 1–41 (2012)
10. El-Haj, M., Kruschwitz, U., Fox, C.: Creating language resources for under-resourced languages: methodologies, and experiments with Arabic. *Lang. Res. Eval.* **49**(3), 549–580 (2015)
11. Luhn, H.P.: The automatic creation of literature abstracts. *IBM J. Res. Dev.* **2**(2), 159–165 (1958)
12. Imam, I., Nounou, N., Abdul Khalek, H.: An ontology-based summarization system for Arabic documents (ossad). *Int. J. Comput. Appl.* **74**(17), 38–43 (2013)
13. Lin, C.Y.: Rouge: a package for automatic evaluation of summaries. In: Moens, M., Szpakowicz, S. (eds.) *Text Summarization Branches Out, Proceedings of the ACL 2004 Workshop*, vol. 8 (2004)
14. Giannakopoulos, G., Karkaletsis, V.: AutoSummENG and MeMoG in evaluating guided summaries. In: *TAC 2011 Workshop, Maryland, USA* (2011)
15. Jones, K.S., Galliers, J.R. (eds.): *Evaluating Natural Language Processing Systems*. LNCS, vol. 1083. Springer, Heidelberg (1995). <https://doi.org/10.1007/BFb0027470>
16. Giannakopoulos, G.: Multi-document multilingual summarization and evaluation tracks. In: *Proceedings of the MultiLing 2013 Workshop on Multilingual Multidocument Summarization, Sofia, Bulgaria*, pp. 20–28 (2013)
17. Mann, W.C., Thompson, S.A.: Rhetorical structure theory: toward a functional theory of text organization. *Text-Interdiscipl. J. Study Discourse* **8**(3), 243–281 (1988)
18. AlSanie, W.: Towards a suitable rhetorical representation for Arabic text summarization. In Kotsis, G., Taniar, D., Bressan, S., Ibrahim, I.K., Mokhtar, S. (eds.) *iiWAS*, vol. 196, pp. 535–542. Austrian Computer Society (2005)
19. Mathkour, H.I.: A Novel rhetorical structure approach for classifying Arabic security documents. *Int. J. Comput. Theory Eng.* **1**(3), 195 (2009)
20. Azmi, A.M., Al-Thanyyan, S.: A text summarizer for Arabic. *Comput. Speech Lang.* **26**(4), 260–273 (2012)
21. Ibrahim, A., Elghazaly, T., Gheith, M.: A novel Arabic text summarization model based on rhetorical structure theory and vector space model. *Int. J. Comput. Linguist. Nat. Lang. Process.* **2**(8), 480–485 (2013)

22. Asher, N., Lascarides, A.: *Logics of Conversation*. Cambridge University Press, Cambridge (2003)
23. Keskes, I.: *Discourse analysis of Arabic documents and application to automatic summarization*. Doctoral dissertation, Université de Toulouse (2015)
24. Oufaida, H., Nouali, O., Blache, P.: Minimum redundancy and maximum relevance for single and multi-document Arabic text summarization. *King Saud Univ.-Comput. Inf. Sci.* **26**(4), 450–461 (2014)
25. Noori Fejer, H., Omar, N.: Automatic Arabic text summarization using clustering and keyphrase extraction. In: *International Conference on Information Technology and Multimedia (ICIMU)*, pp. 293–298. IEEE (2014)
26. Haboush, A., Al-Zoubi, M., Momani, A., Tarazi, M.: Arabic text summarization model using clustering techniques. *Word Comput. Sci. Inf. Technol.* **2**(3), 62–67 (2012)
27. El-Haj, M., Kruschwitz, U., Fox, C.: Exploring clustering for multi-document arabic summarisation. In: Salem, M.V.M., Shaalan, K., Oroumchian, F., Shakery, A., Khelalfa, H. (eds.) *AIRS 2011*. LNCS, vol. 7097, pp. 550–561. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25631-8_50
28. Sobh, I., Darwish, N., Fayek, M.: An optimized dual classification system for Arabic extractive generic text summarization. In the 7th Conference on Language Engineering, pp. 149–154 (2006)
29. Abdel Fattah, M., Ren, F.: GA, MR, FFNN, PNN and GMM based models for automatic text summarization. *Comput. Speech Lang.* **23**(1), 126–144 (2009)
30. Boudabous, M.M., Maaloul, M.H., Belguith, L.H.: Digital learning for summarizing arabic documents. In: Loftsson, H., Rögnvaldsson, E., Helgadóttir, S. (eds.) *NLP 2010*. LNCS (LNAI), vol. 6233, pp. 79–84. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14770-8_10
31. Belkebir, R., Guessoum, A.: A supervised approach to Arabic text summarization using adaboost. In: Rocha, A., Correia, A.M., Costanzo, S., Reis, L.P. (eds.) *New Contributions in Information Systems and Technologies*. AISC, vol. 353, pp. 227–236. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16486-1_23
32. Al-Taani, A.T., Al-Omour, M.M.: An extractive graph-based Arabic text summarization approach. In: *The International Arab Conference on Information Technology*, Jordan (2014)
33. Al-Khawaldeh, F., Samawi, V.: Lexical cohesion and entailment based segmentation for Arabic text summarization (LCEAS). *Word Comput. Sci. Inf. Technol.* **5**(3), 51–60 (2015)
34. Tatar, D., Mihis, A., Lupsa, D., Tamaianu-Morita, E.: Entailment based linear segmentation in summarization. *Int. J. Softw. Eng. Knowl. Eng.* **19**(08), 1023–1038 (2009)
35. Khalifa, I., Feki, Z., Farawila, A.: Arabic discourse segmentation based on rhetorical methods. *Int. J. Electric. Comput. Eng.* **11**(1), 10–15 (2011)