2020 IEEE7thInternational Conference on Engineering Technologies and Applied Sciences (ICETAS)

Natural Language Processing based Question Answering Techniques: A Survey

1st Ammar Arbaaeen
Department of Computer Science.
Faculty of Information and Communication Technology.
International Islamic University Malaysia
Kuala Lumpur, Malaysia
afarbaaeen@uqu.edu.sa

2nd Asadullah Shah
Faculty of Information and Communication Technology.
International Islamic University Malaysia
Kuala Lumpur, Malaysia
asadullah@iium.edu.my

2020 IEEE 7th International Conference on Engineering Technologies and Applied Sciences (ICETAS) | 978-1-6654-2317-5/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICETAS51660.2020.9484290

***Abstract*— The rapid development in the field of information science and the increase in the usage of information retrieval strategies have empowered users to retrieve more accurate information. The availability of information in a different format and across different has presented colossal difficulties for information retrieval using information retrieval techniques. In this paper, an attempt is made at highlighting the question and answering (QA) systems that provide users the platform to express their question using the Natural Language and also retrieve the response from such systems in Natural Language. The four important modules of any QA system are the Natural Language Question (NLQ) processing, document processing, passage processing, and answer processing. Fundamentally, most QA systems combined several techniques from other fields such as information retrieval, knowledge representation, and natural language processing to process NLQ and present the most insightful response based on the stored document. This paper gives a thorough review of the various survey on the QA systems frameworks, their methodologies, types, approaches, and challenges of QA systems.**

Keywords— Question Answering Systems, Question Answering Methodologies, Information Retrieval Natural Language Processing.

I. Introduction

The evolution of information and communication technologies has offered huge value information which is stored in the form of textual documents and data. The advancement in technology is transforming the principles of website architecture from documents (traditional database) to data (semantic data representation). It supports the development of a common framework that enables the users to share and use data across different application areas in a meaningful linking data format. Moreover, it also improves the development of data integration that allows users to manage, share, and extract knowledge contents. Semantic technology such as ontology provides knowledge representation where information is presented and stored in form of rich vocabulary with descriptions of content properties and relations of the domain in an intelligent machine-readable format. Data may consist of valuable information and knowledge to the user in a particular domain. Thus, content description and query processing techniques are essential keys to access information resources and retrieve accurate and concise results for user' query.

One of the notable branches of Artificial Intelligence is the Natural Language Processing (NLP) field which provides the interactions with human languages effectively. The major components of NLP include Natural language understanding and natural language generation. Natural language understanding uses essential techniques that are efficient enough to analyze phrases and terms in a sentence to identify meaningful relations. While on the other hand, natural language generation is used to process text prediction and probabilistic grammars according to the sentence syntaxes. Syntactical and semantic analysis are broadly used in QASs for the major type of NLQ analysis to understand the text in the input question. The syntactical analysis processes the NLQ and document to generate the parse trees and identify grammatical construction of terms in the NLQ. The generated syntactical trees are then used for further semantic analysis process to generate the possible meaning of terms in NLQ.

The emerging field of information retrieval based on question answering system (QA) involves the research and methods from NLP, Information Extraction (IE), Automatic Summarization, Knowledge, and Database Management. In QA applications, the user obtains a concise and most relevant answer of the NLQs from a stored document while in IR applications, the user uses a keyword to perform searches and these keywords serve as input and receive a relevant list of documents based on his query. QA has been attracting many researchers who gain a significant enhancement in different domains such as biomedicine, weather forecast, and tourism [1].

The IR process involves information retrieval based on keywords from different heterogeneous resources [2]. Therefore, it possesses a limitation to gain accurate results as a content meaning of user expression in the query. The user receives a list of relevant documents upon request of the query and he may find the desired information within the list. However, users may request a precise and comprehensive answer rather than a list of documents. Although, the literate showed that IR techniques have shown tremendous success in retrieving relevant information, but the users are still facing many complex challenges for the extraction of desired information in natural language.

The rest of this paper is organized as follows: Section 2 provides a comprehensive survey of QAS. Section 3 describes the general architecture and approaches of QAS. Section 4 presents the classification of QASs and question types are highlighted in section 5. Section 6 provides the challenges of QAS, and section 7 concludes the paper.

II. Related Work

The Research in NLP research was coined by Allen [3] which has been further explored by other researchers in [1,4,5]. The current research in this field is mainly categorized into shallow NLP and Deep NLP. The shallow-NLP (SNLP) is more relevant to partial processing instead of linguistic analyses [6-7]. It comprises several steps including tokenization, POS, chunking, and shallow semantic analysis which are applicable at syntactic and semantic levels [8]. On the other hand, the deep-NLP (DNLP) is more related to semantic and contextual information of the sentence and enables the user to analyze syntactic and semantic of the sentence deeply [9-10]. DNLP is composed of many steps including full syntactic parsing, relation detection, and logical inference. Moreover, it generates many possible interpretations of the sentences which are used to rectify the ambiguity and uncertainty issues [11].

The early work on automated QA was evolved to facilitate the user by providing an interface to structured databases [12]. This work involves several steps in which NLQs are solved by converting them into a formal query, then comparing this query with the structured database, and finally returning the result as a response to the user. In [13], the authors introduced the QA system in which NLQ is processed by IR based on different techniques such as Boolean logic and regular expressions for retrieval of information. Finally, the most appropriate answers are mined from those retrieved documents based on their ranking strategy. In this technique, the quality of retrieved results and NLQs complexity directly affects the performance of the QA system [14]. In [15], the author introduced the testing and evaluation systems that use the F-measure method to assess the quality of output based on precision and recall.

In the early stages of development, [16] introduced the AnswerBus system which utilizes the capabilities of known search engines like google and yahoo to find the most relevant pages. They use the bag of words technique to find the most related page but are unable to find the exact answer to the question. In [17], the author not only analyses the user query but also generates several answers based on simple morphological variations. These answers are generated using n-gram frequencies to provide the most appropriate answer. In [18], the authors used various knowledge sources to develop knowledge-based representation which is used to answer the factoid and open-ended questions by matching structured data.

Recently in [19], the author classified the QA process into five major steps that include question analysis, phrase mapping, disambiguation, query construction, and querying distributed knowledge. This approach uses the user query to extract the named entities and their dependencies, find string and semantic similarities between the query and knowledge base, resolve mapping and segmentation ambiguities, convert the questions into query and answer using multiple knowledge databases. In [20], the author proposed a semantic QA system that focuses on the lexical gaps between words and ambiguity. Several challenges were also observed by the author includes multilingualism in web documents, complex questions, distributed knowledge, procedural, temporal, and spatial questions, and maps these questions to templates.

In [21], the author proposed a method to learn about the cue expressions from social questions-answers. Such cue expressions help in increasing the relevance of the answer. In [22], the authors used the maximum entropy model for the formation of a training set based on search queries for the retrieval of the best answer. Moreover, they also proposed question-like search queries into a set of twenty-six pre-defined semantic categories. In [23], the author proposed a methodology based on a machine-learning algorithm to identify the question class of two types of questions i.e. why and how. Finally, the answer is extracted by returning the starting and ending points of sentences or clauses which shows 90% extraction accuracy.

In the QA system, the most important task is to classify the question and generate the appropriate answer. In [22], the author identified some features such as syntactic and semantic from the query and used these features to identify the semantic class of the question. In [24], the author proposed genetic programming techniques to extract new features from the query. These new features were discovered based on lexical, syntactic, and semantic properties of the sentence. Finally, they also proposed the method to answer the definitional and factoid questions from paragraph and sentence based on the rank of the feature set.

In [25], the author proposed the method to convert the user query into a logical form by utilizing the knowledge of the Freebase database. The logical form contains both the lexical and logical features utilized to identify the category of the question, topic, entity type, and predicate. They proposed the method in which each term is linked with its feature which is used to compare the similarities with the question for the generation of the most appropriate answer. The results on the data set of 2000 question-answer pairs show 75% precision.

The recent research on NLP showed that deep learning resolves problems of complex linguistic processing and feature engineering. These techniques encode the semantic knowledge of documents and questions related to it and produce an answer based on these encodings. In [26], the author proposed a method that utilized the features of a conventional neural network (CNN) on bigram word feature to formulate the problem as a binary classification problem. The results showed that their technique achieved 0.71 MAP (mean average precision) and 0.78 MRR (mean reciprocal rank) scores. In [25] the author used the WikiQA dataset as a benchmark in CNN architecture along with the features of common word counts in the question and answer generates the best results.

In [27], the author used bidirectional long short-term memory (LSTM) network for the selection of answers to the NLQ. They have transformed the questions and answers into distributional representation and a similarity matrix is used to find the similarities between the question and their respective answer. In [28], utilized the concepts of dependency tree recursive neural network to extract question and answer representations jointly.

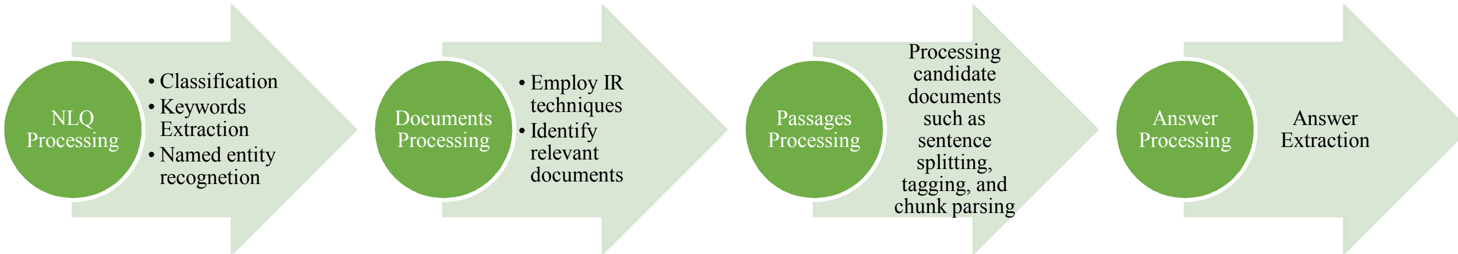

Fig. 1. The general architecture of QA systems

Finally, the representation of each word is fed to the network along with its dependent word and their dependency relation for optimal results. In [29], transformed the questions and QA knowledge base into vectors embeddings and used the scoring function to find the similarity between question and knowledge base triplets. The results of WikiAnswers questions on the ReVerb knowledge base showed a performance of 0.60–0.73 F-measure. In [30], the author utilized the multicolumn technique in a deep learning framework for accurate answer solutions. In [31], the author employed a dynamic memory network for best answer selection. Similarly, [32] also used the CNN architectures for accurate answer selection.

III. QUESTION ANSWERING SYSTEMS

The basic architecture of a QA system consists of four major modules: NLQ processing, documents processing, passages processing, and answer processing which are extracted from either structured databases, semi-structured databases, or unstructured databases [33]. QA systems commonly follow the standard pipeline processing structure as illustrated in Figure 1. The NLQ processing module is responsible for analyzing received NLQs and to understand and identify, what is being asked for? Analyzing and classification of the NLQs is performed based on different approaches which range from keyword extraction to deep linguistic-based approaches [2]. They consider the syntactic and semantic of an NLQ to understand the purpose of that particular NLQ. As a result of the analysis, the extracted keywords are used by IR as the query term for retrieving candidate and relevant information in the document processing module. This relevant information is further analyzed in the passage processing task using methods like sentence splitting, part-of-speech (POS) tagging, and parsing. Finally, the answer processing module uses the result of the analysis to extract the concise precise answer based on the NLQ's representation and ranking of candidate answers.

The goal of a QA system is its ability to retrieve users' exact answers to the questions that are posed using natural language. An integral component of the QA system is the capability of the QA system to understand the question and the context for which the question is generated. Due to the dynamic nature of natural language, developing a QA system has been classified has been a challenging task [34]. One of the usual tasks in every QA system is the analysis of the user question using different techniques, intending to have an overall meaning of the question. The methods that are adopted in the literature include Part Of Speech (POS) tagging, parsing, and the extraction of candidate answer type based on the posed question [35]. There are lots of research works that have adopted these methods however, categorizing the QA approach is very hard even though some literature has attempted in classifying the approaches in the QA system.

In [36], an overview of the actual state in the QA system was captured in the conducted survey. The survey gave an insight into every possible tool that are available for designing a novel QA system and giving a detailed account of the evaluation metrics alongside the various benchmarks. On the other hand, the author in [37] captures various problems in the QA system focusing on the various challenges that are associated with the QA system in the field of Natural Language Processing (NLP) and describing the possible solution based on literature that has provided possible solutions. Moreover, an accurate analysis of the state of the art of the QA system was provided in [38]. The survey listed all the systems that have been proposed for all the benchmarks. The overall structure of the QA system was captured, and each component was further analyzed to describe the various approaches that have been employed and outline the system that uses the benchmarks.

In this section, we attempt to capture the state-of-the-art approaches in the QA systems in literature based on the recent advances in the NLP. These include a rule-based approach, statistical approach, machine learning approach, deep learning approach, and knowledge base QA system.

A. Rule-Based Approach

The rule-based approach is one of the initial methods of the QA system which depends on the rules that are devised from grammatical semantics to determine the appropriate answer for a given question. The set of rules in the rule-based approach is handcrafted and they are dependent on the lexical and semantic insight of a given context [39]. The rule-based QA system is a logical representation of decision trees were the linguistic structures that mirror the human understanding of the text based on grammatical rules. The QA system determines the paths for every instance of question-based on the question type. The path taken for the two types of questions may be different based on the answer extraction process. The POS tagging and Named Entity Recognition were incorporated into the rule-based QA system to improve the answer matching [40]. The rule-based approach to the QA system was a good initiative at developing a good QA system but they have some limitations. One of the notable limitations of the rule-based QA system the fact that the heuristic rules need to be manually hand-crafted all the time. The rule definition required a detailed understanding of the language semantics. The QA system requires that new rules be introduced based on the question type. The introduction of new rules makes the system process to be cumbersome and it is also a reflection that the QA system will not be suitable for a domain that has a high level of volatile data [41]. With an increase in the growth of text-based materials online, the need for a statistical approach to the QA system has increased.

B. Statistical Approach

The limitation of the rule-based QA system brought above the need for a self-learning approach that will help at solving the extensibility challenge. The statistical approach to the QA system requires the formulation of a hypothesis that forms the basis for model development. The statistical approach brings about techniques that can deal with a large amount of data as well as their diverseness [42]. The statistical approach requires a huge amount of satisfactory data to perform definite precise learning. When the data are properly learned there is a very huge possibility of producing a better result. The statistical method can be transferred to a new domain and it is independent of the choice of programming language. Statistical techniques have been applied to different stages of the QA system. Among the notable models that use the statistical technique are the Support Vector Machine (SVM) classifier, Bayesian classifier, and Maximum entropy model. These classifiers have been used for the classification of questions. The statistical models are used to analyze questions for prediction purposes based on users' expected answer type. Annotated documents or a corpus of questions are trained using the models. The use of the statistical technique has brought about a significant improvement in POS tagging for the Sinhala Language [43].

C. Machine Learning Approach

The rapid advancement in the machine learning system allows for the capacity to be able to determine the direction of the data [44]. This brought about the self-induced learning of the QA system. The system uses the training set to build a knowledge base (taxonomy) and use the knowledge to answer the questions. This promotes self-learning capabilities that are not available in the rule-based approach. The statistical technique pave way for the machine learning approach with its ability to analyze an annotated corpus and then build a knowledge base. Machine learning algorithms in the QA system can understand linguistic features implicitly. The machine learning approach is highly scalable as long as there is enough training set. The unique feature of the machine learning approach, when compared with other QA systems, is its learnability. Domains such as Journalism and Judiciary have a huge amount of data and machine learning has become very important in QA. The QA systems in this domain leverage on the speed and coverage factors based on the extensibility of the machine learning approach but these features are not present in the rule-based approach. Recently, the machine learning approach has been combined with statistical approaches in linguistic and sentiment related domains [45]. The linear SVM is usually used in training systems that have strong features and sentiment lexicons. The ensemble approach in machine learning which combines several machine learning algorithms has also been proven to be powerful in the QA system. The approach can calculate the average confidence score of every classifier for positive, negative, and neutral classes [45].

D. Deep Learning Approach

The deep learning approach can automatically discover the representations needed for prediction, detection, or classification when the machine is fed with raw data [46]. There is a distinct difference between deep learning and machine learning because deep learning can learn underlying features in data using a neural network. Every standard neural network is made up of simple, connected nodes/units called neurons. The input neurons in the network are activated by sensors based on the perceived environment, and some other neurons are activated through weighted connections from previously active neurons. In recent times, domains such as natural language processing, computer vision, machine translation, and text summarization have gained significant success in the deep learning domain [47]. The logical representation of the textual context is mapped by a neural network which is later used for answer predictions. The neural network uses a bidirectional Long Short-Time Memory (LSTM) unit which is a variant of the Recurrent Neural Network (RNN) for question processing and answers classification [48]. In Natural Language Processing, the Recurrent Neural Network has shown some high level of significance. The relationships that existed within the hidden layers and the rest of the previous layer is what makes the RNN different form the traditional neural networks. The ability of the RNN to model long span dependencies based on its recurrent property [48]. This enables the network to cluster neurons with similar histories to efficiently represent patterns with variable length [49]. The recurrent networks also have their challenges.

The challenges of RNN are gradient descent and gradient explosion. In mathematical terms, RNN has a challenge with information representation and generalization ability when dealing with complex data structures. One of the notable techniques in overcoming the gradient descent and gradient explosion is the LSTM which was designed to explicitly avoid the long-term dependency problem. Remembering information for long periods is practically the default behavior of LSTM, not something they struggle to learn. Another technique is the Backpropagation Through Time (BPTT). In BPTT, the network errors are propagated through recurrent connections back in time for a specific number of time steps [50]. With the backpropagation, the network learns to remember information for several time steps in the hidden layer.

E. Knowledge Base (KB) QA System

In the Knowledge-Based Question and Answering System, the retrieval of answers to every instance of natural language question involves the mapping of a query over an ontology [51]. The logical form that is derived from the mapping process is used at retrieving facts from the databases. The data source of the knowledge base can be any complex structure, for instance, specific facts or geospatial readings, that require complex SQL queries. Mapping a user asked query is performed by semantic parser [52]. The goal of presenting a knowledge-based approach is to provide the possibility of selecting an appropriate

QA system for a specific kind of problem. The choice of an ontology of a QA system should ensure freedom in requirements definition processes. The ontology should also provide the systematization of the knowledge of the QA system domain and promote time reduction for QA system selection [53]. A detailed understanding of domain knowledge is an important task in setting up a QA system [54]. Knowledge categorization is divided into two which include, general knowledge and specific knowledge. General knowledge is a form of knowledge that has been aggregated over time through various means and measures. This form of knowledge is regarded as a description of the world. The specific knowledge entails the description of a specific linguistic concept as it related to a specific domain [55].

Among notable activities in knowledge, the base system is the extraction of specific information that is later used in the system. The availability of vast knowledge in different domains provides the opportunity to use these vast amounts of knowledge and other available resources in the QA systems. In the development of a knowledge-based system, the users developed some level of preferences that the system will poss. A reasoning mechanism is integrated into the knowledge-based to provide a set of results and compute definitions which helps in fulfilling the predefined requirement of the knowledge base system. At the output side of the knowledge base system, the user obtains a set of results that matches the set of queries that have been passed to the knowledge-based system (QA ontology). The best approach to a knowledge base system is the construction of an ontology for a QA system [56]. The domain of the ontology is relatively small and the analyzed data does not require high-level automation.

IV. CLASSIFICATION OF QUESTION ANSWERING SYSTEMS

The research in the field of QA clearly showed that numerous factors including application domain, type of question, type of answer, knowledge source, type of analysis, and methods utilized affect the QA system. This section highlights the categorization of QAS from an application domain perspective.

A. *Web-based Question Answering system (open domain)*

This answering system is not limited to answer the question related to a single domain rather they can provide a short answer for every type of question. Since the web is the biggest source of information, the utilization of a web-based question answering system along with universal ontology enables the users to get answers about almost everything. Open-domain Question Answering systems are not restricted to any specific domain and provide a short answer to a question, addressed in natural language. [57]. These web-based QA systems utilize the general vocabulary instead of domain-specific vocabulary. Moreover, users can even create a question without any domain-specific knowledge. This system mostly utilizes the knowledge base of sources like Wikipedia as an information source which results in low quality of the answer. These systems are more dependent on basic ontology and world data [58]. Moreover, these systems are domain-independent and can answer questions related to any subject. These systems are very suitable and useful for many casual users.

B. *Restricted Domain Question Answering systems (Closed domain)*

The restricted domain question answering system is the most appropriate system for domain experts requiring specific information in form of answers for the questions. These QA systems are restricted to a specific domain and answer the question only to that specific domain. Moreover, these systems have limited repository and require linguistic support for understanding the text in natural language to answer questions efficiently. As the answers to questions are searched from the particular domain document collection, the quality of the answer is accurate. This versatile QA system enables users to retrieve accurate answers from structured databases, unstructured data, and semi-structured databases [59]. These systems utilize domain-specific ontologies and terminologies to retrieve the answers to their questions. The literate clearly showed that the usage of closed domain QA systems is developed for medical domain QA systems, temporal domain QA systems, geospatial domain QA systems, community-based QA systems, and patent QA systems. This system is very useful and efficient for domain expert users who demand specialized answers for specific questions. In [60], the author proposed a QA system for baseball players which provides answers to the question which are asked about one session baseball data.

When the query is made by the user in form of a question, the NLP analyses the question linguistically by integrating deep and shallow components such as named entity recognition and parser. The NLP is also responsible for the generation of semantic representation and question object that contains the proto query. This proto query is an implementation-independent higher-level representation of a database that is used to construct the instance of a specific database or ontology query. The query information source generates an answer object which results in natural language answer generation [61]

V. CLASSIFICATION BASED ON TYPES OF QUESTIONS

This section highlights all types of possible questions in QAS. The classification of each question according to its type in natural language for producing the appropriate answer is discussed in detail. The NLQs are classified into six major categories which include factoid questions, hypothetical questions, confirmation questions, causal questions, and complex questions. All these types of questions are discussed in the next subsections.

A. *Factoid type questions [what, which, when, who, how]*

The questions which start with wh-word are classified as factoid question. All the questions starting with wh-word are fact-based and very simple to answer in a single sentence or short phrase. For example the factoid type question "What is the capital of Saudi Arabia?" inquiries about the city name of a country and it will narrow down the search space for the name of the capital city of Saudi Arabia. Previous literature showed that the answer to factoid questions are generally named entities [1]. These types of questions are usually quite straightforward and do not require complex and large repositories to answer simple questions. Moreover, they do not require complex natural language processing to retrieve the answer from simple repositories which will result in satisfactory performance in answering. Generally, factoid type questions are a large repository of questions. The previous literate showed that the only hurdle to answer factoid questions is related to their identification and their sub-classification in the Question Answering system. The answers to factoid questions are usually short phrases or sentences such as date, time, location, person, country, etc.

B. List type questions

This type of question requires a list of facts or entities as answers; for example, a list of iPhone models in 2020. The example clearly showed that these types of questions have answers in form of a list of named entities which provide good accuracy for answering the query. This type of question can be easily answered by a simple answering system and does not require deep natural language processing techniques to answer. The previous literature in this field showed that all techniques which are applied to factoid type questions can work perfectly for list type questions [21]. The only complexity that occurred during answering this type of question is related to setting the threshold value for the quantity of the entity or the number.

C. Confirmation Questions [yes or no]

This type of questions is quite simple and straight forward as the possible answer is either yes or no. For example, “Is iWatch waterproof” is a confirmation type question that requires either yes or no as an answer. This type of question can be answered accurately and concisely when the Knowledge base has world knowledge, inference mechanism, and common-sense reasoning. The previous literature clearly showed that one of the major advantages of this type of questions asked in QA systems builds a new knowledge base whenever some expert users enquire for information that needs common sense reasoning and world knowledge [62]. The only drawback or limitation in answering this type of question is the requirement of a higher level of knowledge and complex retrieval techniques that are not mature enough and are still under development phase.

D. Causal Questions [why or how]

Casual questions ask about an entity and require a detailed answer which provides a complete description of that entity in form of a sentence, paragraph, or even whole document. Usually, these questions are asked by users who expect answers as reasons, descriptions, elaborations, etc related to particular objects or events. e.g., Why night is dark? Therefore, to answer this type of question, the user requires descriptive answers with reasons expanding from paragraphs to a whole document [23].

E. Hypothetical Questions

Hypothetical questions are the most complex type of question as they require information associated with any hypothetical event. Therefore, they do not have specific answers. For instance, the hypothetical question “what would happen if all oceans dry” do not have any specific answer because of a hypothetical event. These types of complex questions are hard to answer due to which the reliability and accuracy of these questions are low and depend upon users and context. The previous literature clearly showed that the expected answer type is spread for hypothetical type questions which result in less accuracy in answering these questions [60]

F. Complex Questions

Finally, the most difficult question to answer is known as complex questions which usually contain a list of "nuggets" that requires many references or documents to answer. For instance, the question “what is the possible usage of computer in our life?” often requires gathering and synthesizing information from multiple documents to generate multiple nuggets as answers. Usually, these complex questions require complex procedures and methods to answer complex questions. The literature showed that complex questions are composed of multiple independent questions and each question searches for many documents and knowledge bases to find the appropriate answer [62].

VI. Challenges of QA System in Research

One of the notable challenges confronting the QA system is the ability to scale large data. QA systems are expected to be robust in handling heterogeneous and conflicting data that are collected from multiple sources [63]. Ontology and the semantic web are the key drivers of the QA system, ontology is an integral component of the semantic web, it enables knowledge sharing and exchange, and research has shown that it is becoming crucial for mythology is representing specific domain knowledge to provide the required semantic support for a QA system [64]. The NLP techniques are used by the QA system and an ontology as input to process a given question. The QA system searches the knowledge base for the required information in the query to identify the answer and present the answer to the user in a structural and non-structural data collection [65]. The QA system does not require that user to learn all the entire vocabularies that are required in querying the ontology.

There are lots of challenges in adopting Semantic Web technologies for web data. It is very easy to point at the various unique challenges that are largely overlooked by the Semantic Web, the challenges are inconsistency, scale, unreliability, and disruptions [63]. The understanding of the natural language question is an important component of the QA system. In recent times, QA systems have taken a key role in the search engine optimization concept. From a technology point of view, the building blocks to QA systems are statistical or natural language processing, knowledge representation, information retrieval, and reasoning [56]. The QA system's potential building blocks involve text classification, information extraction, and summarization technologies.

Another challenge of the QA system involves answering a question. It is expected that a search engine should analyze a given question in the context of ongoing interaction. It is expected that a QA system when presented with a question should find one or more answers that are related to the question by consulting on links in the QA system, then it is expected that a perfect answer should be presented to the user in an appropriate form [65]. Another challenge is the number of processes of searching and querying is massive and largely heterogeneous. Existing approaches to querying semantic data have numerous difficulties in measuring their model effectiveness in dealing with the ever-increasing amount of distributed semantic data available online [66].

VII. Conclusion And Future Directions.

This paper provides a comprehensive detailed survey of emerging Question Answering systems and also presents the general architecture of this system. The paper discusses in detail various types of Question Answering systems including open and closed domain question answering systems. Moreover, the paper also describes different categories of questions posed by humans in natural language. The extensive literature review showed that various researchers worked on a Question-Answering system with different languages but still the extraction of accurate information to NLQ remains challenging in the QA system. One of the major challenges faced by all previous techniques is that they are not capable to deal effectively with semantic issues which results in ambiguity in NLQ that would cause failure in returning accurate results. We have also observed that all the ambiguities cannot be easily recognized, and they require deep linguistic and expression analysis to define the accurate meaning of these ambiguities.

The previous literature clearly showed that a lot of work still needs to be done in the field to achieve full reasoning capability. Moreover, the need of the hour is to look at ways to solve ambiguous questions as they are very common in real-life scenarios. There is also a need to devise a model that can serve as a multilingual question answering system to handle the queries in all languages. Besides this, the developments in the field have been rapid and there is a need to carefully analyze all that is going on. The literature clearly showed that some hidden factors affect the performance of QASs that includes Psychology and the skill of the user who is asking the question etc. The need of the hour is to develop future QASs that can perform knowledge survey tasks to retrieve the results which could satisfy the needs of the customer. These dialogs based QASs are being developed to understand the need of users. Finally, there are requirements of intelligent QASs that can track the browsing history and behavioral activities of users and present answers to questions in a more effective manner.

References

[1] O. Al-harbi, S. Jusoh, and N. Norwawi, “Lexical Disambiguation in Natural Language Questions (NLQs),” *Int. J. Comput. Sci. Issues*, vol. 8, no. 4, pp. 143–150, 2011.

[2] H. Hu, “A Study on Question Answering System Using Integrated Retrieval Method,” pp. 28–36, 2006.

[3] J. Allen, “Natural language understanding. Pearson,” 1995.

[4] U. Schäfer, “Integrating Deep and Shallow Natural Language Processing Components – Representations and Hybrid Architectures, Diss.,” 2006.

[5] D. S. Wang, “A domain-specific question answering system based on ontology and question templates,” *Proc. - 11th ACIS Int. Conf. Softw. Eng. Artif. Intell. Netw. Parallel/Distributed Comput. SNPD2010*, pp. 151–156, 2010.

[6] U. Hermjakob, “Parsing and Question Classi cation for Question Answering,” 2001.

[7] M. Lease, "Natural language processing for information retrieval: the time is ripe (again)." In *Proceedings of the ACM first Ph. D. workshop in CIKM*, pp. 1-8. ACM, 2007.

[8] S. P. Abney, “Parsing By Chunks,” pp. 257–278, 2011.

[9] P. Zweigenbaum, “Question answering in biomedicine,” *Proc. 10th Conf. Eur. Chapter Assoc. Comput. Linguist.*, vol. EACL2003, pp. 1–4, 2003.

[10] F. Benamara, “Cooperative question answering in restricted domains: the WEBCOOP experiment,” *Work. Quest. Answering Restricted Domains. 42nd Annu. Meet. Assoc. Comput. Linguist.*, pp. 31–38, 2004.

[11] D. Jurafsky, R. Ranganath, and D. McFarland, “Extracting social meaning: Identifying interactional style in spoken conversation,” *NAACL ’09 Proc. Hum. Lang. Technol.*, no. June, pp. 638–646, 2009.

[12] M. R. Joshi and R. A. Akerkar, “Algorithms to improve performance of natural language interface,” vol. 5, no. 2, pp. 52–68, 2003.

[13] L. Hirschman and R. Gaizauskas, “Natural language question answering: The view from here,” *Nat. Lang. Eng.*, vol. 7, no. 4, pp. 275–300, 2001.

[14] D. Moldovan, M. Paşca, S. Harabagiu, and M. Surdeanu, “Performance issues and error analysis in an open-domain question answering system,” *ACM Trans. Inf. Syst.*, vol. 21, no. 2, pp. 133–154, 2003.

[15] E. M. Voorhees, “Overview of the TREC 2003 Web Track.,” *Trec 2003*, vol. 3, pp. 1–15, 2003.

[16] Z. Zheng, "AnswerBus question answering system." In *Proceedings of the second international conference on Human Language Technology Research*, pp. 399-404. Morgan Kaufmann Publishers Inc., 2002.

[17] E. Brill, S. Dumais, and M. Banko, “An analysis of the AskMSR question-answering system,” no. July, pp. 257–264, 2007.

[18] B. Katz, “Annotating the World Wide Web using Natural Language Introducing S-rules,” *Proc. 5th RIAO Conf. Comput. Assist. Inf. Search. Internet (RIAO ’97)*, pp. 136–159, 1997.

[19] D. Diefenbach, P. Maret, V. Lopez, and K. Singh, “Core Techniques of Question Answering Systems over Knowledge Bases: a Survey QALD-QA benchmarks View project PowerAqua View project Core Techniques of Question Answering Systems over Knowledge Bases: a Survey,” *Knowl. Inf. Syst. Core*, pp. 1–41, 2017.

[20] M. Sabou *et al.*, “Survey on Challenges of Question Answering in the Semantic Web,” *Semant. Web*, vol. 0, pp. 1–26, 2016.

[21] Y. Wu, C. Hori, H. Kashioka, and H. Kawai, “Leveraging social Q&A collections for improving complex question answering,” *Comput. Speech Lang.*, vol. 29, no. 1, pp. 1–19, 2015.

[22] A. Figueroa and G. Neumann, Context-Aware Semantic Classification of Search Queries for Browsing Community Question-Answering Archives, vol. 96. 2016.

[23] C. Pechsiri and R. Piriyakul, “Developing a Why–How Question Answering system on community web boards with a causality graph including procedural knowledge,” *Inf. Process. Agric.*, vol. 3, no. 1, pp. 36–53, 2016.

[24] I. Khodadi and M. S. Abadeh, “Genetic programming-based feature learning for question answering,” *Inf. Process. Manag.*, vol. 52, no. 2, pp. 340–357, 2016.

[25] M. C. Yang, D. G. Lee, S. Y. Park, and H. C. Rim, “Knowledge-based question answering using the semantic embedding space,” *Expert Syst. Appl.*, vol. 42, no. 23, pp. 9086–9104, 2015.

[26] L. Yu, K. M. Hermann, P. Blunsom, and S. Pulman, “Deep Learning for Answer Sentence Selection,” pp. 1–9, 2014.

[27] D. Wang and E. Nyberg, “A Long Short-Term Memory Model for Answer Sentence Selection in Question Answering,” pp. 707–712, 2015.

[28] M. Iyyer, J. Boyd-Graber, L. Claudino, R. Socher, and H. Daumé III, “A Neural Network for Factoid Question Answering over Paragraphs,” pp. 633–644, 2015.

[29] A. Bordes, J. Weston, and N. Usunier, “Open question answering with weakly supervised embedding models,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8724 LNAI, no. PART 1, pp. 165–180, 2014.

[30] L. Dong, F. Wei, M. Zhou, and K. Xu, “Question Answering over Freebase with Multi-Column Convolutional Neural Networks,” pp. 260–269, 2015.

[31] C. Xiong, S. Merity, and R. Socher, “Dynamic Memory Networks for Visual and Textual Question Answering,” 2016.

[32] M. Feng, B. Xiang, M. R. Glass, L. Wang, and B. Zhou, “Applying deep learning to answer selection: A study and an open task,” *2015 IEEE Work. Autom. Speech Recognit. Understanding, ASRU 2015 - Proc.*, pp. 813–820, 2016.

[33] D. Moldovan, et al. "The structure and performance of an open-domain question answering system." *Proceedings of the 38th annual meeting on association for computational linguistics*. Association for Computational Linguistics, 2000.

[34] Rodrigo, Alvaro, and Anselmo Penas. "A study about the future evaluation of Question-Answering systems." *Knowledge-Based Systems* 137 (2017): 83-93.

[35] Siciliani, Lucia. "Question Answering over Knowledge Bases." In *European Semantic Web Conference*, pp. 283-293. Springer, Cham, 2018.

[36] Unger, Christina, André Freitas, and Philipp Cimiano. "An introduction to question answering over linked data." In *Reasoning Web International Summer School*, pp. 100-140. Springer, Cham, 2014.

[37] Höffner, Konrad, Sebastian Walter, Edgard Marx, Ricardo Usbeck, Jens Lehmann, and Axel-Cyrille Ngonga Ngomo. "Survey on challenges of question answering in the semantic web." *Semantic Web* 8, no. 6 (2017): 895-920.

[38] Diefenbach, Dennis, Vanessa Lopez, Kamal Singh, and Pierre Maret. "Core techniques of question answering systems over knowledge bases: a survey." *Knowledge and Information systems* 55, no. 3 (2018): 529-569.

[39] Madabushi, Harish Tayyar, and Mark Lee. "High accuracy rule-based question classification using question syntax and semantics." In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pp. 1220-1230. 2016.

[40] Riloff, Ellen, and Michael Thelen. "A rule-based question answering system for reading comprehension tests." In *ANLP-NAACL 2000 Workshop: Reading Comprehension Tests as Evaluation for Computer-Based Language Understanding Systems*. 2000.

[41] Ishwari, K. S. D., A. K. R. R. Aneeze, S. Sudheesan, H. J. D. A. Karunaratne, A. Nugaliyadde, and Y. Mallawarrachchi. "Advances in Natural Language Question Answering: A Review." *arXiv preprint arXiv:1904.05276* (2019).

[42] Dwivedi, Sanjay K., and Vaishali Singh. "Research and reviews in question answering system." *Procedia Technology* 10 (2013): 417-424.

[43] Gunasekara, Dilmi, W. V. Welgama, and A. R. Weerasinghe. "Hybrid part of speech tagger for sinhala language." In *2016 Sixteenth International Conference on Advances in ICT for Emerging Regions (ICTer)*, pp. 41-48. IEEE, 2016.

[44] Cohn, David A., Zoubin Ghahramani, and Michael I. Jordan. "Active learning with statistical models." *Journal of artificial intelligence research* 4 (1996): 129-145.

[45] Muthutantrige, Supun R., and A. R. Weerasinghe. "Sentiment Analysis in Twitter messages using constrained and unconstrained data categories." In *2016 Sixteenth International Conference on Advances in ICT for Emerging Regions (ICTer)*, pp. 304-310. IEEE, 2016.

[46] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521, no. 7553 (2015): 436-444.

[47] Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." *Neural networks* 61 (2015): 85-117.

[48] Mikolov, Tomas, and Geoffrey Zweig. "Context dependent recurrent neural network language model." In *2012 IEEE Spoken Language Technology Workshop (SLT)*, pp. 234-239. IEEE, 2012.

[49] Mikolov, Tomáš, Stefan Kombrink, Lukáš Burget, Jan Černocký, and Sanjeev Khudanpur. "Extensions of recurrent neural network language model." In *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 5528-5531. IEEE, 2011.

[50] Williams, Ronald J., and David Zipser. "A learning algorithm for continually running fully recurrent neural networks." *Neural computation* 1, no. 2 (1989): 270-280.

[51] Jurafsky, Daniel, and James H. Martin. "Semantic role labeling." *Speech and Language Processing* (2015).

[52] Höffner, Konrad, Sebastian Walter, Edgard Marx, Jens Lehmann, Axel-Cyrille Ngonga Ngomo, and Ricardo Usbeck. "Overcoming challenges of semantic question answering in the semantic web." *Semantic Web Journal* (2016).

[53] Konys, Agnieszka. "Knowledge-based approach to question answering system selection." In *Computational Collective Intelligence*, pp. 361-370. Springer, Cham, 2015.

[54] Allahyari, Mehdi, Seyedamin Pouriyeh, Mehdi Assefi, Saied Safaei, Elizabeth D. Trippe, Juan B. Gutierrez, and Krys Kochut. "A brief survey of text mining: Classification, clustering and extraction techniques." *arXiv preprint arXiv:1707.02919* (2017).

[55] Ahsan, Imran, Wasi Haider Butt, Mudassar Adeel Ahmed, and Muhammad Waseem Anwar. "A comprehensive investigation of natural language processing techniques and tools to generate automated test cases." In *Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing*, pp. 1-10. 2017.

[56] Konys, Agnieszka. "An ontology-based knowledge modelling for a sustainability assessment domain." *Sustainability* 10, no. 2 (2018): 300.

[57] V. Sreelakshmi, J. Sangeetha "Open Domain Question Answering System Using Semantic Role Labeling," *Int. J. Res. Eng. Technol.*, vol. 03, no. 27, pp. 104–107, 2015.

[58] P. Biswas, A. Sharan, and N. Malik, "A framework for restricted domain Question Answering System," *Proc. 2014 Int. Conf. Issues Challenges Intell. Comput. Tech. ICICT 2014*, pp. 613–620, 2014.

[59] R. K. Harpreet Kaur, "Novel Scoring System for Identify Accurate Answers for Factoid Questions," *Int. J. Sci. Res.*, vol. 2, no. 9, pp. 294–297, 2013.

[60] A. Mishra and S. K. Jain, "A survey on question answering systems with classification," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 28, no. 3, pp. 345–361, 2016.

[61] I. Srba and M. Bielikova, "A Comprehensive Survey and Classification of Approaches for Community Question Answering," *ACM Trans. Web*, vol. 10, no. 3, pp. 1–63, 2016.

[62] [1] P. Tanwar, D. T., and D. Kamlesh, "An Effective Reasoning Algorithm for Question Answering System," *Int. J. Adv. Comput. Sci. Appl.*, vol. 4, no. 1, pp. 52–57, 2014.

[63] Hogan, Aidan, Andreas Harth, Jürgen Umbrich, Sheila Kinsella, Axel Polleres, and Stefan Decker. "Searching and browsing linked data with swse: The semantic web search engine." *Journal of Web Semantics* 9, no. 4 (2011): 365-401.

[64] Guo, Qinglin, and Ming Zhang. "Question answering system based on ontology and semantic web." In *International Conference on Rough Sets and Knowledge Technology*, pp. 652-659. Springer, Berlin, Heidelberg, 2008.

[65] Nalawade, Shekhar, Shivraj Kumar, and Deepak Tiwari. "Question answering system." *International Journal of Science and Research (IJSR)* 3, no. 5 (2014).

[66] Lopez, Vanessa, Victoria Uren, Marta Sabou, and Enrico Motta. "Is question answering fit for the semantic web?: a survey." *Semantic web* 2, no. 2 (2011): 125-155.