Journal of King Saud University – Computer and Information Sciences 32 (2020) 635–646

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

A literature review on question answering techniques, paradigms and systems

Marco Antonio Calijorne Soares *, Fernando Silva Parreiras

LAIS – Laboratory for Advanced Information Systems, FUMEC University, Av. Afonso Pena 3880, 30130 009 Belo Horizonte, MG, Brazil

ARTICLE INFO

Article history:
Received 26 January 2018
Revised 5 June 2018
Accepted 4 August 2018
Available online 12 August 2018

Keywords:
Question answering systems
Natural language processing
Information retrieval

ABSTRACT

Background: Question Answering (QA) systems enable users to retrieve exact answers for questions posed in natural language.
Objective: This study aims at identifying QA techniques, tools and systems, as well as the metrics and indicators used to measure these approaches for QA systems and also to determine how the relationship between Question Answering and natural language processing is built.
Method: The method adopted was a Systematic Literature Review of studies published from 2000 to 2017.
Results: 130 out of 1842 papers have been identified as describing a QA approach developed and evaluated with different techniques.
Conclusion: Question Answering researchers have concentrated their efforts in natural language processing, knowledge base and information retrieval paradigms. Most of the researches focused on open domain. Regarding the metrics used to evaluate the approaches, Precision and Recall are the most addressed.

Contents

* Corresponding author.
E-mail address: marco.calijorne@techsysinfo.com (M.A. Calijorne Soares).
Peer review under responsibility of King Saud University.

ELSEVIER Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2018.08.005
1319-1578/

1. Introduction

Question Answering (QA) systems have emerged as powerful platforms for automatically answering questions asked by humans in natural language using either a pre-structured database or a collection of natural language documents (Chali et al., 2011; Dwivedi and Singh, 2013; Ansari et al., 2016; Lende and Raghuwanshi, 2016). In other words, QA systems make it possible asking questions and retrieve the answers using natural language queries (Abdi et al., 2016) and may be considered as an advanced form of Information Retrieval (IR) (Cao et al., 2010).

With the efforts from academic research, Question Answering is a growing research field worldwide (Voorhees and Tice, 2000; Wang et al., 2000). The demand for this kind of system increases day by day since it delivers short, precise and question-specific answers (Pudaruth et al., 2016). Nevertheless, a systematic approach for understanding the algorithms, techniques and systems around Question Answering is lacking so far. Although previous literature reviews have focused on specific aspects of Question Answering like domain (Athenikos and Han, 2010; Kolomiyets and Moens, 2011), information retrieval paradigm (Gupta and Gupta, 2012) and hybrid based paradigm (Kalyanpur et al., 2012), the relationships between domains, algorithms, techniques and systems have not been established.

We provide an a holistic view of QA according to the literature. We performed a systematic mapping study to enlighten the paradigms, technologies, domains, metrics and concepts that surround this field of research. As result, we identified how the research community addresses the theme, the main paradigms observed, how the approaches fit in different kinds of domains, the results obtained by the implementations of these approaches.

We outline this paper as follows: we explained the main concepts about the theme in Section 2, then we detailed the systematic literature review process in Section 3 and presented the results of the SLR in Section 4. We also analyzed the related works found on literature in Section 5.

2. Background

2.1. Question Answering

Question Answering systems in information retrieval are tasks that automatically answer the questions asked by humans in natural language using either a pre-structured database or a collection of natural language documents (Chali et al., 2011; Dwivedi and Singh, 2013; Ansari et al., 2016; Lende and Raghuwanshi, 2016). In other words, QA systems enable asking questions and retrieving answers using natural language queries (Abdi et al., 2016). (Cao et al., 2010) consider QA systems an advanced form of information retrieval. The demand for this kind of system increases on a daily basis since it delivers short, precise and question-specific answers. Pudaruth et al. (2016). With the efforts from academic research, the QA subject has attracts growing interest around the world (Voorhees and Tice, 2000; Wang et al., 2000) and the main evidence of this is the IBM Watson (Ferrucci, 2012).

To understand the Question Answering subject, we firstly define the associated terms. A *Question Phrase* is the part of the question that says what is searched. The term *Question Type* refers to a categorization of the question for its purpose. In the literature the term *Answer Type* refers to a class of objects which are sought by the question. *Question Focus* is the property or entity being searched by the question. *Question Topic* is the object or event that the question is about. *Candidate Passage* can broadly be defined as anything from a sentence to a document retrieved by a search engine in response to a question. *Candidate Answer* is the text ranked according to its suitability to as an answer (Prager, 2007).

Previous studies mostly defined a architecture of Question Answering systems in three macro modules (Malik et al., 2013; Bhoir and Potey, 2014; Neves and Leser, 2015): *Question Processing*, *Document Processing* and *Answer Processing* as showed in Fig. 1.

Question Processing receives the input from the user, a question in natural language, to analyze and classify it. The analysis is to find out the type of question, meaning the focus of the question. This is necessary to avoid ambiguities in the answer (Malik et al., 2013). The classification is one of the main steps of a QA system. There are two main approaches for question classification, manual and automatic (Ray et al., 2010). Manual classification applies handmade rules to identify expected answer types. These rules may be accurate but they are time-consuming, tedious, and non-extensible in nature. There are approaches that classify the question type as *What*, *Why*, *Who*, *How*, *Where* questions and so on Moldovan et al. (2003) and Gupta and Gupta (2012). This type of definition helps on a better answer detection. Automatic classifications, in contrast, are extensible to new questions types with acceptable accuracy (Liang et al., 2007; Ray et al., 2010).

Question processing is divided in two main procedures. The first one is to analyze the structure of the user's question. The second one it to transform the question into a meaningful question formula compatible with QA's domain (Hamed and Ab Aziz, 2016). Questions can also be defined by the type of answer expected. The types are *factoid*, *list*, *definition* and *complex question* (Kolomiyets and Moens, 2011). Factoid questions are the ones that ask about a simple fact and can be answered in a few words (Heie et al., 2012), for instance, *How far is it from Earth to Mars?*. List Question demands as an answer a set of entities that satisfies a given criteria (Heie et al., 2012), *When did Brazil win Soccer World Cups?* illustrates this point clearly. Definition questions expect a summary or a short passage in return (Neves and Leser, 2015): *How does the mitosis of a cell work?* is a good illustration of it. In contrast, Complex Question is about information in a context. Usually, the answer is a merge of retrieved passages. This merge is implemented using algorithms, such as: *Normalized Raw-Scoring*, *Logistic Regression*, *Round-Robin*, *Raw Scoring* and *2-step RSV* (Garĉa-Cumbreras et al., 2012).

Different from question processing that is execute on every question asked by the user, *Document Processing* has as its main feature the selection of a set of relevant documents and the extraction of a set of paragraphs depending on the focus of the question or text understanding throw natural language processing (Malik et al., 2013). This task can generate a dataset or a neural model which will provide the source for the answer extraction. The retrieved data can be ranked according to its relevance for the question (Neves and Leser, 2015).

The *Answer Processing* is the most challenging task on a Question Answering system. This module uses extraction techniques on the result of the *Document Processing* module to present an answer (Bhoir and Potey, 2014). The answer must be a simple answer for the question, but it might require merging information

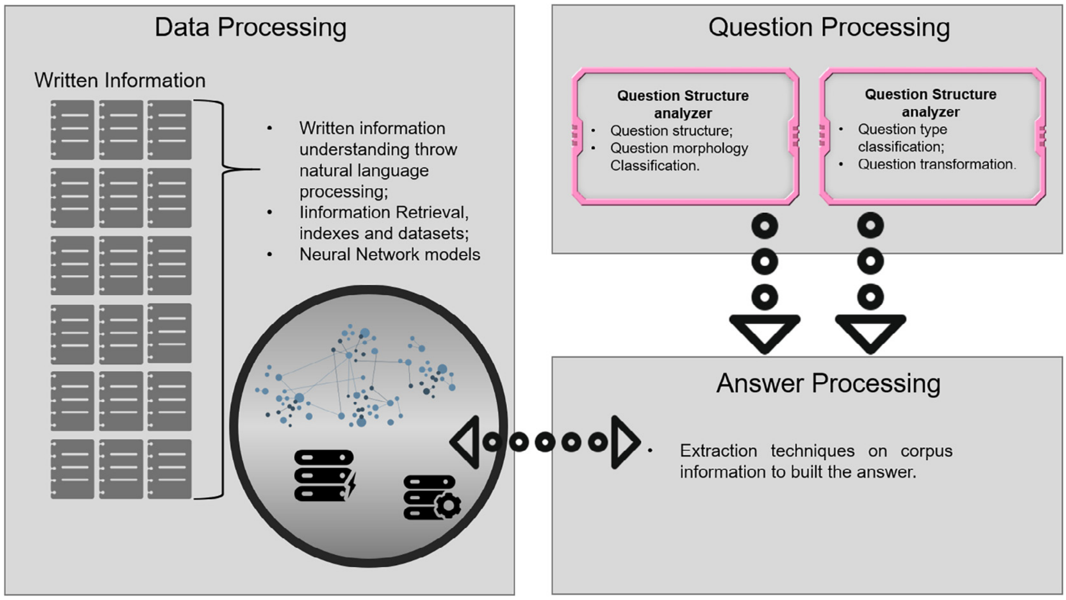

Fig. 1. Architecture of the three macro question answering modules. Question processing module classifies the question by its type and morphology. Answer processing module uses the classification and transformation made in question processing module to extract the answer from the result of the Document Processing module that executes previously to create datasets, indexes or neural models.

from different sources, summarization, dealing with uncertainty or contradiction.

Besides the main architecture, QA systems can be defined by the paradigm each one implements (Yao, 2014):

1. *Information Retrieval QA:* Usage of search engines to retrieve answers and then apply filters and ranking on the recovered passage.
2. *Natural Language Processing QA:* Usage of linguistic intuitions and machine learning methods to extract answers from retrieved snippet.
3. *Knowledge Base QA:* Find answers from structured data source (a knowledge base) instead of unstructured text. Standard database queries are used in replacement of word-based searches (Yang et al., 2015). This paradigm, make use of structured data, such as ontology. An ontology describes a conceptual representation of concepts and their relationships within a specific domain. Ontology can be considered as a knowledge base which has a more sophisticated form than a relational database (Abdi et al., 2016). To execute queries in order to retrieve knowledge from the ontology, structured languages are proposed and one of then is SPARQL.[1]
4. *Hybrid QA:* High performance QA systems make use of as many types of resources as possible, especially with the prevailing popularity of modern search engines and enriching community-contributed knowledge on the web. A hybrid approach is the combination of IR QA, NLP QA and KB QA. The main example of this paradigm is IBM Watson (Ferrucci et al., 2013)

2.1.1. Evaluation methods

Evaluation methods are part of a Question Answer system. As QA approaches are developed rapidly, reliable evaluation metrics to compare these implementations are needed. According to Yao (2014), the evaluation metrics used in QA are F_1 and accuracy. To understand these measures we have to keep in mind a 2×2 contingency table. For any particular piece of data being evaluated, this table will classify it in two classes, a fragment that was correctly selected (true positive) or was correctly not selected (false negative) and a fragment that was not correctly selected (false positive) or incorrectly not selected (true negative). Using accuracy as a measure metric means the application of the formula 1.

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + False\ Negative + True\ Negative}$$

The issue observed on this metric for Question Answering systems evaluation is the high rates of true negative a system can find, i. e., when a fact question is made, there is one correct answer, anything else would be incorrect and not selected. In this case, a system could have a high calculated accuracy but unmeaningful. To fix this issue, the *f measure* is addressed and it is based on the same 2×2 contingency table and two measures: *Precision* and *Recall*. Precision (Formula 2) is the percentage of selected answers that are correct and Recall (Formula 3) is the opposite measure, it is the percentage of correct answers selected. Using Precision and Recall, the fact of a high rate of true negative answers is not relevant anymore (Kumar et al., 2005; Yao, 2014).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{2}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{3}$$

Analyzing Precision and Recall and knowing that they are opposite, the key concept here is the trade-off researchers must do in each measure, looking for the best metrics to evaluate their systems. Most of fact QA systems should use Recall as a measure metric since it does not matter how high the false positive rates are, if there are high true positive rates, the result will be good. However, for list or definition QA systems, maybe Precision would be better. To balance this trade-off, the *f measure* is presented (Zhang et al., 2008; Yao, 2014) (Formula 4).

$$F_1 = \frac{2(Precision.Recall)}{Precision + Recall} \tag{4}$$

[1] https://www.w3.org/TR/rdf-sparql-query/.

This measure implements a weighted way of assessing the Precision and Recall trade-off.

There are metrics that can be used to evaluate QA systems, such as Mean Average Precision (MAP) presented on Formula 5, which is a standard measure for Information Retrieval, Mean Reciprocal Rank (MRR) showed in Formula 7 used to calculate the answer relevance (Li et al., 2007; Zhang et al., 2008; Yao, 2014). This one is mainly used in Information Retrieval paradigms.

$$MAP = \frac{\sum_{Q}^{q=1} AveP(q)}{Q} \tag{5}$$

The *AveP*, Average Precision, is given by the Eq. 6:

$$AveP = \frac{\sum_{N}^{k=1} P(k) \times rel(q)}{|\{relevant\}|} \tag{6}$$

$$MRR = \frac{1}{N}\sum_{i=1}^{N} RR_{(q_i)} \tag{7}$$

3. Research

This Systematic Literature Review (SLR) was based on guidelines provides by Okoli and Schabram (2010) and Keele (2007). The review tasks are based on their eight steps, and here we will describe: *Purpose of the Literature Review*, *Searching the Literature*, *Practical Screen*, *Quality Appraisal* and *Data Extraction*.

A exponential growth in written digital information led us to the need for increasingly sophisticated search tools (Bhoir and Potey, 2014; Pinto et al., 2014). The amount of unstructured data is increasing and it has been collected and stored at unprecedented rates (Chali et al., 2011; Bakshi, 2012; Malik et al., 2013). The challenge is to create ways to consume this data, extract information and knowledge having an interesting experience in the process. In this context the Question Answering systems emerge, providing a natural language interaction between humans and computers to answer as many questions as possible and enabling the retrieval of these answers from unstructured data sets. We created five research questions, showed in Table 1, to guide this SLR as an attempt to understand how Question Answering systems techniques, tools, algorithms and systems work and perform, how reliable implementing tasks they are and the relationship between QA and Natural Language Processing (NLP).

Table 1
Research Questions.

ID	Question
RQ1	What is the representativeness of each QA paradigm?
RQ2	Which are the QA techniques addressed?
RQ3	Which metrics or indicators are used to compare the different QA algorithms, techniques and systems?
RQ4	Which are the fields in which QA systems and NLP are used?
RQ5	How the relationship between QA systems and NLP is built?

Table 2
Number of papers retrieved in each Digital Library after search strings execution.

Digital Library	Number Of Returned Papers
ACM Digital Library	138
IEEE Xplore	182
Science Direct – Elsevier	863
Springer Link	638
Wiley	21
TOTAL	1842

3.1. Data retrieval

We indexed journals and papers written in English. Besides the language factor, a date filter was applied. Papers published from 2000 up to December in 2017 were accepted. We defined this date criteria based on the observations of Wang et al. (2000). After the first TREC[2] that addressed QA systems in 1999, the number of studies about the theme increased and also represent its evolution. Book chapters were also excluded from the search. The searches were conducted in five digital libraries, ACM Digital Library,[3] IEEE Xplore,[4] Science Direct – Elsevier,[5] Springer Link[6] and Wiley.[7] To execute this task, a conceptual research string was developed containing the main keyword of the theme. We executed the search strings on December 1st 2016 in each digital library and the results are presented in Table 2.

3.2. Screening of papers

We applied inclusion and exclusion criteria to be explicit about the studies we considered in our review. We kept in this study papers that satisfy the inclusion and exclusion criteria, described in Table 3. With the practical screening we selected 221 papers from the initial 1842 set.

3.3. Eligibility

To achieve the commitment to include relevant papers in this review synthesis, we created quality assessment questions (QAQ) shown in Table 4 as proposed by (Kitchenham et al., 2009).

Each QAQ was answered with Yes (Y), Partial (P) or No (N). We assigned values to these answers in which Yes = 1.5, Partial = 0.5 and No = 0. From this moment on, 203 papers were read and evaluated according to the Quality Assessment Questions. The articles with scores lower than 0.5 were excluded from the result set. Finally, we selected 130 papers, that are listed in Table 5.

3.4. Data extraction and classification

With the studies included in the review identified, we started a data extraction from the papers full readings. We created 6 categories to classify the studies set. Table 6 describes these categories.

4. Results and discussion

This section describes and discusses the findings from the data extraction and classification activities. The findings are presented in a graphical view and are organized by research question.

Using the categories described in Section 3.4, we analyzed the information throw its time line. When we consider this time line of the publications, it is possible to see relevant works published each year in the Question Answering field. When crossing this information with the obtained results these works got by year, showed in Fig. 2, we can highlight a variation on its average, which revels that QA paradigm is an open question.

4.1. RQ1 – What is the representativeness of QA paradigm?

To answer this question, we considered the observations brought up by Yao (2014) to classify the papers. From the 130

[2] http://trec.nist.gov/pubs/trec8/t8_proceedings.html.
[3] http://dl.acm.org.
[4] http://ieeexplore.ieee.org/Xplore/home.jsp.
[5] http://www.sciencedirect.com.
[6] http://link.springer.com.
[7] http://onlinelibrary.wiley.com/.

Table 3
Inclusion and Exclusion Criteria defined for screening.

Inclusion Criteria	Exclusion Criteria
Papers written in English	Papers written in other languages rather than English
Academic papers published in conferences or journals	Duplicated papers found on the digital libraries
Papers that describe a Question Answering algorithm, technique or system	Books, thesis, editorials, prefaces, article summaries, interviews, news, reviews, correspondences, discussions, comments, reader's letters and summaries of tutorials, workshops, panels, and poster sessions
Papers that establish a relationship between Question Answering and Natural Language Processing	
Papers published until December 1st 2017	
Papers published from 2000 on wards	

Table 4
Question Assessment Questions.

ID	Research Question
QAQ 1	Is the main objective of the paper to deal with the question answering topic?
QAQ 2	Does the paper describe a question answering algorithm, technique or system?
QAQ 3	Is the paper able to establish the relationship between question answering and natural language processing?
QAQ 4	Does the paper describe the results of the proposed QA algorithm, technique or system with an evaluation metric?

studies, 34.59% of the papers implemented a knowledge base paradigm, 33.08% implemented a natural language processing paradigm, 28.57% implemented an information retrieval paradigm and 3.76% implemented a Hybrid paradigm. We observed a slightly higher usage of knowledge base and natural language processing implementations.

As one may notice, knowledge base and natural language processing paradigms are implemented at almost the same rate (Fig. 2), so we went further in this relation and crossed the amount of implementations of three paradigms (we excluded Hybrid

Table 5
124 Papers addressed in this study. They were categorized by the question answering paradigm implemented.

Addressed Paradigm	Papers
NLP	(Pack and Weinstein, 2001; Li et al., 2002; Girju, 2003; Solorio et al., 2004; Jung et al., 2005; Metzler and Croft, 2005; Pustejovsky et al., 2005; Xie and Liu, 2005; Saquete et al., 2006; Beg et al., 2007; Liu et al., 2007b; Lin et al., 2007a; Liang et al., 2007; Ko et al., 2007; Hartawan and Suhartono, 2015; Vazquez-Reyes and Black, 2008; Mansouri et al., 2008; Wang and Manning, 2010; Cao et al., 2010; Surdeanu et al., 2011; Silva et al., 2011; Oh et al., 2011; Moreda et al., 2011; Chali et al., 2011; Pinto et al., 2014; Bhoir and Potey, 2014; Bonnefoy et al., 2011; Fikri and Purwarianti, 2012; Perera and Perera, 2012; Aarabi, 2013; Ferrucci et al., 2013; Sagara and Hagiwara, 2014; Biswas et al., 2014; Ilvovsky, 2014; Chali et al., 2015; Hristovski et al., 2015; Ansari et al., 2016; Lende and Raghuwanshi, 2016; Nanda et al., 2016; Pechsiri and Piriyakul, 2016; Archana et al., 2016; Yue et al., 2017; Chandurkar and Bansal, 2017; Atkinson and Maurelia, 2017)
IR	(Moldovan et al., 2000; Hammo et al., 2002; Moldovan et al., 2003; Dumais, 2003; Lin and Katz, 2003; Azari et al., 2004; Hammo et al., 2004; Kumar et al., 2005; Li et al., 2005; Radev et al., 2005; Xie and Liu, 2005; Yu et al., 2007; Moldovan et al., 2007; Li et al., 2007; Zhang et al., 2008; Wen et al., 2008; Tapeh and Rahgozar, 2008; Kosseim and Yousefi, 2008; Figueroa and Neumann, 2008; Wenyin et al., 2009; Athenikos et al., 2009; Moriceau and Tannier, 2010; Balahur et al., 2010; Zong et al., 2011; Richardson et al., 2011; Cao et al., 2011; Vila et al., 2011; Heie et al., 2012; Garća-Cumbreras et al., 2012; Barskar et al., 2012; Komiya et al., 2013; Kamal et al., 2014; xxxx, 2014; Pavlić et al., 2015; Bakari et al., 2016; Seena et al., 2016; Galitsky, 2017)
KB	(Harabagiu et al., 2000; Hsu et al., 2001; Cheng et al., 2002; Li et al., 2002; Chu-Carroll et al., 2003; Li et al., 2003; Beale et al., 2004; Montero and Araki, 2004; Badia, 2007; Terol et al., 2007; Lopez et al., 2007; Rabiah et al., 2007; Grau, 2006; Rabiah et al., 2008; Guo and Zhang, 2009; Guo and Zhang, 2009; Guo, 2009; Ferrández et al., 2009; Ferrández et al., 2009; Dietze and Schroeder, 2009; Varathan et al., 2010; Peng et al., 2010; Furbach et al., 2010; Buscaldi et al., 2010; Malik et al., 2013; Tartir et al., 2011; Al-Nazer and Helmy, 2015; Ferrandez et al., 2011; Moré et al., 2012; Kuchmann-Beauger et al., 2013; Spranger and Labudde, 2014; Toti, 2014; Shekarpour et al., 2015; Abacha and Zweigenbaum, 2015; Zayaraz, 2015; Yang et al., 2015; Molino et al., 2015; Pudaruth et al., 2016; Hakkoum and Raghay, 2016; Abdi et al., 2016; Cui et al., 2017; Yin et al., 2017)
Hybrid	(Delmonte, 2006; Frank et al., 2007; Sucunuta and Riofrio, 2010; Chandrasekar, 2014; Walke and Karale, 2013)

Table 6
Six Categories created to classify the Studies.

Category	Description
01	Metadata of the papers, comprising the name of the paper and DOI, the home country of the author, University, the year of publication and the digital library where the paper came from
02	Defined which main approach was selected on the paper. Four paradigms were observed, Natural Language Processing based, Information Retrieval based, Knowledge Base based and Hybrid Based (Yao, 2014)
03	Was based on natural language processing technique used on the approach (Wang et al., 2000; Beg et al., 2007; Cao et al., 2011; Monner and Reggia, 2012; Bhoir and Potey, 2014; Chandrasekar, 2014; Pinto et al., 2014; Chali et al., 2015; Ansari et al., 2016; Hakkoum and Raghay, 2016)
04	Dealt with the metrics used to evaluate the accuracy of the question answering systems (Solorio et al., 2004; Bonnefoy et al., 2011; Pinto et al., 2014; Hakkoum and Raghay, 2016)
05	Determined whether the approach used a training set or which kind of corpora or Data Source was addressed. Examples are, Wordnet[8], Hownet,[9] GoiTaikei[10] and MedLine[11]
06	Detailed classifications such as: language of application, fields of application, usage of a chat bot or dialogue approach to improve the quality of the information from users, information merging algorithms used to answer complex questions and question type (Hsu et al., 2001; Garća-Cumbreras et al., 2012; Biswas et al., 2014; Neves and Leser, 2015; Mishra and Jain, 2016)

[8] https://wordnet.princeton.edu/.
[9] http://www.keenage.com/.
[10] http://www.kecl.ntt.co.jp/icl/lirg/resources/GoiTaikei/index-en.html.
[11] https://www.medline.com/.

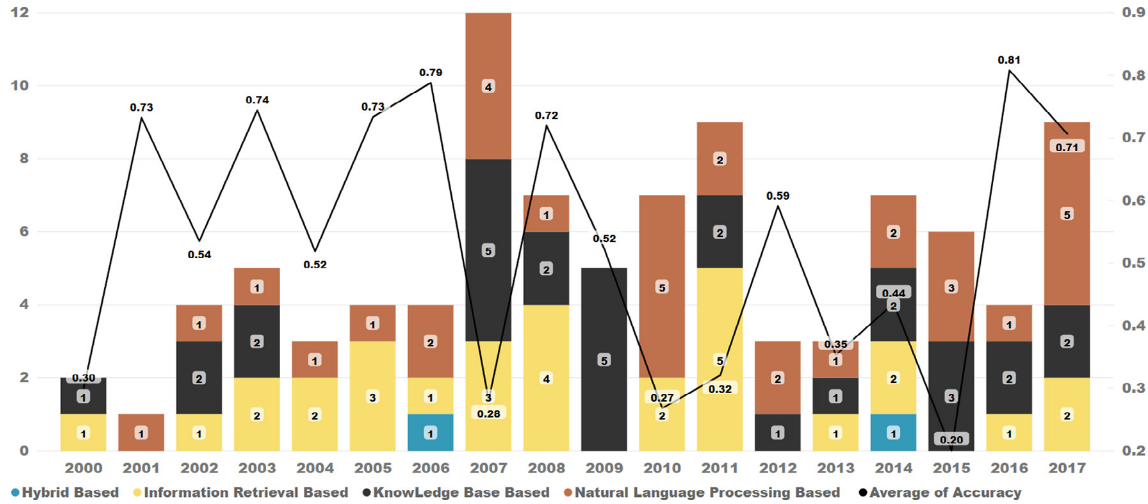

Fig. 2. Publication Time-Line since 2000 and the average of the results obtained by the approaches implementation. The average oscillation over the years show a constantly seek to better option on QA research.

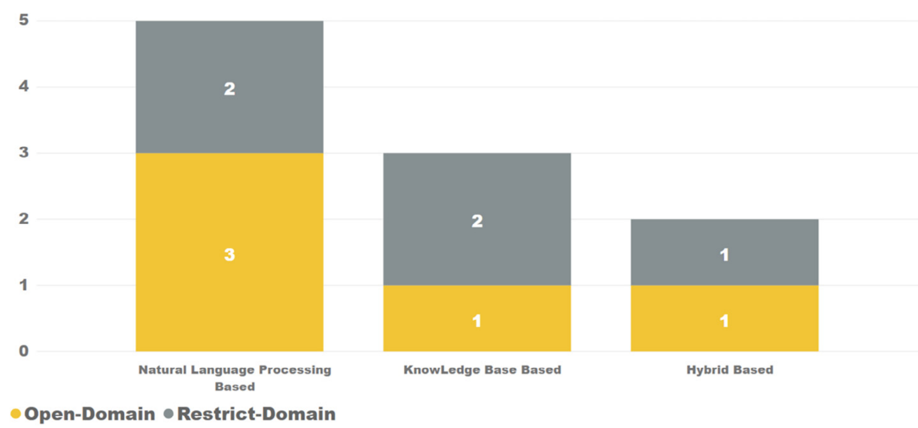

Fig. 3. Number of Papers that implemented a Independence Language approach by each related paradigm. The papers are grouped by the Domain they implemented.

Paradigms due to its lower rates of implementation) against the following dimensions: the first dimension is the language implemented on the approach. The approaches are implemented to understand a natural language question, and answer in the same way. With this analysis, we were able to see that English was the most adopted language and natural language processing paradigms implemented independent language approaches. The second dimension that we had crossed with the representativeness of each QA paradigm was the accuracy each one reaches. We could observe an average accuracy per year higher for natural language processing paradigm in 53% of the cases. One important finding for this study is the accuracy reached by approaches that implemented language Independence. Analyzing these approaches, we were able to identify that, NLP paradigms were more used by researchers to develop QA systems with any language as it focus (Fig. 3), but, as important as the rate of each paradigm was used, is the average accuracy that the approaches reached, and from this point of view, KB had a better performance than NLP if we consider Restrict-Domain QA system, as we can see in Fig. 4.

4.2. RQ2 – Which are the most frequently applied QA techniques?

To implement the modules of a QA software architecture (*Question Processing*, *Document Processing* and *Answer Processing*), researchers have used techniques, algorithms, frameworks and systems related to information extraction, natural language processing and machine learning. We categorized these tools in order to analyze and summarize their relevance to the study. We clustered, in Fig. 5, the 15 most addressed techniques, algorithms, frameworks and tools observed in the SLR by the paradigms they are part of.

4.3. RQ3 – Which metrics or indicators are used to compare different Question Answering algorithms, techniques or systems?

In our research, we extracted the metrics that researchers used to evaluate their implementations in each works. In Fig. 6, we clustered the identified metrics by the QA paradigms.

We are able to understand from this analysis that since natural language processing paradigm had the best average evaluation performance over the years, the *Precision* and *Recall* metrics are indicated to evaluate this kind of implementation.

The elapsed time performed by the approach on answer extraction and user delivery was not evaluated by most of researchers, only 6% of papers addressed this metric.

It is important to observe how the evaluation was performed in the approaches when analyzing the obtained results. In order to provide a comparison among implementations, a common dataset for training and testing is crucial. We crossed metrics extracted by the usage of a testing or training set provided by QA conference, such

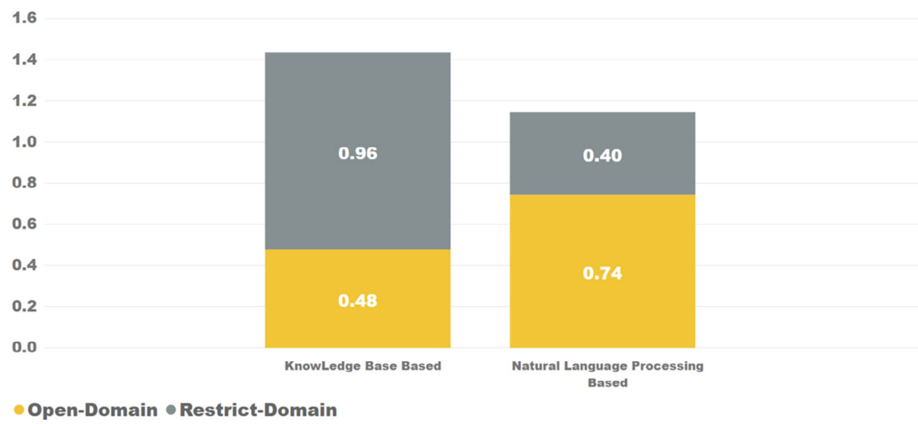

Fig. 4. Relationship between the high average accuracy reached and independence language implementations, showing that Natural Language Processing paradigms are a better fit if Open-Domain QA systems are needed and Knowledge Based paradigms are better if Restrict-Domain QA systems are the focus.

	Hybrid Based	Information Retrieval Based	KnowLedge Base Based	Natural Language Processing Based
Deep Neural Network			1	2
Graph Based				3
Lemmatization			3	
LSA		1		1
Multi Document Summarization		1		2
Naive Bayes			1	2
Named Entity Recognition	2	14	8	10
Parser		4	5	4
POS Tagging	2	15	11	14
Relation Finding (Similarity Distance)	1	3	4	5
Shallow Syntactical			1	1
Stemming		1	1	2
Support Vector Machine		4		9
Text Chunking		5	2	1
Tokenization		4	3	6

Papers That Addres..
1 15

Fig. 5. TOP 15 Techniques, algorithms, frameworks and tools observed in SLR.

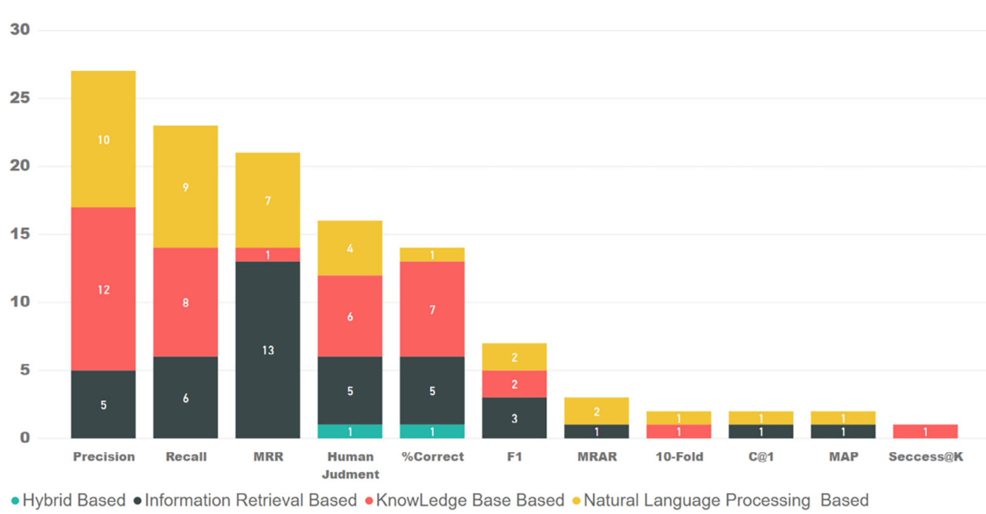

Fig. 6. QA paradigms and their metrics.

as *TREC* (Pack and Weinstein, 2001; Moldovan et al., 2003; Chu-Carroll et al., 2003; Girju, 2003; Lin and Katz, 2003; Li et al., 2005; Lin and Demner-Fushman, 2006; Kunichika et al., 2007; Kosseim and Yousefi, 2008; Hartawan and Suhartono, 2015) and *CLEF* (Figueroa and Neumann, 2008; Buscaldi et al., 2010), and we found out that 24% of the approaches used a training or test set like that.

4.4. RQ4 – Which are the fields in which Question Answering systems and Natural Language Processing are being applied?

Question Answering has been implemented in different fields of knowledge. The domains in which QA systems are implemented can be divided in *Open-Domain*, *Restrict-Domain* and *Closed-Domain*. Our research summarized where and in which context researchers implemented their systems. We did not find any applications in *Closed-Domain*. We found out with this analysis that *Open-Domain* based on World Wide Web implementations are the biggest part of the researches and medicine subject is also treated at an important rate, Fig. 7 reveals that. From this analysis, we see that question answering intersects with many areas and domains, showing how this systems can be important on knowledge extraction for any kind of user and need.

4.5. RQ5 – How the relationship between Question Answering systems and Natural Language Processing is built?

When we analyze the QA implementations starting from 2000, it is possible to see the paradigms in each approach. The paradigms are named over their main task, but they use natural language processing techniques whether to classify the question or to retrieve the answer. In other words, NLP paradigm is not the only relationship between QA and natural language processing. Information Retrieval and Knowledge Base make use of NLP techniques that can help on the implementation of these paradigms. Techniques such as POS Tagging, Tokenization, Named Entity Recognition, Semantic Parser and Similarity Distance are described in papers which addressed IR or KB paradigms as we can see on Fig. 8. This analysis shows us how natural language processing is important

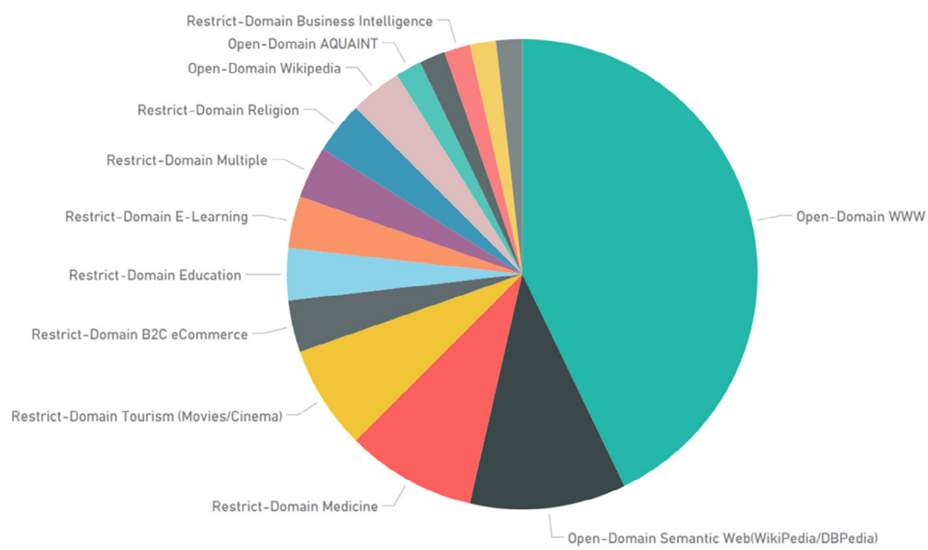

Fig. 7. Fields where researcher are implementing their approaches. There is a wide set of fields, besides the amount of implementations in open domain on *www* and medicine on restricted domain.

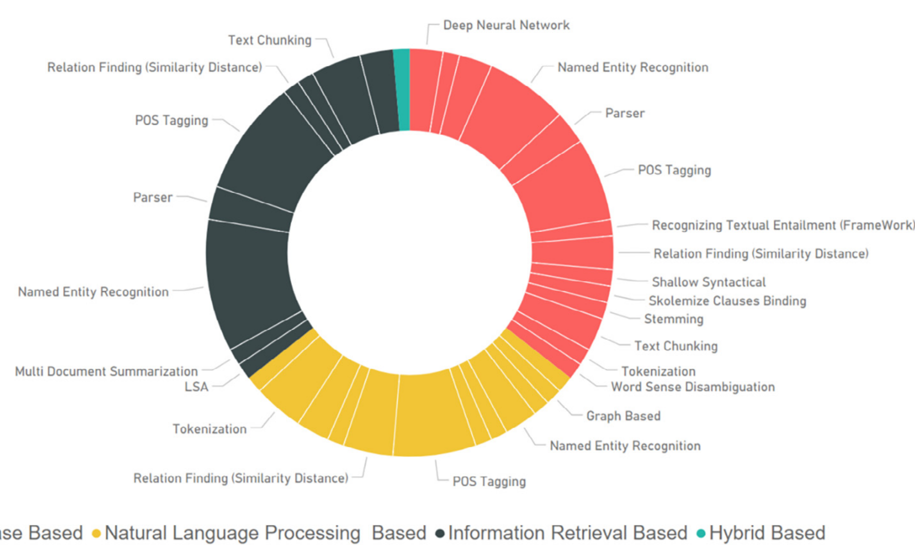

Fig. 8. Natural language processing techniques used in other paradigms besides de NPL based one.

for QA systems, they are essential to create the understanding between the user and the machine and most part of the researcher are using NLP on their approaches.

5. Related work

In the past decades, automated QA has generated interesting of information extraction researchers (Maybury, 2008). Question Answering systems have become challenging due to the complexity and applicability of these systems. Studies published over the past fifteen years addressed points of view, surveys and reviews about the theme (Hamed and Ab Aziz, 2016).

From the works we found, we could see how question answering and natural language processing interacts and intersects (Hirschman and Gaizauskas, 2001) and a comparison on approaches based on natural language processing, information retrieval and question templates, analyzing the differences among the QA approaches, their accuracy and applicability (Andrenucci and Sneiders, 2005; Prager, 2007).

We found reviews for each kind of paradigm. A knowledge-based approach in order to implement a QA systems in the Bio-medical domain (Athenikos and Han, 2010; Kolomiyets and Moens, 2011), a information retrieval paradigm complete analysis (Gupta and Gupta, 2012) and a hybrid based paradigm review (Kalyanpur et al., 2012). One of the limitations with these studies is that they do not look after all paradigms together, comparing then.

There were works that discuss the techniques, defines the core components and proposes trends for the future of QA field (Bouziane et al., 2015; Hamed and Ab Aziz, 2016; Mishra and Jain, 2016), describes the QA core components and provides a comparison among the Question Answering systems implemented in 16 papers (Mishra and Jain, 2016).

These works helped on the development of this systematic literature review. This study aims to fill a gap on systematic literature reviews on question answering, analyzing the paradigms and their behavior, making available a evaluation of them to establish when which one can be better used. We considered in this study a classification for the paradigms, their implemented techniques, their metrics and fields of usage, which no other work considered.

6. Conclusion

We summarized on this study the current state of the Question Answering literature. This SLR identified 1842 studies on the search results for the defined search strings. After the screening phases, a total of 130 studies were selected as studies about question answering approaches and their applications. The selected studies were analyzed and the extracted data was classified into categories focusing on research aspects (research type and empirical research), as well as technical features, such as: the paradigm type, domain of application, natural language processing techniques applied, accuracy, and language of application.

We observed that researchers have equally used the paradigms in their approaches, and the Hybrid approach has not been used at the same rates because of its complex nature. We were able to see that natural language approach is the most flexible to independent language implementations which is noteworthy, since we did not find any implementations in Portuguese (Brazil) on the literature.

Besides the paradigms and their approaches, we were able to enlighten the metrics used to evaluate QA systems, the domains in which researchers apply their systems and how NLP is used in this task, apart from the paradigm adopted.

The contributions of this work are a systematic outline of question answering systems research summarizing the research and technical features found on literature and definition to be followed by researchers that want to understand and implement QA systems.

Future research should explore the question answering when applied to the extraction of complex answers, answers that should be given in a context. Algorithms such as *Round-Robin*, *Raw Scoring*, *2-step RSV*, and *Normalized Raw-Scoring* require further research in order to reach this kind of knowledge.

References

Aarabi, P., 2013. Virtual cardiologist—a conversational system for medical diagnosis. 2013 26th Annual IEEE Canadian Conference on Electrical and Computer Engineering (CCECE). IEEE, pp. 1–4.

Abacha, A.B., Zweigenbaum, P., 2015. Means: a medical question-answering system combining nlp techniques and semantic web technologies. Inf. Process. Manage. 51 (5), 570–594. https://doi.org/10.1016/j.ipm.2015.04.006.

Abdi, A., Idris, N., Ahmad, Z., 2016. Qapd: an ontology-based question answering system in the physics domain. Soft. Comput., 1–18 https://doi.org/10.1007/s00500-016-2328-2.

Al-Nazer, A., Helmy, T., 2015. Personalizing health and food advices by semantic enrichment of multilingual cross-domain questions. GCC Conference and Exhibition (GCCCE), 2015 IEEE 8th. IEEE, pp. 1–6. https://doi.org/10.1109/ieeegcc.2015.7060095.

Andrenucci, A., Sneiders, E., 2005. Automated question answering: Review of the main approaches. Third International Conference on Information Technology and Applications, ICITA 2005, vol. 1. IEEE, pp. 514–519.

Ansari, A., Maknojia, M., Shaikh, A., 2016. Intelligent question answering system based on artificial neural network. 2016 IEEE International Conference on Engineering and Technology (ICETECH). IEEE, pp. 758–763. https://doi.org/10.1109/icetech.2016.7569350.

Archana, S., Vahab, N., Thankappan, R., Raseek, C., 2016. A rule based question answering system in malayalam corpus using vibhakthi and pos tag analysis. Procedia Technol. 24, 1534–1541. https://doi.org/10.1016/j.protcy.2016.05.124.

Athenikos, S.J., Han, H., 2010. Biomedical question answering: a survey. Computer Methods Programs Biomed. 99 (1), 1–24. https://doi.org/10.1016/j.cmpb.2009.10.003.

Athenikos, S.J., Han, H., Brooks, A.D., 2009. A framework of a logic-based question-answering system for the medical domain (loqas-med). Proceedings of the 2009 ACM symposium on Applied Computing. ACM, pp. 847–851. https://doi.org/10.1145/1529282.1529462.

Atkinson, J., Maurelia, A., 2017. Redundancy-based trust in question-answering systems. Computer 50 (1), 58–65. https://doi.org/10.1109/mc.2017.18.

Azari, D., Horvitz, E., Dumais, S., Brill, E., 2004. Actions, answers, and uncertainty: a decision-making perspective on web-based question answering. Inf. Process. Manage. 40 (5), 849–868. https://doi.org/10.1016/j.ipm.2004.04.013.

Badia, A., 2007. Question answering and database querying: Bridging the gap with generalized quantification. J. Appl. Logic 5 (1), 3–19. https://doi.org/10.1016/j.jal.2005.12.007.

Bakari, W., Bellot, P., Neji, M., 2016. Aqa-webcorp: web-based factual questions for arabic. Procedia Computer Sci. 96, 275–284. https://doi.org/10.1016/j.procs.2016.08.140.

Bakshi, K., 2012. Considerations for big data: architecture and approach. Aerospace Conference, 2012 IEEE. IEEE, pp. 1–7. https://doi.org/10.1109/aero.2012.6187357.

Balahur, A., Boldrini, E., Montoyo, A., Martínez-Barco, P., 2010. Going beyond traditional qa systems: challenges and keys in opinion question answering. Proceedings of the 23rd international conference on computational linguistics: Posters. Association for Computational Linguistics, pp. 27–35.

Barskar, R., Ahmed, G.F., Barskar, N., 2012. An approach for extracting exact answers to question answering (qa) system for english sentences. Procedia Eng. 30, 1187–1194. https://doi.org/10.1016/j.proeng.2012.01.979.

Beale, S., Lavoie, B., McShane, M., Nirenburg, S., Korelsky, T., 2004. Question answering using ontological semantics. Proceedings of the 2nd Workshop on Text Meaning and Interpretation. Association for Computational Linguistics, pp. 41–48. https://doi.org/10.3115/1628275.1628281.

Beg, M.S., Thint, M., Qin, Z., 2007. Pnl-enhanced restricted domain question answering system. Fuzzy Systems Conference, FUZZ-IEEE 2007. IEEE International. IEEE, pp. 1–7. https://doi.org/10.1109/fuzzy.2007.4295551.

Bhoir, V., Potey, M., 2014. Question answering system: a heuristic approach. 2014 Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT). IEEE, pp. 165–170. https://doi.org/10.1109/icadiwt.2014.6814704.

Biswas, P., Sharan, A., Malik, N., 2014. A framework for restricted domain question answering system. 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT). IEEE, pp. 613–620. https://doi.org/10.1109/icicict.2014.6781351.

Bonnefoy, L., Bellot, P., Benoit, M., 2011. The web as a source of evidence for filtering candidate answers to natural language questions. Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, vol. 01. IEEE Computer Society, pp. 63–66. https://doi.org/10.1109/wi-iat.2011.226.

Bouziane, A., Bouchiha, D., Doumi, N., Malki, M., 2015. Question answering systems: survey and trends. Procedia Computer Sci. 73, 366–375. https://doi.org/10.1016/j.procs.2015.12.005.
Buscaldi, D., Rosso, P., Gómez-Soriano, J.M., Sanchis, E., 2010. Answering questions with an n-gram based passage retrieval engine. J. Intelligent Inform. Syst. 34 (2), 113–134.
Cao, Y.-G., Cimino, J.J., Ely, J., Yu, H., 2010. Automatically extracting information needs from complex clinical questions. J. Biomed. Inf. 43 (6), 962–971. https://doi.org/10.1016/j.jbi.2010.07.007.
Cao, Y., Liu, F., Simpson, P., Antieau, L., Bennett, A., Cimino, J.J., Ely, J., Yu, H., 2011. Askhermes: an online question answering system for complex clinical questions. J. Biomed. Inform. 44 (2), 277–288. https://doi.org/10.1016/j.jbi.2011.01.004.
Chali, Y., Hasan, S.A., Joty, S.R., 2011. Improving graph-based random walks for complex question answering using syntactic, shallow semantic and extended string subsequence kernels. Inf. Process. Manage. 47 (6), 843–855. https://doi.org/10.1016/j.ipm.2010.10.002.
Chali, Y., Hasan, S.A., Mojahid, M., 2015. A reinforcement learning formulation to the complex question answering problem. Inf. Process. Manage. 51 (3), 252–272. https://doi.org/10.1016/j.ipm.2015.01.002.
Chandrasekar, R., 2014. Elementary? Question answering, ibm's watson, and the jeopardy! challenge. Resonance 19 (3), 222–241. https://doi.org/10.1007/s12045-014-0029-7.
Chandurkar, A., Bansal, A., 2017. Information retrieval from a structured knowledgebase. 2017 IEEE 11th International Conference on Semantic Computing (ICSC). IEEE, pp. 407–412. https://doi.org/10.1109/icsc.2017.95.
Cheng, J., Kumar, B., Law, K.H., 2002. A question answering system for project management applications. Adv. Eng. Inform. 16 (4), 277–289. https://doi.org/10.1016/s1474-0346(03)00014-4.
Chu-Carroll, J., Czuba, K., Prager, J., Ittycheriah, A., 2003. In question answering, two heads are better than one. Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, vol. 1. Association for Computational Linguistics, pp. 24–31. https://doi.org/10.3115/1073445.1073449.
Cui, W., Xiao, Y., Wang, H., Song, Y., Hwang, S.-W., Wang, W., 2017. Kbqa: learning question answering over qa corpora and knowledge bases. Proc. VLDB Endowment 10 (5), 565–576. https://doi.org/10.14778/3055540.3055549.
Delmonte, R., 2006. Hybrid systems for information extraction and question answering. Proceedings of the Workshop on How Can Computational Linguistics Improve Information Retrieval? Association for Computational Linguistics, pp. 9–16. https://doi.org/10.3115/1629808.1629811.
Dietze, H., Schroeder, M., 2009. Goweb: a semantic search engine for the life science web. BMC Bioinform. 10 (10), S7. https://doi.org/10.1186/1471-2105-10-s10-s7.
Dumais, S., 2003. Data-driven approaches to information access. Cognitive Sci. 27 (3), 491–524. https://doi.org/10.1207/s15516709cog27037.
Dwivedi, S.K., Singh, V., 2013. Research and reviews in question answering system. Procedia Technol. 10, 417–424. https://doi.org/10.1016/j.protcy.2013.12.378.
Ferrández, S., Toral, A., Ferrández, Ó., Ferrández, A., Muñoz, R., 2009. Exploiting wikipedia and eurowordnet to solve cross-lingual question answering. Inf. Sci. 179 (20), 3473–3488. https://doi.org/10.1016/j.ins.2009.06.031.
Ferrández, O., Izquierdo, R., Ferrández, S., Vicedo, J.L., 2009. Addressing ontology-based question answering with collections of user queries. Inf. Process. Manage. 45 (2), 175–188. https://doi.org/10.1016/j.ipm.2008.09.001.
Ferrandez, O., Spurk, C., Kouylekov, M., Dornescu, I., Ferrandez, S., Negri, M., Izquierdo, R., Tomas, D., Orasan, C., Neumann, G., et al., 2011. The qall-me framework: a specifiable-domain multilingual question answering architecture. Web Semantics: Sci., Services Agents World Wide Web 9 (2), 137–145. https://doi.org/10.2139/ssrn.3199513.
Ferrucci, D.A., 2012. Introduction to this is watson. IBM J. Res. Dev. 56 (3.4). https://doi.org/10.1147/jrd.2012.2184356. 1-1.
Ferrucci, D., Levas, A., Bagchi, S., Gondek, D., Mueller, E.T., 2013. Watson: beyond jeopardy! Artif. Intell. 199, 93–105. https://doi.org/10.1016/j.artint.2012.06.009.
Figueroa, A.G., Neumann, G., 2008. Genetic algorithms for data-driven web question answering. Evol. Comput. 16 (1), 89–125. https://doi.org/10.1162/evco.2008.16.1.89.
Fikri, A., Purwarianti, A., 2012. Case based indonesian closed domain question answering system with real world questions. 2012 7th International Conference on Telecommunication Systems, Services, and Applications (TSSA). IEEE, pp. 181–186. https://doi.org/10.1109/tssa.2012.6366047.
Frank, A., Krieger, H.-U., Xu, F., Uszkoreit, H., Crysmann, B., Jörg, B., Schäfer, U., 2007. Question answering from structured knowledge sources. J. Appl. Logic 5 (1), 20–48. https://doi.org/10.1016/j.jal.2005.12.006.
Furbach, U., Glöckner, I., Helbig, H., Pelzer, B., 2010. Logic-based question answering. KI-Künstliche Intelligenz 24 (1), 51–55. https://doi.org/10.1007/s13218-010-0010-x.
Galitsky, B., 2017. Matching parse thickets for open domain question answering. Data Knowl. Eng. 107, 24–50. https://doi.org/10.1016/j.datak.2016.11.002.
Garća-Cumbreras, M., Martínez-Santiago, F., Ureña-López, L., 2012. Architecture and evaluation of bruja, a multilingual question answering system. Inf. Retrieval 15 (5), 413–432.
Girju, R., 2003. Automatic detection of causal relations for question answering. Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering, vol. 12. Association for Computational Linguistics, pp. 76–83. https://doi.org/10.3115/1119312.1119322.
Grau, B., 2006. Finding an answer to a question. Proceedings of the 2006 international workshop on Research issues in digital libraries. ACM, p. 7. https://doi.org/10.1145/1364742.1364751.
Guo, Q.-L., 2009. A novel approach for agent ontology and its application in question answering. J. Central South Univ. Technol. 16 (5), 781–788. https://doi.org/10.1007/s11771-009-0130-3.
Guo, Q.-L., Zhang, M., 2009. Semantic information integration and question answering based on pervasive agent ontology. Expert Syst. Appl. 36 (6), 10068–10077. https://doi.org/10.1016/j.eswa.2009.01.056.
Guo, Q., Zhang, M., 2009. Question answering based on pervasive agent ontology and semantic web. Knowl.-Based Syst. 22 (6), 443–448. https://doi.org/10.1016/j.knosys.2009.06.003.
Gupta, P., Gupta, V., 2012. A survey of text question answering techniques. Int. J. Computer Appl. 53 (4). https://doi.org/10.5120/8406-2030.
Hakkoum, A., Raghay, S., 2016. Semantic q&a system on the qur'an. Arabian J. Sci. Eng. 41 (12), 5205–5214. https://doi.org/10.1007/s13369-016-2251-y.
Hamed, S.K., Ab Aziz, M.J., 2016. A question answering system on holy quran translation based on question expansion technique and neural network classification. J. Computer Sci. 12 (3), 169–177. https://doi.org/10.3844/jcssp.2016.169.177.
Hammo, B., Abu-Salem, H., Lytinen, S., 2002. Qarab: a question answering system to support the arabic language. Proceedings of the ACL-02 workshop on Computational approaches to semitic languages. Association for Computational Linguistics, pp. 1–11. https://doi.org/10.3115/1118637.1118644.
Hammo, B., Abuleil, S., Lytinen, S., Evens, M., 2004. Experimenting with a question answering system for the arabic language. Comput. Humanit. 38 (4), 397–415. https://doi.org/10.1007/s10579-004-1917-3.
Harabagiu, S.M., Paşca, M.A., Maiorano, S.J., 2000. Experiments with open-domain textual question answering. Proceedings of the 18th conference on Computational linguistics, vol. 1. Association for Computational Linguistics, pp. 292–298. https://doi.org/10.3115/990820.990863.
Hartawan, A., Suhartono, D., et al., 2015. Using vector space model in question answering system. Procedia Computer Sci. 59, 305–311. https://doi.org/10.1016/j.procs.2015.07.570.
Heie, M.H., Whittaker, E.W., Furui, S., 2012. Question answering using statistical language modelling. Computer Speech Lang. 26 (3), 193–209. https://doi.org/10.1016/j.csl.2011.11.001.
Hirschman, L., Gaizauskas, R., 2001. Natural language question answering: the view from here. Nat. Lang. Eng. 7 (04), 275–300. https://doi.org/10.1017/s1351324901002807.
Hristovski, D., Dinevski, D., Kastrin, A., Rindflesch, T.C., 2015. Biomedical question answering using semantic relations. BMC Bioinf. 16 (1), 6. https://doi.org/10.1186/s12859-014-0365-3.
Hsu, W.-L., Wu, S.-H., Chen, Y.-S., 2001. Event identification based on the information map-infomap. 2001 IEEE International Conference on Systems, Man, and Cybernetics, vol. 3. IEEE, pp. 1661–1666.
Ilvovsky, D., 2014. Using semantically connected parse trees to answer multi-sentence queries. Autom. Doc. Math. Ling 48 (1), 33–41. https://doi.org/10.3103/s0005105514010063.
Jung, H., Yi, E., Kim, D., Lee, G.G., 2005. Information extraction with automatic knowledge expansion. Inf. Process. Manage. 41 (2), 217–242. https://doi.org/10.1016/s0306-4573(03)00066-9.
Kalyanpur, A., Boguraev, B.K., Patwardhan, S., Murdock, J.W., Lally, A., Welty, C., Prager, J.M., Coppola, B., Fokoue-Nkoutche, A., Zhang, L., et al., 2012. Structured data and inference in deepqa. IBM J. Res. Dev. 56 (3.4). https://doi.org/10.1147/jrd.2012.2188737. 10-1.
Kamal, A.I., Azim, M.A., Mahmoud, M., 2014. Enhancing arabic question answering system. 2014 International Conference on Computational Intelligence and Communication Networks (CICN). IEEE, pp. 641–645. https://doi.org/10.1109/cicn.2014.143.
Keele, S., 2007. Guidelines for performing systematic literature reviews in software engineering, in: Technical report, Ver. 2.3 EBSE Technical Report. EBSE, sn.
Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J., Linkman, S., 2009. Systematic literature reviews in software engineering–a systematic literature review. Inf. Software Technol. 51 (1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009.
Ko, J., Nyberg, E., Si, L., 2007. A probabilistic graphical model for joint answer ranking in question answering. Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, pp. 343–350. https://doi.org/10.1145/1277741.1277801.
Kolomiyets, O., Moens, M.-F., 2011. A survey on question answering technology from an information retrieval perspective. Inf. Sci. 181 (24), 5412–5434. https://doi.org/10.1016/j.ins.2011.07.047.
Komiya, K., Abe, Y., Morita, H., Kotani, Y., 2013. Question answering system using q & a site corpus query expansion and answer candidate evaluation. SpringerPlus 2 (1), 396. https://doi.org/10.1186/2193-1801-2-396.
Kosseim, L., Yousefi, J., 2008. Improving the performance of question answering with semantically equivalent answer patterns. Data Knowl. Eng. 66 (1), 53–67. https://doi.org/10.1016/j.datak.2007.07.010.
Kuchmann-Beauger, N., Brauer, F., Aufaure, M.-A., 2013. Quasl: a framework for question answering and its application to business intelligence. 2013 IEEE Seventh International Conference on Research Challenges in Information Science (RCIS). IEEE, pp. 1–12. https://doi.org/10.1109/rcis.2013.6577686.
Kumar, P., Kashyap, S., Mittal, A., Gupta, S., 2005. A hindi question answering system for e-learning documents. Third International Conference on Intelligent Sensing

and Information Processing, ICISIP 2005. IEEE, pp. 80–85. https://doi.org/10.1109/icisip.2005.1619416.
Kunichika, H., Honda, M., Hirashima, T., Takeuchi, A., 2007. A method for evaluating answers by comparing semantic information in a question and answer interaction. Syst. Computers Jpn. 38 (7), 84–97. https://doi.org/10.1002/scj.20432.
Lende, S.P., Raghuwanshi, M., 2016. Question answering system on education acts using nlp techniques. World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave). IEEE, pp. 1–6. https://doi.org/10.1109/startup.2016.7583963.
Liang, Z., Lang, Z., Jia-Jun, C., 2007. Structure analysis and computation-based chinese question classification. Sixth International Conference on Advanced Language Processing and Web Information Technology, ALPIT 2007. IEEE, pp. 39–44. https://doi.org/10.1109/alpit.2007.52.
Li, S., Zhang, J., Huang, X., Bai, S., Liu, Q., 2002. Semantic computation in a chinese question-answering system. J. Computer Sci. Technol. 17 (6), 933–939. https://doi.org/10.1007/bf02960786.
Li, W., Srihari, R.K., Li, X., Srikanth, M., Zhang, X., Niu, C., 2002. Extracting exact answers to questions based on structural links. Proceedings of the 2002 conference on multilingual summarization and question answering, vol. 19. Association for Computational Linguistics, pp. 1–9. https://doi.org/10.21236/ada457785.
Li, H., Fan, X., Li, L., 2003. A new approach to design the universal chinese question parser. Proceedings. 2003 International Conference on Natural Language Processing and Knowledge Engineering. IEEE, pp. 671–676.
Li, P., Wang, X.L., Guan, Y., Zhao, Y.M., 2005. Extracting answers to natural language questions from large-scale corpus. Proceedings of 2005 IEEE International Conference on Natural Language Processing and Knowledge Engineering, IEEE NLP-KE'05. IEEE, pp. 690–694.
Li, X., Hu, D., Li, H., Hao, T., Chen, E., Liu, W., 2007. Automatic question answering from web documents. Wuhan Univ. J. Nat. Sci. 12 (5), 875–880. https://doi.org/10.1007/s11859-007-0046-4.
Lin, J., Demner-Fushman, D., 2006. Methods for automatically evaluating answers to complex questions. Inf. Retrieval 9 (5), 565–587. https://doi.org/10.1007/s10791-006-9003-7.
Lin, J., Katz, B., 2003. Question answering from the web using knowledge annotation and knowledge mining techniques. Proceedings of the twelfth international conference on Information and knowledge management. ACM, pp. 116–123. https://doi.org/10.1145/956863.956886.
Lin, H., Yang, Z., Zhao, J., 2007a. Question-answering system based on concepts and statistics. Front. Electr. Electron. Eng. China 2 (1), 23–28. https://doi.org/10.1007/s11460-007-0004-0.
Liu, C.-H., Wang, Y.-C., Zheng, F., Liu, D.-R., 2007b. Using lsa and text segmentation to improve automatic chinese dialogue text summarization. J. Zhejiang Univ. Sci. A 8 (1), 79–87. https://doi.org/10.1631/jzus.2007.a0079.
Lopez, V., Uren, V., Motta, E., Pasin, M., 2007. Aqualog: An ontology-driven question answering system for organizational semantic intranets. Web Semantics: Sci., Serv. Agents World Wide Web 5 (2), 72–105. https://doi.org/10.1016/j.websem.2007.03.003.
Malik, N., Sharan, A., Biswas, P., 2013. Domain knowledge enriched framework for restricted domain question answering system. 2013 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC). IEEE, pp. 1–7. https://doi.org/10.1109/iccic.2013.6724163.
Mansouri, A., Affendey, L.S., Mamat, A., Kadir, R.A., 2008. Semantically factoid question answering using fuzzy svm named entity recognition. International Symposium on Information Technology, ITSim 2008, vol. 2. IEEE, pp. 1–7.
Maybury, M., 2008. New directions in question answering. Advances in open domain question answering. Springer, pp. 533–558. https://doi.org/10.1007/978-1-4020-4746-618.
Metzler, D., Croft, W.B., 2005. Analysis of statistical question classification for fact-based questions. Inf. Retrieval 8 (3), 481–504. https://doi.org/10.1007/s10791-005-6995-3.
Mishra, A., Jain, S.K., 2016. A survey on question answering systems with classification. J. King Saud Univ.-Computer Inform. Sci. 28 (3), 345–361. https://doi.org/10.1016/j.jksuci.2014.10.007.
Moldovan, D., Harabagiu, S., Pasca, M., Mihalcea, R., Girju, R., Goodrum, R., Rus, V., 2000. 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, pp. 563–570. https://doi.org/10.3115/1075218.1075289.
Moldovan, D., Paşca, M., Harabagiu, S., Surdeanu, M., 2003. Performance issues and error analysis in an open-domain question answering system. ACM Trans. Inf. Syst. (TOIS) 21 (2), 133–154. https://doi.org/10.1145/763693.763694.
Moldovan, D., Clark, C., Harabagiu, S., Hodges, D., 2007. Cogex: a semantically and contextually enriched logic prover for question answering. J. Appl. Logic 5 (1), 49–69. https://doi.org/10.1016/j.jal.2005.12.005.
Molino, P., Lops, P., Semeraro, G., de Gemmis, M., Basile, P., 2015. Playing with knowledge: A virtual player for who wants to be a millionaire? that leverages question answering techniques. Artif. Intell. 222, 157–181. https://doi.org/10.1016/j.artint.2015.02.003.
Monner, D., Reggia, J.A., 2012. Neural architectures for learning to answer questions. Biol. Inspired Cognitive Arch. 2, 37–53. https://doi.org/10.1016/j.bica.2012.06.002.
Montero, C.A., Araki, K., 2004. Information-demanding question answering system. IEEE International Symposium on Communications and Information Technology, ISCIT 2004, vol. 2. IEEE, pp. 1177–1182.
Moreda, P., Llorens, H., Saquete, E., Palomar, M., 2011. Combining semantic information in question answering systems. Inf. Process. Manage. 47 (6), 870–885. https://doi.org/10.1016/j.ipm.2010.03.008.
Moré, J., Climent, S., Coll-Florit, M., 2012. An answering system for questions asked by students in an e-learning context. Int. J. Educ. Technol. Higher Educ. 9 (2), 229–239. https://doi.org/10.7238/rusc.v9i2.1161.
Moriceau, V., Tannier, X., 2010. Fidji: using syntax for validating answers in multiple documents. Inform. Retrieval 13 (5), 507–533. https://doi.org/10.1007/s10791-010-9131-y.
Nanda, G., Dua, M., Singla, K., 2016. A hindi question answering system using machine learning approach. 2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT). IEEE, pp. 311–314. https://doi.org/10.1109/icctict.2016.7514599.
Neves, M., Leser, U., 2015. Question answering for biology. Methods 74, 36–46. https://doi.org/10.1016/j.ymeth.2014.10.023.
Oh, H.-J., Sung, K.-Y., Jang, M.-G., Myaeng, S.H., 2011. Compositional question answering: a divide and conquer approach. Inf. Process. Manage. 47 (6), 808–824. https://doi.org/10.1016/j.ipm.2010.03.011.
Okoli, C., Schabram, K., 2010. A guide to conducting a systematic literature review of information systems research. Sprouts Work. Pap. Inf. Syst. 10 (26). https://doi.org/10.2139/ssrn.1954824.
Pack, D., Weinstein, C., 2001. The use of dynamic segment scoring for language-independent question answering. Proceedings of the first international conference on Human language technology research, Association for Computational Linguistics, pp. 1–5. https://doi.org/10.21236/ada460584.
Pavlić, M., Han, Z.D., Jakupović, A., 2015. Question answering with a conceptual framework for knowledge-based system development node of knowledge. Expert Syst. Appl. 42 (12), 5264–5286. https://doi.org/10.1016/j.eswa.2015.02.024.
Pechsiri, C., Piriyakul, R., 2016. Developing a why–how question answering system on community web boards with a causality graph including procedural knowledge. Inform. Process. Agric. 3 (1), 36–53. https://doi.org/10.1016/j.inpa.2016.01.002.
Peng, X., Chen, Y., Huang, Z., 2010. A chinese question answering system using web service on restricted domain. 2010 International Conference on Artificial Intelligence and Computational Intelligence (AICI), vol. 1. IEEE, pp. 350–353. https://doi.org/10.1109/aici.2010.80.
Perera, R., Perera, U., 2012. Question answering through unsupervised knowledge acquisition. International Conference on Advances in ICT for Emerging Regions (ICTer2012). https://doi.org/10.1109/icter.2012.6421416.
Pinto, D., Gómez-Adorno, H., Vilarino, D., Singh, V.K., 2014. A graph-based multi-level linguistic representation for document understanding. Pattern Recogn. Lett. 41, 93–102. https://doi.org/10.1016/j.patrec.2013.12.004.
Prager, J. et al., 2007. Open-domain question-answering. Foundations Trends Inf. Retrieval 1 (2), 91–231.
Pudaruth, S., Boodhoo, K., Goolbudun, L., 2016. An intelligent question answering system for ict. International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT). IEEE, pp. 2895–2899. https://doi.org/10.1109/iceeot.2016.7755228.
Pustejovsky, J., Knippen, R., Littman, J., Saurí, R., 2005. Temporal and event information in natural language text. Lang. Resou. Eval. 39 (2), 123–164. https://doi.org/10.1007/978-1-4020-5958-213.
Rabiah, A.K., Sembok, T.M.T., Halimah, B., 2007. Towards skolemize clauses binding for reasoning in inference engine. International Conference on Computational Science and its Applications, ICCSA 2007. IEEE, pp. 273–282. https://doi.org/10.1109/iccsa.2007.24.
Rabiah, A.K., Sembok, T., Halimah, B., 2008. Utilization of external knowledge to support answer extraction from restricted document using logical reasoning. International Symposium on Information Technology, ITSim 2008, vol. 2. IEEE, pp. 1–8. https://doi.org/10.1109/itsim.2008.4631702.
Radev, D., Fan, W., Qi, H., Wu, H., Grewal, A., 2005. Probabilistic question answering on the web. J. Assoc. Inform. Sci. Technol. 56 (6), 571–583. https://doi.org/10.1002/asi.20146.
Ray, S.K., Singh, S., Joshi, B.P., 2010. A semantic approach for question classification using wordnet and wikipedia. Pattern Recogn. Lett. 31 (13), 1935–1943. https://doi.org/10.1016/j.patrec.2010.06.012.
Richardson, K.D., Bobrow, D.G., Condoravdi, C., Waldinger, R., Das, A., 2011. English access to structured data. 2011 Fifth IEEE International Conference on Semantic Computing (ICSC). IEEE, pp. 13–20. https://doi.org/10.1109/icsc.2011.67.
Sagara, T., Hagiwara, M., 2014. Natural language neural network and its application to question-answering system. Neurocomputing 142, 201–208. https://doi.org/10.1016/j.neucom.2014.04.048.
Saquete, E., Munoz, R., Martínez-Barco, P., 2006. Event ordering using terseo system. Data Knowl. Eng. 58 (1), 70–89. https://doi.org/10.1016/j.datak.2005.05.011.
Seena, I., Sini, G., Binu, R., 2016. Malayalam question answering system. Procedia Technol. 24, 1388–1392. https://doi.org/10.1016/j.protcy.2016.05.155.
Shekarpour, S., Marx, E., Ngomo, A.-C.N., Auer, S., 2015. Sina: semantic interpretation of user queries for question answering on interlinked data. Web Semantics: Sci. , Services Agents World Wide Web 30, 39–51. https://doi.org/10.1016/j.websem.2014.06.002.
Silva, J., Coheur, L., Mendes, A.C., Wichert, A., 2011. From symbolic to sub-symbolic information in question classification. Artif. Intell. Rev. 35 (2), 137–154.
Solorio, T., Pérez-Coutino, M., Montes-y Gémez, M., Villasenor-Pineda, L., López-López, A., 2004. A language independent method for question classification. Proceedings of the 20th international conference on Computational Linguistics.

Association for Computational Linguistics, p. 1374. https://doi.org/10.3115/1220355.1220556.

Spranger, M., Labudde, D., 2014. Establishing a question answering system for forensic texts. Procedia-Social Behav. Sci. 147, 197–205. https://doi.org/10.1016/j.sbspro.2014.07.152.

Sucunuta, M.E., Riofrio, G.E., 2010. Architecture of a question-answering system for a specific repository of documents. 2010 2nd International Conference on Software Technology and Engineering (ICSTE), vol. 2. IEEE, pp. V2–V12. https://doi.org/10.1109/icste.2010.5608753.

Surdeanu, M., Ciaramita, M., Zaragoza, H., 2011. Learning to rank answers to non-factoid questions from web collections. Comput. Ling. 37 (2), 351–383. https://doi.org/10.1162/colia00051.

Tapeh, A.G., Rahgozar, M., 2008. A knowledge-based question answering system for b2c ecommerce. Knowl.-Based Syst. 21 (8), 946–950. https://doi.org/10.1109/itng.2008.262.

Tartir, S., Arpinar, I.B., McKnight, B., 2011. Semanticqa: exploiting semantic associations for cross-document question answering. 2011 Fourth International Symposium on Innovation in Information & Communication Technology (ISIICT). IEEE, pp. 1–6. https://doi.org/10.1109/isiict.2011.6149593.

Terol, R.M., Martínez-Barco, P., Palomar, M., 2007. A knowledge based method for the medical question answering problem. Computers Biol. Med. 37 (10), 1511–1521. https://doi.org/10.1016/j.compbiomed.2007.01.013.

Toti, D., 2014. Aqueos: a system for question answering over semantic data. 2014 International Conference on Intelligent Networking and Collaborative Systems (INCoS). IEEE, pp. 716–719. https://doi.org/10.1109/incos.2014.13.

Varathan, K.D., Sembok, T.M.T., Kadir, R.A., 2010. Automatic lexicon generator for logic based question answering system. 2010 Second International Conference on Computer Engineering and Applications (ICCEA), vol. 2. IEEE, pp. 349–353. https://doi.org/10.1109/iccea.2010.219.

Vazquez-Reyes, S., Black, W.J., 2008. Evaluating causal questions for question answering. Mexican International Conference on Computer Science, ENC'08. IEEE, pp. 132–142. https://doi.org/10.1109/enc.2008.14.

Vila, K., Maźon, J.-N., Ferŕandez, A., 2011. Model-driven adaptation of question answering systems for ambient intelligence by integrating restricted-domain knowledge. Procedia Computer Sci. 4, 1650–1659. https://doi.org/10.1016/j.procs.2011.04.178.

Voorhees, E.M., Tice, D.M., 2000. Building a question answering test collection. Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval. ACM, pp. 200–207. https://doi.org/10.1145/345508.345577.

Walke, P.P., Karale, S., 2013. Implementation approaches for various categories of question answering system. 2013 IEEE Conference on Information & Communication Technologies (ICT). IEEE, pp. 402–407. https://doi.org/10.1109/cict.2013.6558129.

Wang, M., Manning, C.D., 2010. Probabilistic tree-edit models with structured latent variables for textual entailment and question answering. Proceedings of the 23rd International Conference on Computational Linguistics. Association for Computational Linguistics, pp. 1164–1172.

Wang, W., Auer, J., Parasuraman, R., Zubarev, I., Brandyberry, D., Harper, M., 2000. A question answering system developed as a project in a natural language processing course. Proceedings of the 2000 ANLP/NAACL Workshop on Reading comprehension tests as evaluation for computer-based language understanding systems, vol. 6. Association for Computational Linguistics, pp. 28–35. https://doi.org/10.3115/1117595.1117600.

Wen, D., Jiang, S., He, Y., 2008. A question answering system based on verbnet frames. International Conference on Natural Language Processing and Knowledge Engineering, NLP-KE'08. IEEE, pp. 1–8. https://doi.org/10.1109/nlpke.2008.4906769.

Wenyin, L., Hao, T., Chen, W., Feng, M., 2009. A web-based platform for user-interactive question-answering. World Wide Web 12 (2), 107–124.

Xie, N., Liu, W., An answer fusion model for web-based question answering, 2005. First International Conference on Semantics, Knowledge and Grid, 2005. SKG'05. IEEE. https://doi.org/10.1109/skg.2005.36 (pp. 8–8).

Shallow parsing natural language processing implementation for intelligent automatic customer service system, 2014. In: 2014 International Conference on Advanced Computer Science and Informtion System. IEEE, Jakarta, pp. 274–279. https://doi.org/10.1109/ICACSIS.2014.7065845.

Yang, M.-C., Lee, D.-G., Park, S.-Y., Rim, H.-C., 2015. Knowledge-based question answering using the semantic embedding space. Expert Syst. Appl. 42 (23), 9086–9104. https://doi.org/10.1016/j.eswa.2015.07.009.

Yao, X., 2014. Feature-driven Question Answering with Natural Language Alignment. John Hopkins University (Ph.D. thesis).

Yin, J., Zhao, W.X., Li, X.-M., 2017. Type-aware question answering over knowledge base with attention-based tree-structured neural networks. J. Computer Sci. Technol. 32 (4), 805–813. https://doi.org/10.1007/s11390-017-1761-8.

Yue, C., Cao, H., Xiong, K., Cui, A., Qin, H., Li, M., 2017. Enhanced question understanding with dynamic memory networks for textual question answering. Expert Syst. Appl. 80, 39–45. https://doi.org/10.1016/j.eswa.2017.03.006.

Yu, H., Lee, M., Kaufman, D., Ely, J., Osheroff, J.A., Hripcsak, G., Cimino, J., 2007. Development, implementation, and a cognitive evaluation of a definitional question answering system for physicians. J. Biomed. Inform. 40 (3), 236–251. https://doi.org/10.1016/j.jbi.2007.03.002.

Zayaraz, G. et al., 2015. Concept relation extraction using naïve bayes classifier for ontology-based question answering systems. J. King Saud Univ.-Computer Inform. Sci. 27 (1), 13–24. https://doi.org/10.1016/j.jksuci.2014.03.001.

Zhang, X., Hao, Y., Zhu, X.-Y., Li, M., 2008. New information distance measure and its application in question answering system. J. Computer Sci. Technol. 23 (4), 557–572. https://doi.org/10.1007/s11390-008-9152-9.

Zong, H., Yu, Z., Guo, J., Xian, Y., Li, J., 2011. An answer extraction method based on discourse structure and rank learning. 2011 7th International Conference on Natural Language Processing andKnowledge Engineering (NLP-KE). IEEE, pp. 132–139. https://doi.org/10.1109/nlpke.2011.6138181.