

Researches and Reviews in Arabic Question Answering: principal approaches and systems with classification

Wided Bakari¹, Patrice Bellot², and Mahmoud Neji¹

¹Faculty of Economics and Management 3018, Sfax Tunisia MIR@CL, Sfax, Tunisia

²Aix-Marseille University, F-13397, Marseille Cedex 20, LSIS, Marseille, France

Abstract: As, we live in the world of knowledge, there is increasingly high advancement in technology. Our thirst for knowledge enhances our integrity to ask questions to various search engines. A similar search engine is a question-answering system which provides perfect solution to retrieve valid and accurate answers to user question asked in natural language instead of a query. Internet includes websites in different languages; the need for such systems is higher in the context of the Arabic. Furthermore, the community of the Arabic language users is still obliged to manually looking for precise answers to their questions which is a tedious task regarding the great amount of available Web information. Indeed, Arabic has a complex structure, which makes it difficult to apply natural language processing (NLP). Much research on Arabic NLP does exist. However, it is not as mature as that of other languages. This paper provides a comprehensive and comparative overview of the Arabic question answering technology. Specifically, it introduces an overview on principal existing approaches and discusses the different proposed systems with a classification.

Keywords: Arabic language, question-answering, approaches, literature review, classification.

1. Introduction

In the 90s, with the emergence of the Internet and the expanding to the public of their use, the need for information search tools has become crucial. Indeed, one of the major challenges of information retrieval concerned the determination of the precise information sought by a user. This focus became a reality in recent years by bringing interest to the question-answering systems. Indeed, teaching machines how to automatically answer questions asked in natural language on any topic or in any domain has always been a long standing goal in Artificial Intelligence [16]. A question answering system works closely with the field of information retrieval. This requires information retrieval and natural language techniques. It is a difficult form of information retrieval characterized by information needs that are at least somewhat expressed as natural language statements or questions, and was used as one of the most natural type of human computer communication [23]. A question is defined as a natural language sentence, which usually starts with an interrogative word and expresses some information need of the user. Sometimes, a question has a form of an imperative construct and starts with a

verb. In such a case, the information request is called statement.

The basic idea behind the question-answering system is that the users just have to enter the question and the system will retrieve the most appropriate and precise answer for the question and return it to the user [15]. It was frequently produced in several companions and validation workshops, such as, TREC, CLEF, NTCIR, MUC, etc. At present, there is no standard approach to the question-answering systems; the architectures can be quite varied. A question-answer system has emerged as a good solution for the search engine which produces the notion of accurate information. In addition, the main objective of any question-answering system is to get answers to questions rather than full documents or the best corresponding passages. Nevertheless, we note a common structure or typical architecture for a question-answering system that provides generally three or four steps. Thus, there are many approaches and methodologies that have been proposed for different languages (Bulgarian, Dutch, English, French, German, Indonesian, Italian, Japanese, Portuguese, Arabic and Spanish, etc.). The main goal of this paper is to provide a comprehensive and comparative overview of question-

answering technology. It surveys the principal approaches adopted by the Arabic question-answering systems and classifies them based on different criteria. Thus, the main contributions of our work are four-fold. First, we propose a stacked attention of the question-answering process. Second, we perform a detailed analysis of the Arabic language and the question-answering. Third, we expose a comprehensive and comparative overview of question answering technology and principal approaches. Finally, and before concluding, we will make a discussion.

2. The question-answering process

Research in the area of question-answering began in 1960s and since then, a large number of question-answering systems have been developed. This discipline promotes systems for multiple domains, different data sources, different types of questions, specific formats of the answers, etc. The number of systems available each year is too large; these systems are experienced in multiple languages. Indeed, English has been considered the most studied language in this area. Subsequently, the question-answering tends to grow in favor of other languages, Chinese, Japanese, etc. From then on, the question-answering attempts to introduce new languages, European languages have been studied mainly in CLEF, Asian languages in NTCIR, French language in EQueR, Arabic in TREC and CLEF, etc. The research on the question-answering are relatively advanced for English and other Latin languages in relation to Arabic. Arabic question-answering systems are gaining great importance due to the increasing amounts of Arabic content on the Internet and the increasing demand for information that regular Information Retrieval techniques cannot satisfy [21]. Therefore, to meet the methodologies of these systems and their ability to meet user needs for accurate, systematic investigation of all these paradigms and approaches is needed.

In general, a question-answering system has three components such as question analysis, documents processing, and answer extraction. Question analysis is the first phase which consists of construction of question representation, it classifies user questions, derives expected answer types, extracts keywords, and reformulates a question into semantically equivalent multiple questions. Some systems also perform question reformulation, where a question is transformed into a number of declarative equivalents. Documents processing typically includes keyword expansion, document retrieval, and passage identification, it is the task in which, given a collection of documents where the answers can be found. Finally,

the answer extraction module is raising topics in the question-answering systems where these systems are often requiring ranking and validating a candidate's answer [23]. Answering natural language questions requires more complex processing of text than employed by current information retrieval systems. Arabic question-answering has not reached the same level due to the specific challenges of this language. As there are some challenges that are specific to Arabic language that forced the researchers to tackle them differently. Nevertheless, a number of question-answering systems have been developed which are capable of carrying out the processing required to achieve high levels of accuracy. Some attempts were made to reach an acceptable result in the Arabic question-answering technology. Most of these attempts suffered from over-fitting and subjective, non-realistic evaluation [37].

Users input the question from the web and return exact answer by various search engines; this also called as question-answering. Retrieving relevant documents on the web is a difficult task mainly when the query words do not convey what the user needs [24]. Several researchers, such as [19], [39], [10], [31] supports the idea that the question-answering is a specialized form of information retrieval, it is not only interested in obtaining relevant pages, but to get specific answers to a list of questions in natural language. It is necessary to know the input question type in order to suggest answers to users. For instance, question type classification is the task of determining the correct type of the expected answer to a given question [29]. Several types of questions exist and are often linked to a particular type of answers as well as different approaches to respond. A question-answering as a problem deals with many types of questions. In Arabic question-answering, factual questions have been the subject of many researches. Thus, the factual questions are one type that is concerned with questions that ask mainly about a Named Entity (NE) like questions using the words: When, Where, How much/many, Who, and What, which ask about a date/time, a place, a person, and an organization respectively [21]. Identifying the type of the question is an important step in the development of question-answering systems. For each question type there is a set of possible answers.

As shown in Table 1, the challenges of question types differ from the question type perspective. Obviously, when the project only concerns factoid questions and/or list and definition questions about NEs, the task as well the used techniques may not require advanced processes. On the contrary, when the system is open to other allowed types of questions, it is necessary to integrate modules that allow obtaining an in-depth understanding of question and documents through different techniques among which we can cite

Type	Example	Answer type	Description
Definition	Who is Vlad the Impaler?	Definition of an object or a person or organization	Questions that ask for important information's or definitions of someone or something.
Factual	Who invented the camera?	Named entity or specific information	Questions that fall into several distinct types and the answer is a single fact, such as "person", "location", "date", "organization", etc.
Boolean	Is gold mined in Canada?	Yes or NO	Questions that generally expect an answer either YES or NO.
List	What are the symptoms of diabetes?	List of entities.	Questions that imitate a dialogue that could have a person with a machine and the answer is present in list format.
Complex	How is body weight determined?	Detailed answers and not just named entities	Questions that use the words "Why" or "How" are another type that is hard to answer and there are very little if any attempts done to answer this type of questions [17].

deduction, text entailment, semantic reasoning, etc. A typical definition question-answering system extracts definition nuggets that contain the most descriptive information about the question target (the concept for which information is being sought is called the target term, or simply, the target) from multiple documents [41]. Answering factual questions is given a description of an entity; identify the person, place, or thing discussed. The system extracts answers as text

snippets from one or more documents. The system is often functionally equivalent to information retrieval. Given a knowledge base and a query, the goal is to return an answer [28].

Table 1. Question types classification

Nevertheless, complex questions cannot be answered using the same techniques that apply to factual questions. Complex questions refer to relations between entities or events; they refer to complex processes and model scenarios that involve deep knowledge of the topic under investigation [26]. Consequently, complex questions that require inferencing and synthesizing information from multiple documents can be seen as a kind of topic oriented, informative multi-document summarization. A list question expects a list as its answer. The ability to answer list questions has been tested as part of the yearly TREC question-answering evaluation [20]. Alternatively, for the Boolean questions, an answer can be YES or NO. For instance, as it is mentioned in [40], answering YES/NO questions requires deeper linguistic and semantic understanding of the questions and relevant snippets, which includes leveraging semantic networks of concepts to identify antonyms, hypernyms, hyponyms, and utilizes dependency relations between the concepts, as well as sentiment analysis of the facts.

3. Arabic language and the question answering

The question-answering conducted a comprehensive performance in various languages, since its emergence, including English and some other Latin languages that are much benefited by the progress in the area of question-answering. However, in Arabic question-answering, challenges are not limited to those commonly faced by systems developed for other languages. Thus, each integrated Arabic NLP component may positively or negatively impact the performance of the system unless it considers the particularities of this language such as its complex morphology and syntax, its high ambiguity, the absence of capital letters, etc. More specifically, Arabic Language is the mother tongue for 23 countries and more than 350 million persons. Moreover, since it is the language of the Holy Quran, many other Islamic countries, teach Arabic as a second language. Nevertheless, it is noticed that the Arabic content on the web is less than what should be expected. It is the official language of millions of people and is the

religious language of all Muslims. It is a Semitic language of 28 alphabets, and it is one of the United Nations official languages [14].

Among these specialties, its complex morphological, grammatical, and semantic aspects since it is a highly inflectional and derivational language. Furthermore, the Arabic Language is highly ambiguous by several reasons. One of these is that the Arabic Language has not capitalization property, which directly influences and complicates the identification of the Arabic Named Entities. More specifically, the importance of the technology of question-answering is very obvious and relevant for the information retrieval. Therefore, the NLP tools that designed for the English Language can't directly accommodate the needs of the Arabic Language. Moreover, the studies that have been proposed for Arabic cannot keep the same pace of change because of the specific challenges of this language. It is practically, from 2004, a few exploratory attempts have been made around the Arabic question-answering systems.

Given that, according to a literature review, we note that almost all of the Arabic question-answering systems process morphosyntactic approaches and few of them provide approaches based on inference and logic. However, despite this lack, there are some attempts in the literature that have achieved acceptable results in this area of research. To the best of our knowledge, the Arabic question-answering technology was studied, among others, since the 1990s (Figure 1). Due to the rarity of these systems, the need for such question-answering systems is higher in the context of Arabic [30]. Indeed, the developed Arabic question-answering systems are still few and immature because of the unique aspects of the Arabic language. This is mainly due to two reasons: the lack of accessibility to language resources and tools (i.e. the corpus and the Arabic NLP tools) and the very complex nature of the language itself (i.e. Arabic is not inflected - concatenative). Arabic has a tremendous religious significance in Islam considering that: the Quran, one of the four Holy books, was revealed in Arabic, and that over 1.2 billion Muslims in the world pray five times a day using the Arabic language. Geographically, Arabic is an official language in 25 countries including the members of the Arab league. Therefore, we illustrate in various works [3], [4] that there are several motivating factors for choosing the Arabic.

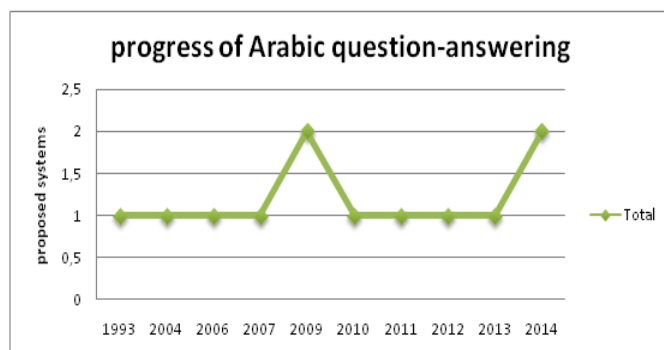


Figure 1. Evolution of the question-answering in Arabic since its appearance

4. Researches and reviews in Arabic approaches and systems

A good overview of interactive Arabic question-answering can be found in [5]. A number of question-answering systems have been developed which are capable of carrying out the processing required to achieve high levels of accuracy. Existing question-answering systems are unable to handle variety of questions and reasoning based question. Also, because of the lack of the data sources, the question-answering systems fail to answer the question. When discussing the main methods and tools used in question-answering, many different classifications can potentially be followed. Question-answering is a form of information retrieval, and any retrieval model is defined by the representation of the information need, by the representation of the retrievable object, and by the retrieval or ranking function. The most instructive way to describe the methods starts from the simplest representations derived from the natural language statements and corresponding retrieval functions, and gradually moves to more advanced analysis techniques and retrieval functions that incorporate some degree of reasoning. In this paper, a comparative analysis of several approaches adopted by the question-answering systems with classification is presented.

4.1. Approaches

There is an extensive literature on Arabic question-answering, in which researchers were proposed several approaches, developed their systems and evaluated them by different questions types to produce answers. Through, a detailed analysis of existing studies on Arabic question-answering systems, we found that for more than a decade, Arabic question-answering has been important by researchers in implementing new approaches and new systems. The approaches show that there is not really a standard method to be adopted when it comes to extract the correct answers to a question. Everything will depend on the type of problems we want to treat, and the context of work in which you want to register.

Referring to previous research works, we can divide the attempts of main approaches into three parts: the morph syntactic approaches, the hybrid approaches and the semantic approaches. These approaches are based on a representation of questions and documents that can be varied depending on the system and the choice of tools used. In [7], [5], [8], [6], we follow the category of hybrid approaches and propose a new

logical approach to implement a question-answering system to answer factual questions in Arabic.

4.1.1. Morphosyntactic approaches

In these kinds of approaches, generating an answer based extensively on techniques of Natural Language Processing and Information Retrieval and does not necessarily imply a semantic analysis of the question or recovered documents [12]. We must remember that in the existing of proposed approaches, virtually all research focuses on morph syntactic approaches. The answer search usually requires an indexing step to identify named entities or phrases that best exhibit the content of a document. Indeed, linguistic processing such as morphological analysis and / or syntactic of documents can also enrich these kinds of approaches. First, we can highlight a first work which introduces an approach based on the Man Machine Interface [32]. Another family of works that are based on linguistic approaches used the named entity recognition [25], [13], [11], [30] and on the platform Nooj [18]. The authors in [38] proposed an approach based on lexical patterns to identify definitions and heuristic rules to filter these definitions. In addition, to the named entity recognition, the automatic text comprehension has been the subject of some works [22]. Indeed, automatic text comprehension serves to answer questions based on a single document or a text with the correct answers require some type of inference and a knowledge base review previously acquired [9].

4.1.2. Semantic approaches

Like any other language, Arabic semantics needs to interpret the meaning of the question and the documents to extract the exact answer. Indeed, language is seen as a major challenge in terms of answering questions which every language has its own morphology and semantic. Indeed, a semantic approach typically performs a semantic analysis of the question and the documents and produced a formal representation of their meaning [12]. Very little works has been done in the Arabic question answering using semantic approaches to serve the semantic analysis of the questions or the documents. For example, the work of [2] have used the density distance of the N-gram model and semantic expansion using Arabic WordNet and even contributed to the richness of this ontology. The approach presented in [1] is to interpret the semantics of the Quran to answer the user's questions using the Quran and its interpretation pounds.

4.1.3. Hybrid approaches

Logic is a level between syntax and semantic [33], it is the most important and difficult level in natural language processing. Approaches based the logic are a rich topic of research for the Arabic question-answering. Referring to this definition, we can say that the approaches based on the logic-inference are considered as hybrid approaches between those morph syntactic and semantic ones. Although, more logical approaches have been proposed for other languages, there are few similar systems for the Arabic language. To our knowledge, there is only one attempt to work presented by [35]. This approach promotes a logical representation of questions to retrieve the paragraph that contains relevant answers to these questions. The approach presented in our work fills this gap [7]. Actually, this kind of approaches has been shown to give good results for English. Indeed, the use of the logic based approaches and inference started long time, especially for English, we therefore apply the same method to the Arabic. Examples of approaches that have inspired us those proposed by [27], [34], [36], etc. Our approach involves a step of logic transformation. At this level, this is performed to convert information in natural language into a set of logical predicates [33].

4.2. Systems

To our knowledge, existing Arabic question-answering systems are limited either in terms of their scope as well as in terms of performance regarding the types of questions that they are designed to answer. Nowadays, most of the question-answering systems address questions in Latin languages (English, French, Chinese Japanese, etc.). The need to develop question answering systems dedicated to the Arabic language is becoming increasingly necessary due to certain difficulties inherent in the language itself; as well as lack of tools available to help researchers. In this section, we identified some systems developed in this language from their birth with AQAS [32] to Al-Bayan [1]. In this case, the experimental literature in Arabic gives us the ability to classify these systems into four groups. Moreover, our classification was carried out at the base of many tools and techniques extensively used by each system.

4.2.1. Systems based on the interrogation of databases

The systems that bring this category introduce an approach based on the human machine interface. Indeed, these systems transform question into a query and interrogate databases to extract the answer. AQAS [32] is considered among the first systems that have

emerged in the 60s; it seeks answers from structured databases. However, QARAB [25] seeks answers to questions from unstructured documents extracted from the Al-Raya newspaper.

4.2.2. Systems based on NLP and IR techniques

We illustrate, in this category, the main systems that rely extensively on NLP and IR techniques to extract the answer. Indeed, almost all systems treat a particular type of questions as to the factual question and are based on morphosyntactic approaches. For example, ArabiQA [13] uses the techniques of named entities; QASAL [18] based on the Nooj platform to extract the answer from an educational book. By against, AquASys [11] analyzes the question and extracts the answer from a corpus. In addition, JAWEB [30] has been built on the basis of AquASys by providing a user interface as an extension.

4.2.3. Systems based on automatic text comprehension

A third categorization is based on the automatic text comprehension to answer questions. Extracting the right answers requires a certain type of inference and knowledge base review previously acquired [9]. In this context, most of the systems favoring an automatic comprehension of a text then use preprocessing, such as anaphora resolution, coreference, or the named entity recognition. In fact, IDRAAQ [2] participates in QA4MRE @ CLEF task. This task is introduced for the first time, in 2012, for the Arabic. Moreover, ALQASIM proposed by [22] is based on the selection and validation of the answer, it answered multiple choice questions.

4.2.4. Systems based on logic and inference

Another class of question-answering systems is based on logical reasoning and inference to find the right answers to a question in natural language. To our knowledge, this kind of approaches is little used until now in the Arabic question-answering. The use of logic and inference in this field has been the subject of few studies to know [35]. The latter aims to promote a semantic representation constraint using an explicit unification framework based on semantic similarities and expansion of the query (synonyms and antonyms). Also, our study focuses in this context. Referring to our detailed analysis above, our classification can be well illustrated in (Table 2) which we present some examples for each category of systems. Thus, for every system we show the techniques and tools it adopts.

Table 2. Classification of Arabic question-answering systems

Category	criteria/Technical used	Examples of systems
Systems based on the interrogation of databases.	Structured database	AQAS [32]
	Unstructured database	QARAB [25]
Systems based on NLP and IR techniques.	Named entity recognition	ArabiQA [13]
	Platform Nooj	QASAL [18]
	Lexical patterns	DefArabicQA [38]
	Named entity recognition	AquASys [11]
	Automatic Language Processing NLP	JAWEB [30]
	Named entity recognition	AL-Bayan [1]
Systems based on automatic text comprehension.	Distance density N-gram model and semantic expansion	IDRAAQ [2]
	Selection and validation of the answer	ALQASIM [22]
Systems based on logic and inference.	IR information retrieval and artificial intelligence AI and NLP	System of [35]

By summarizing, our concern is based on the specific approaches based on the logic. Our investigation proposes a new purely logical approach to address the Arabic questions. Our approach combines automatic processing techniques of natural language, the information retrieval and artificial intelligence techniques. It promotes transformation of Arabic statements in logical forms. Our approach makes the integration of logic and inference for the Arabic question-answering systems.

4. Discussion

The question-answering systems can be differentiated according to their approaches, techniques and studied tools. Therefore, since the introduction of question-answering in conferences and workshops as soon TREC-8 in 1999, there has been a dramatic increase in question-answering systems both theoretical and empirical level. In addition, English was once considered the universal language, particularly on the Web. From then on, question-answering should evolve to handle new languages; few studies have focused on the Arabic. The scarcity of studies for Arabic may, in some way, explains the difficulty of the language. In fact, this survey paper, like any other survey, provides

our effort that takes a comprehensive overview of the question-answering research to meet the challenges due to information explosion in this information and communication technology era. As we can see the approaches taken by these systems are diverse. There is also evidence that these approaches often share common points.

Anyway, we cannot say that there is not a more efficient system than the other. Also, the production of a new system does not preclude the existence of an old one. Despite the benefits of these systems they had some limitations. Moreover, Arabic is among the languages that are less affected by the researchers in question-answering. So, we found that the success rate of these Arabic question-answering systems is less developed compared to other languages, such as the English. After a thorough review of approaches and systems to answer questions in the Arabic, we explore upstream that almost all previous works focuses on the analysis of the question to better select the correct answer. We observed also that the choice of NLP techniques which are used for analysis is highly problem specific.

5. Conclusion

The research in question answering has emerged since the 1960s. There are many approaches and methodologies that have been proposed for different languages (Bulgarian, Dutch, English, French, German, Indonesian, Italian, Japanese, Portuguese, Arabic and Spanish, etc.). Arabic question answering is not received from several studies in theoretical and empirical level; this lack is due to the specifics of this language. This paper illustrates a view of sets of existing studies in the field of Arabic question answering. Especially, it presents a review of the literature of different systems and approaches proposed for the Arabic. Our survey considered most of the systems and approaches. It includes different classifications of Arabic question-answering systems and their adopted approaches. To the best of our knowledge, which are concerned with the morph syntactic approaches in Arabic question-answering are more studied rather than those based on the logic-inference and semantic. It can be deduced that the Arabic Language is still far away from this kind of approaches. This was the motivation to propose a new approach and provide tracks for this approach.

References

- [1] Abdelnasser, H., Mohamed, R., Ragab, M., Mohamed, A., Farouk, B., El-Makky, N., & Torki, M. (2014). Al-Bayan: An Arabic Question Answering System for the Holy Quran. ANLP 2014, 57.
- [2] Abouenour, L., Bouzoubaa, K. and Rosso, P. , "IDRAAQ: New Arabic Question Answering System Basedon Query Expansion and Passage Retrieval", CLEF (Online Working Notes/Labs/Workshop). 2012.
- [3] Abufardeh, S. and Magel, K., "Software localization: The Challenging Aspects of Arabic to the Localization Process (Arabization)", IASTED Proceeding of the Software Engineering SE 2008. Innsbruck, Austria. 2008. Pp: 275-279.
- [4] Akour, M., Abufardeh, S., Magel, K. and Al-Radaideh, A. Q., "QArabPro: A Rule Based Question Answering System for Reading Comprehension Tests in Arabic", American Journal of Applied Sciences 8 (6): 652-661, 2011, ISSN 1546-9239.
- [5] Bakari, W., Bellot, P., & Neji, M. (2015). Literature Review of Arabic Question-Answering: Modeling, Generation, Experimentation and Performance Analysis. In Flexible Query Answering Systems 2015 (pp. 321-334). Springer International Publishing.
- [6] Bakari, W., Bellot, P., & Neji, M. (2016). AQA-WebCorp: Web-based Factual Questions for Arabic . Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 20th International Conference KES-2016. Procedia Computer Science, Volume 96, 2016, Pages 275–284. (b).
- [7] Bakari, W., Trigui, O., and Neji, M. (2014, December). Logic-based approach for improving Arabic question answering. In Computational Intelligence and Computing Research (ICCIC), 2014 IEEE International Conference on (pp. 1-6). IEEE.
- [8] Bakari, W., Bellot, P. and Neji, M. A Preliminary Study for Building an Arabic Corpus of Pair Questions-texts from the Web: AQA-WebCorp. iJES 2016, Volume 4, 2016. (a)
- [9] Banerjee, S., Bhaskar, P., Pakray, P., Bandyopadhyay, S., & Gelbukh, A. F. (2013, September). Multiple Choice Question (MCQ) Answering System for Entrance Examination. In CLEF (Working Notes).
- [10] Bauer, M. A., & Berleant, D. (2012). Usability survey of biomedical question answering systems. Human genomics, 6(1), 1.
- [11] Bekhti, S., Rehman, A., AL-Harbi, M. and Saba, T., "AQUASYS: an arabic question-answering system based on extensive question analysis and

- answer relevance scoring”, In International Journal of Academic Research; Jul2011, Vol. 3 Issue 4, p45.
- [12] Ben-Abacha, A., “Recherche de réponses précises à des questions médicales : le système de questions-réponses MEANS”, PhD thesis. Université PARIS-SUD 11 LIMSI-CNRS. JUIN 2012.
 - [13] Benajiba, Y., Rosso, P. and Lyhyaoui, A., “Implementation of the ArabiQA Question Answering System's components”, In: Proc. Workshop on Arabic Natural Language Processing, 2nd Information Communication Technologies Int. Symposium. ICTIS-2007, Fez, Morocco, April 3-5, 2007.
 - [14] Beseiso, M., Ahmad, A. R., & Ismail, R. (2011, June). An Arabic language framework for semantic web. In 2011 International Conference on Semantic Technology and Information Retrieval (pp. 7-11). IEEE.
 - [15] Biswas, P., Sharan, A., & Malik, N. (2014, February). A framework for restricted domain Question Answering System. In Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 International Conference on (pp. 613-620). IEEE.
 - [16] Bordes, A., Chopra, S., & Weston, J. (2014). Question answering with subgraph embeddings. arXiv preprint arXiv:1406.3676.
 - [17] Bouziane, A., Bouchiha, D., Doumi, N., & Malki, M. (2015). Question Answering Systems: Survey and Trends. Procedia Computer Science, 73, 366-375.
 - [18] Brini, W., Ellouze, M., Mesfar, S. and Belguith, L. H., “An Arabic Question-Answering system for factoid questions”, In IEEE International Conference on Natural Language Processing and Knowledge Engineering (IEEE NLP-KE'09). Dalian, China. 2009.
 - [19] Cao Y., Liu F., Simpson P., Antieau L., Bennett A., Cimino JJ., Ely J. and Yu H. (2011) “AskHERMES: an online question answering system for complex clinical questions”, J Biomed Inform, 44(2):277–288.
 - [20] Dang, H.T., Kelly, D., and Lin, J.. 2007. Overview of the TREC 2007 question answering track. Proceedings of the Sixteenth Text REtrieval Conference.
 - [21] Ezzeldin A. M. and Shaheen M. (2012). “A survey of Arabic question answering: challenges, tasks, approaches, tools, and future trends”, the 13th International Arab Conference on Information Technology ACIT'2012. Dec.10-13. ISSN 1812-0857.
 - [22] Ezzeldin, A. M., Kholief, M. H. and El-Sonbaty, Y. , “ALQASIM: Arabic Language Question Answer Selection in Machines”, CLEF 2013. 2013. P 100-103.
 - [23] Gupta, P., & Gupta, V. (2012). A survey of text question answering techniques. International Journal of Computer Applications, 53(4).
 - [24] Hammo, B. (2009). Towards enhancing retrieval effectiveness of search engines for diacritized Arabic documents. Information Retrieval, 12(3), 300–323.
 - [25] Hammo, B., Ableil, S., Lytinen, S. and Evens, M. (2004). “Experimenting with a Question Answering system for the Arabic language”, In Computers and the Humanities. Vol. 38, N°4. Pages 397—415.
 - [26] Harabagiu, S., Lacatusu, F., & Hickl, A. (2006, August). Answering complex questions with random walk models. In Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 220-227). ACM.
 - [27] Harabagiu, S.M., Pasca, M.A., and Maiorano, S.J., (2000), Experiments with open-domain textual question answering, in: Proceedings of the 16th Int'l Conf. Comp. Ling. (COLING 2000), Saarbrücken, Germany, 2000, pp. 292–298.
 - [28] Iyyer, M., Boyd-Graber, J. L., Claudino, L. M. B., Socher, R., & Daumé III, H. (2014). A Neural Network for Factoid Question Answering over Paragraphs. In EMNLP (pp. 633-644).
 - [29] Khoury, R. (2011, June). Question type classification using a part-of-speech hierarchy. In International Conference on Autonomous and Intelligent Systems (pp. 212-221). Springer Berlin Heidelberg.
 - [30] Kurdi, H., Alkhaider, S. and Alfaif, N. (2014). “Development and evaluation of a web based question answering system for arabic language”, In Computer Science & Information Technology (CS & IT), 04(02), 187 – 202.
 - [31] Mishra, A., & Jain, S. K. (2015). A survey on question answering systems with classification. Journal of King Saud University-Computer and Information Sciences.
 - [32] Mohammed, F., Nasser, K. and Harb, H. (1993). “A knowledge based Arabic Question Answering system (AQAS)”, ACM SIGART Bulletin, pp. 21-33.
 - [33] Moldovan, D., Rus, B., (2001), Logic form transformations of WordNet and its applicability to question answering, in: Proceedings of the 39th Ann. Meeting Assoc. Comp. Ling. (ACL 2001), ACL, Toulouse, France, 2001, pp. 402–409.
 - [34] Mollá, D., Schwitter, R., Hess, M., Fournier, R., (2000), ExtrAns, an answer extraction system, TAL 41 (2000) 495–522.

- [35] N Bdour, W., and Gharaibeh, N. K. (2013) "Development of Yes/No Arabic Question Answering System", In International Journal of Artificial Intelligence and Applications (IJAIA),. Vol.4. No.1, January 2013. DOI: 10.5121/ijaia.2013.4105 51.
- [36] Rinaldi, F., Dowdall, J., Schneider, G. (2004). "Answering questions in the genomics domain", Dans Proc. ACL04 Workshop on Question Answering in Restricted Domains.
- [37] Shaheen, M., & Ezzeldin, A. M. (2014). Arabic question answering: Systems, resources, tools, and future trends. Arabian Journal for Science and Engineering, 39(6), 4541-4564.
- [38] Trigui, O., Belguith, L.H and Rosso, P., "DefArabicQA: Arabic Definition Question Answering System", In Workshop on Language Resources and Human Language Technologies for Semitic Languages, 7th LREC. Valletta, Malta. 2010.
- [39] Wren, JD. (2011). "Question answering systems in biology and medicine – the time is now", Bioinformatics, 27(14):2025–2026.
- [40] Yang, Z., Zhou, Y., Nyberg, E. "Learning to Answer Biomedical Questions: OAQA at BioASQ 4B." ACL 2016 (2016): 23.
- [41] Zhang, Z., Zhou, Y., Huang, X., & Wu, L. (2005, October). Answering definition questions using web knowledge bases. In International Conference on Natural Language Processing (pp. 498-506). Springer Berlin Heidelberg.

Wided BAKARI is a PhD student at Faculty of Economic Sciences and Management of Sfax (FSEGS)/Tunisia. She is currently a member of the Multimedia and InfoRmation Systems and Advanced Computing Laboratory (MIR@CL).

Patrice BELLOT is a professor at Aix-Marseille University, Head of the project-team DIMAG, Aix-Marseille University - CNRS, UMR 7296 LSIS.

Mahmoud Neji is an associate professor at the Faculty of Economics and Management of Sfax, University of Sfax, in Tunisia. He is a member of the Multimedia and InfoRmation Systems and Advanced Computing Laboratory (MIR@CL).