

Review

Ontology-Based Approach to Semantically Enhanced Question Answering for Closed Domain: A Review

Ammar Arbaeen ^{1,*}  and Asadullah Shah ² 

¹ Department of Computer Science, Faculty of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur 53100, Malaysia

² Faculty of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur 53100, Malaysia; asadullah@iium.edu.my

* Correspondence: arbaeen.ammar@live.iium.edu.my

Abstract: For many users of natural language processing (NLP), it can be challenging to obtain concise, accurate and precise answers to a question. Systems such as question answering (QA) enable users to ask questions and receive feedback in the form of quick answers to questions posed in natural language, rather than in the form of lists of documents delivered by search engines. This task is challenging and involves complex semantic annotation and knowledge representation. This study reviews the literature detailing ontology-based methods that semantically enhance QA for a closed domain, by presenting a literature review of the relevant studies published between 2000 and 2020. The review reports that 83 of the 124 papers considered acknowledge the QA approach, and recommend its development and evaluation using different methods. These methods are evaluated according to accuracy, precision, and recall. An ontological approach to semantically enhancing QA is found to be adopted in a limited way, as many of the studies reviewed concentrated instead on NLP and information retrieval (IR) processing. While the majority of the studies reviewed focus on open domains, this study investigates the closed domain.

Keywords: knowledge-based; natural language processing; ontology-based approach; question answering systems; semantic similarity



Citation: Arbaeen, A.; Shah, A. Ontology-Based Approach to Semantically Enhanced Question Answering for Closed Domain: A Review. *Information* **2021**, *12*, 200. <https://doi.org/10.3390/info12050200>

Academic Editor: Diego Reforgiato Recupero

Received: 4 April 2021
Accepted: 21 April 2021
Published: 1 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Technological advancements in the field of information communication technology (ICT) support the conversion of the conventional web structures of internet documents to semantic web-linking data, enabling novel semantic web data representation and integration retrieval systems [1]. Globally, millions of internet searches are made every minute, and obtaining precise answers to queries can be challenging, due to the continually expanding volume of information available [2].

Information retrieval (IR) technology supports the process of searching for and retrieving information. The general method used by IR employs keywords to retrieve relevant information from different heterogeneous resources [3], thereby providing a limitation that captures an accurate conceptualization of the user's expression and the meaning of their content. Upon searching, the user receives a list of related documents that may contain the information required. However, in some cases, users might be seeking for a specific answer, rather than a list of documents. Although IR techniques can be highly successful and retrieve relevant information, users nevertheless face complex linguistic challenges to extract the desired information. Question answering (QA) techniques can address this by enabling users to express and extract precise information and knowledge in natural language (NL).

Typically, QA systems aim to create methods with functionality that extends beyond the retrieval of appropriate texts, in order to provide the correct response to NL queries posed by a user. Answering such queries requires composite processing of documents

that exceeds that which is achievable using IR systems [4]. With QA systems, answers are obtained from large document collections by grouping required answer types, to determine the precise semantic representation of the question's keywords, which are essential to carry out further tasks such as document processing, answer extraction, and ranking. To achieve high semantic accuracy, precise analysis of the questions posed by users in NL is essential to define the intended sense in the context of the user's question. This challenge motivated us to investigate the various ontology-based approaches and existing methodologies designed to semantically enhance QA systems.

The semantic web implements an ontological methodology intended to retain data in the form of expressions associated with the domain, as well as procedures, referencing, and comments, and properties that are automated or software agent readable, and plausible for use processing and retrieving data [5,6]. This facilitates the power of widely distributed data, by linking it in a standardized and meaningful manner so that it is available on different web servers, or on the centralized servers of an organization [7]. Researchers usually utilize measures of similarity to discover identical notions in free text notes and records. A measure of semantic similarity adopts two notions as input, and then provides a numerical score, which calculates how similar they are in terms of senses [8]. A variety of semantic similarity measures have been developed to describe the strength of relationships between concepts. These existing semantic similarity measures generally fall into one of two categories: ontology-based or corpus-based approaches (i.e., a supervised method).

Ontology-based semantic similarities usually rely on various graph-based characteristics, for instance, the shortest path length between notions, as well as the location of lowest prevalent ancestors when identifying the similarity in meaning. Such ontology-based methods rely on the quality and totality of fundamental ontologies [8]; nevertheless, maintaining and curating domain ontologies is a challenging and labour-intensive task. It is worth noting that corpus-based approaches as an alternative to identifying ontology-based semantic similarities, are grounded on the co-occurrences and distributional semantics of terms in free texts [9]. Such corpus-based models depend on the linguistic principle that the sense of a lexical term can be identified on the basis of the surrounding context.

Although corpus-based approaches can outperform ontology-based approaches, the literature has shown that the performance gap between both kinds of approaches has narrowed [8,10]. Today, some ontology-based approaches reportedly achieve better performance than the corpus-based approach. This is because corpus-based approaches require a set of sense-annotated training data, which is extremely time-consuming and costly to produce. Moreover, corpora do not exist for all domains [11,12]. In order to explore the various approaches to overcoming this challenge, this study employs the following research methodology:

2. Research Methodology

The principal aim of this paper is to briefly review existing searching methods, focusing on the role of different technologies, and highlighting the importance of the ontology-based approach in searches, as well as trying to identify the major technology gaps associated with existing methods. The following steps describe the research methodology employed.

2.1. Research Objectives

The following objectives were determined:

1. To review the various methodologies currently available in the ontology-based approach to QA systems.
2. To understand the unique features of the majority of ontology-based approaches, across QA domains.
3. To determine what methodology best suits the development of a closed domain "Pilgrimage" ontology-based QA.

4. To establish the major challenges and study gaps affecting current methods used to semantically enhance QA systems to establish future directions.

2.2. Research Questions

This review comprises four questions:

- a. What are the methodologies available in the ontology-based approach to QA systems?
- b. What are the unique features of current ontology-based approaches, across QA domains?
- c. Which methodology best suits the development of closed domain ontology-based QA systems?
- d. What are the major challenges and study gaps affecting the present methodology across QA systems in terms of future directions?

2.3. Criteria for Inclusion and Exclusion Study

To examine the evidence for ontology-based methods for semantically enhancing QA for a closed domain, this study employed a tree table (Table 1) after rapidly examining 124 articles related to ontology-based approaches located on the Web of Science (WoS), limited by date of publication (2000–2020). The data were collected from the WoS Core Collection, constituted by the Social Sciences Citation Index (SSCI), Science Citation Index Expanded (SCI-EXPANDED), Conference Proceedings Citation Index - Science (CPCI-S), Conference Proceedings Citation Index-Social Science & Humanities (CPCI-SSH), Arts & Humanities Citation Index (A&HCI), and the newly included Emerging Sources Citation Index (ESCI). None of the selected articles afforded us exact data. Therefore, we excluded information deemed irrelevant by adding inclusion and exclusion criteria. In the first phase, we remove 145 irrelevant articles out of 269. In the second phase, we selected articles based on titles, abstracts and keywords, and rejected a further 66 papers. In this phase, we excluded 45 articles because the abstract did not align with the study's inclusion criteria. In the fourth phase, 19 were excluded for unusable data, and finally 26 were chosen for the literature review, as mentioned in Figure 1. The main categories created to classify the papers reviewed included the metadata of the paper, the problem identified, methods used, findings, future studies, techniques, and evaluation metrics (Table 1).

Table 1. Five categories created to classify the papers reviewed.

Category	Description
1	Metadata of the paper, comprising the title of the paper, author, date, problems identified, methods used, findings and future studies [13]
2	Defined which main techniques were selected for the paper e.g., Ontology based, Natural Language Processing, Semantic Web Links Data/ Knowledge based, hybrid based, etc. [14]
3	Was based on ontology-based techniques used for the approach [13]
4	Deals with the metrics used to evaluate the precision and recall for the paper reviewed on question answering (QA) [13]
5	Detailed classifications such as method and findings used [13]

The remainder of this paper has been organized as follows, Section 2 states the research methodology in which the research process, research questions and criteria for the inclusion and exclusion study were given along with search strategy. Section 3 introduces related works. Section 4 concerns the challenges of the QA systems. Section 5 presents the evaluation of QA systems, and a detailed discussion is presented in Section 6. Finally, conclusions and suggestions for recommended future work are set out in Section 7.

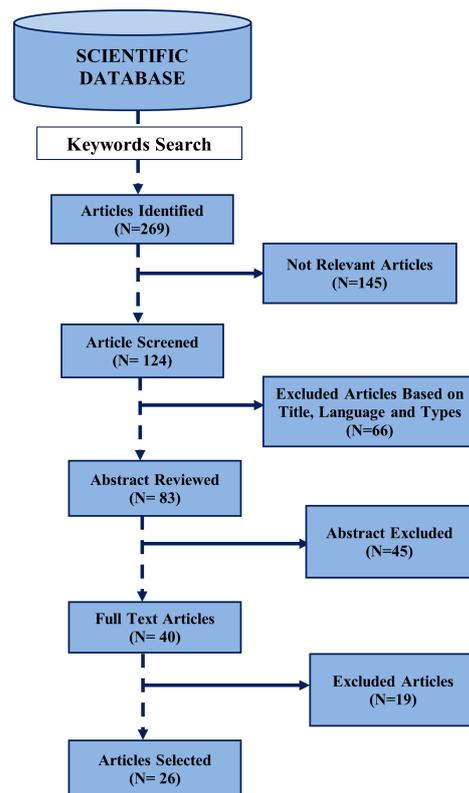


Figure 1. Criteria for Inclusion and Exclusion Study.

3. Related Works

Semantic technology considers ontology to be a form of knowledge representation, in which data are described and stored in the form of rich vocabulary, expressive properties, and the object relationships of a domain in machine-readable format. While QA systems seek to improve upon existing technology in order to identify relevant results and enhance the user experience, the majority of systems face many difficulties, due to the continual expansion in the amount of web content [15,16], even when the results returned may contain appropriate information that offers valuable knowledge to users in a specific domain. Therefore, effective content description and query processing techniques are vital for developing appropriate information resources that provide accurate results. Early QA systems in the 1960s, like BASEBALL [17], LUNAR [18], START [19] were simply NL front-ends for structured database query systems. The queries provided to such systems were often analysed utilizing NLP methods to generate a canonical form, that could subsequently be employed to form a standard database question. These systems are typically used to search a specific domain when requiring long-term data.

The majority of previous studies concerning open domain systems have addressed various QA systems. Among the earliest systems were MUDLER [20] and AnswerBus [21]; MUDLER uses NL parsing and a new voting approach to obtain reliable answers by employing multiple search engine queries. Its recall requires Google, and approximately 6.6 times the effort in searches by users to obtain the desired recall. Similarly, the AnswerBus QA system is a sentence level web IR method that can accept users' NL questions in different languages, such as English, French, Spanish, Italian, and Portuguese, and provides answers in English. It uses five different search engines and directories to provide relevant answers by skimming the directories to locate sentences that contain appropriate answers. AskJeeves [22] is another open domain QA system that directs users to relevant web pages, using an NL parser to solve the question posed in a query. It does not provide direct answers to the user, rather, it refers users to different web pages in the same way as a search engine, requiring only half the work of a QA system [20]. Meanwhile, Story Understanding Systems, such as the QUALM QA system employs an IR system approach and works by

asking questions about simple paragraph-length stories, providing a question analysis module that connects each question with a question-type that guides the processing of IR. In addition, an ontology-based domain exists that uses the AQUALOG QA system that processes input queries and classifies them into 23 categories, defined by the system [23]. QA systems by [24,25] have developed the web as their knowledge resource. Such systems use their own heuristics to store information from web documents in a local knowledge database which can be accessed later and relies on linguistic techniques for generating answers. With the growth of text-based materials online, the need for a statistical method for the QA systems has been largely augmented. Clark [26] provided a method for supplementing online texts (dynamic manual) with a knowledge-based question answering capability. This combined method makes it possible for users to access not only answers to routine questions, but also questions that were unpredicted at the time the system was constructed. This particular feature of the QA system is accomplished via an inference engine constituent.

3.1. Knowledge-Based QA Systems

Knowledge-based systems are a subset of AI that were designed to extract and generate knowledge from a defined data source, known as a Knowledge Base (KB), as an alternative to unstructured text or documents [27]. As [28] explained, standard database questions are used in exchange for word-based queries. This formation employs defined data, such as ontology in QA. It is widely understood that knowledge is central to many common tasks in a QA system. The main types of knowledge that may require representation in AI systems are objects, events, performance, and meta-knowledge [29,30]. According to [31], knowledge can be categorized into two forms: general knowledge and specific knowledge. The former describes world knowledge, and the latter a specific situation in which linguistic communication occurs. Knowledge-based systems generally extract certain information in advance for later use as in [32], and [33] observed that QA researchers have increasingly tended to employ the ontology-based method when developing semantically enhanced QA for natural language questions (NLQs). The availability of knowledge resources in different domains provides opportunities to utilize these sources in QA systems. The ontology domain is typically small and the data analysis does not require high-level automation. In connection with this, [34] offered a template for a representation model, in which the task developed concerned mapping the question proposed to the available template. To perform this, the entity in question is substituted with selected notions. This process is significant to consider, and is accomplished through a mechanism called Conceptualization. Utilizing templates, a complex question could be disintegrated into a sequence of questions, wherein every question corresponds to one predicate.

3.2. Representation of Knowledge

The representation of knowledge has a long history of development in the areas of AI as well as logic. The notion of graphically representing knowledge dates back to 1956, when the idea of the semantic net was proposed by Richens in 1956. The processing of logic knowledge symbolically dates to the creation of the General Problem Solver in 1959 [35]. The foundation of knowledge was initially employed alongside knowledge-based systems for the purposes of problem-solving and reasoning. MYCIN [36] is one of the most popular rule-based expert systems currently utilized for medical diagnosis. It has a knowledge base comprising approximately 600 rules. In later years, the community of human representation of knowledge progressed to hybrid representations, rule-based and frame-based language. Web Ontology Language (OWL) as well as a resource description framework (RDF) were subsequently developed, and came to serve as significant Semantic Web standards. At that time, several open bases of knowledge or ontologies were established; for example, Freebase, YAGO, DBpedia, and WordNet. The authors of [37] presented a current concept for knowledge structures in a graph. Subsequently, the knowledge graph (KG) notion

became popularized by Google in 2012, when a framework of knowledge fusion termed Knowledge Vault [38] was provided to develop large-scale KGs.

Integration of human knowledge is a current avenue of AI study directions. Knowledge representation and reasoning, simulated from data regarding human problem solving, is a knowledge priority for intelligent systems responsible for performing sophisticated tasks [36]. Recently, as a form of structured human knowledge, KGs have garnered considerable attention from both industry and academia. A KG delivers a structured mode of fact representation, whereby facts can either be annotated manually by human experts or automatically removed from huge data structures and plaintexts. KGs are essentially composed of semantic descriptions, relationships, and entities. These entities can be abstract concepts and real-life objects, while relationships embody the association between descriptions of semantic entities, and their relationships include properties and categories with a distinct meaning. These attributed graphs or property graphs are extensively employed in instances where relations and nodes have attributes or features. For example, KGs play a significant role in several applications and tasks associated with artificial intelligence such as question answering, extraction of information, document ranking, indexing of topics, semantic parsing, disambiguation of entity, word sense disambiguation, and computation of word similarity [39].

With the advances in deep learning, distributed representation learning has revealed its capabilities in the contexts of NLP and computer vision. Currently, KGs for distributed representation learning are being investigated, for their capability to represent knowledge in terms of knowledge inference, relation extraction, and other knowledge-driven applications including QA and IR. Such applications are believed to offer depth and accuracy in terms of understanding the requirements of users, and so provide appropriate responses. However, to work well they require specific external knowledge. There remain some gaps in the knowledge retained in KGs, and the knowledge employed in knowledge-driven applications. To address this, researchers have used KG representations to close the gap. Moreover, KG representations can be utilized to resolve the sparsity of data when modelling the relatedness that exists between relations and entities [39]. Moreover, they are suitable for inclusion in deep learning approaches, as they by nature possess potential, in connection with diverse information.

In the context of question answering, KGs can provide answers to questions posed in NL. Neural network-based methods represent answers and questions in dispersed semantic space, and some also perform a symbolic injection of knowledge for prevalent sense reasoning. Taking a KG as a peripheral intellectual source, simple factoid QA can answer a simple question that involves a single fact present on the KG. The authors of [40] offered a restricted focused neural network provided with focused pruning, in order to decrease the space allocated to searching. Recently, the study in [41] has modelled two-way contact between KG questions and a bidirectional attention device. Despite being extensively used in KG-QA, deep learning techniques increase the complexity of models. According to [42], through the assessment of simple KG-QA with and without neural networks, complex deep models like gated recurrent unit (GRU) and long short-term memory (LSTM) with heuristics attain state of the art standards, and non-neural models perform reasonably well.

3.3. Ontology

A theoretical representation of notions and their associations within a particular area are known as ontology. As [4] clarified, ontology can be regarded as a database. Nevertheless, it has a more complicated KB. In order to perform enquiries to retrieve knowledge from ontology, structured languages such as SPARQL have been developed. Domain ontology describes how knowledge is associated with linguistic structures, such as the lexicon and grammar. It can be employed to resolve the obscurity of NLQs, and also to develop the accuracy and expressiveness of queries. Ontology has a significant role to play in developing the accuracy and effectiveness of QA systems [43]. Moreover, ontologies

characterize and describe domain knowledge clearly in an auto-readable structure, and thus can be incorporated into machine-based formations designed to accelerate assessment, aid the reflexive annotation of net resources, and perform logical work [44]. Reference [45] suggested a novel question analysis process that employed ontologies, signifying a standard structure of queries by applying type-recognized graphs, and combining domain the ontologies and a lexicon-syntactic structure for query reformulation.

Many previous studies have revealed the effectiveness of semantic search engines over conventional keyword search engines when dealing with NLQ in several domains, among which are the biomedical field [46], expert systems [47], transportation [48], and forensics [49]. Meanwhile, reference [46] offered an automatic generation framework for novel ontology, called the linked open data (LOD) method for automatic biomedical ontology generation (LOD-ABOG), which is powered by LOD. The LOD-ABOG was utilized with the aim of extracting concepts utilizing KBs, mainly the medical system of language UMLS and LOD, together with the operation of NLP. Reference [47] investigated the modelling of a mobile object trajectory in the semantic context web. Their study presented a form of semantic trajectory episode (STEP) ontology to represent the generic spatiotemporal episode, using a FrameSTEP as a fresh framework to explain semantic trajectories, on the basis of episodes. Meanwhile, reference [48] proposed a novel method for intelligent transportation systems (ITSs), using novel fuzzy ontology-based semantic knowledge with a Word2vec model to extract transportation features and classify texts, utilizing a bi-directional long short-term memory (Bi-LSTM) method. The method supported semantic knowledge description in terms of features as well entities and their association within the transportation area. Elsewhere, reference [49] attempted to utilize ontology for classifying a digital forensic area based on several dimensions, utilizing an RDF and an OWL which was searchable via SPARQL. The ontological design could incorporate current ontologies. The semantic web was employed to carry out reasoning and assist digital investigators in selecting a suitable instrument.

A comparative study was conducted by [50], in which keyword search engines, e.g., Yahoo and Google, were tested against semantic search engines, e.g., Hakia and DuckDuckGo. Their study revealed the semantic searches provided more relevant answers. In their research, reference [51] developed a generated framework of automated ontology, on the basis of N-Gram algorithms, classification, and semantic medical oncology, to enhance relevant biomedical concepts and identification of terms. The same author also proposed another framework for automated biomedical ontology generation utilizing LOD to reduce the requirement for ontology engineers and domain experts. However, reference [52] presented an analysis method as well as ontology-based rollover monitoring, which improved the decision support system for engineering vehicles. This model was developed to represent stability data for monitored rollover with semantic features, and to build semantic relevance across several notions present in the rollover domain. The querying and reasoning of the ontology approach were grounded in SPARQL, as well as RDF query language, and the rules of semantic web rule language (SWRL) were employed for rollover evaluation and to identify proposed measures. Concomitantly, reference [53] presented a framework permitting a large amount of data to be accessed over the internet. An ontology that can hold web resource description was established, and a composition engine of automated servicing was conducted to establish the benefits of the ontology. The engine had two execution stages, namely, the processing stage and the start-up stage.

In conclusion, due to the increasing number of mobile and fixed sensors engaged in producing the large amounts of geospatial data responsible for altering human lifestyles, and influencing human and ecological health, an ontology-driven method to incorporate intelligence for handling human health as well as ecological health risk in the geospatial sensor web (GSW) was presented by [27]. The study described a web-based model, the Human and Ecological Health Risk Management System, formulated to assist human experts and assess human ecological risk.

3.4. Review of the Most Existing Methodologies

This section reviews the methodologies used for ontology development in QA systems, namely the set of activities that implement the basic principles of ontology design. While there is a growing requirement for an ontology based on the existing or new methodologies [54], any future development will involve the inherent challenges associated with standard generic methods of building knowledge representation, due to the inadequacies present in the guiding principles of software development. This study reviewed various ontology construction methodologies for different domains.

The goal of the study conducted by [55] was to understand the application of ontology in different domains, and to ascertain the methodologies that apply to each specific domain. Significantly, the study captured the steps involved in ontology construction, namely ontology scope, ontology capture, ontology encode, ontology integration, ontology evaluation, and ontology documentation, together with the respective merits and failings of each of the methodologies. Meanwhile, reference [16] proposed an experimental environment for the semantic web, with custom rules and scalable datasets. The experiment was evaluated according to the performance of an optimized chaining ontology reasoning system, and a comparison was made of the experimental results and other ontology reasoning systems, to assess the performance and scalability of the ontology reasoning system.

In a different study, reference [56] presented an open-source math-aware QA system, based on Ask Platypus, which returned a single mathematical formula for an NLQ in the English language or Hindi. The formulae originated from the knowledge-based Wikidata and were translated into computable data by integrating the calculation engine into the system. Approximately 80% of the results were correct and significant. Although QA using an ontology-based methodology has recently received significant interest, current methods are incompatible with Arabic. One of the first NL QA interfaces was proposed in the study by [57], which recommended a system called AR2SPARQL that interpreted Arabic questions paired against RDF data in triplicate to answer the given question. In order to address the imperfect support for Arabic NLP, rather than making exhaustive use of new semantic techniques, the system employed an expertise-based ontology to secure the framework for Arabic questions, and create a relevant RDF illustration. The technique attained a 61% average recall, and 88.14% average precision [57]. Similarly, reference [58] proposed a simple Natural Language Interface (NLI) for Arabic ontologies that transformed user queries expressed in Arabic into formal language queries. The keywords were extracted using NLP techniques, then mapped to the ontology entities, and a valid SPARQL Protocol and RDF Query Language (SPARQL) query generated based on the ontology definition and the reasoning capabilities of the ontology web language (OWL). Meanwhile, reference [59] used a framework to perform a concept and keyword-based English search of the Holy Quran. It employed the Quran English WordNet, a WordNet driven database for the English translation of the Quran, to form a semantically rich Quran search system.

A university system ontology-based IR method and reasoning was developed by [60], who aimed to identify a general framework for ontology development. A general framework for the ontology searching mechanism was proposed using the university ontology. The initial step in the ontology development required information gathering, based on the targeted domain (the university), involving the identification and definition of all super and subclasses of the targeted domain. This was followed by the identification and definition of properties, such as object, data, and annotations, that existed among the classes defined. The interaction among these classes was crucial; therefore, the identification and application of the appropriate restriction constraints to the classes were essential to conduct proper reasoning. The various stages were then synchronized to identify design inconsistencies. The methodology was implemented using Protégé 4.3.1, and the reasoner (Fact ++ and Hermet 1.3.8) was used to check consistency and reliability. At the end of the implementation, the ontology developed was exported to the desired location. The study also compared the various ontologies discussed in the previous literature at that

time. While the ontology developed by the study was intended for specific segments of the university, such as the examination system, course management system, scientific research management, and data warehouse, the authors also amalgamated all aspects of the university ontology previously captured by the different studies into a single ontology to assist university staff in addressing various concerns. The subsequent ontology was able to assist users in obtaining all forms of reasoning support and could be reused by altering the restriction parameters [60]. The framework proposed facilitated efficient and effective search of and access to information through the description logic DL query.

A recent study by [61] sought to develop a conceptual model for a QA system based on the semantic web. The study combined centralized and distributed data, choosing the former due to its ability to speed up the data retrieval process, and the latter because it can retain its original format when accessed by search engines and QA systems. The study's test case domain was the automotive QA system (AQAS), and it formulated a model that facilitated the features available in the system. The main functions of AQAS were the collection of data in the data center, and obtaining answers based on the user's questions. The AQAS data knowledge employed the OWL/XML or RDF/Turtle form, and the data used were from the automotive manufacturers, dealers, and rentals on the internet. The AQAS conceptual model contained the following three main elements. Firstly, users who posed questions to the AQAS, using the Indonesian language expected that the AQAS would respond to all questions asked. Secondly, the system task, namely the task of the AQAS, which was the ability to understand the question, search related information, and provide an appropriate response. Finally, the distributed data were spread across the internet to provide reference answers to questions posed. The KB of the AQAS was divided into the following four groups:

a. *OntoKendaraan*: This was based on the OWL DL standard, and was the main ontology used to develop the other ontologies; b. *OntoManufacturer*: This was an ontology and collection of motorized vehicle instances provided by the manufacturer; c. *OntoDealers*: This was a file containing a collection of instances provided by motor vehicle sales service providers (dealers); d. *OntoRentals*: This was a file containing a collection of motor vehicle rental instances.

The novelty of the study was its focus on resolving the weaknesses associated with both distributed and centralized data, and the results demonstrated that it is possible to combine both forms of data in a pool, and then to query the pool periodically. The study contributed to the development of a conceptual model for the Indonesian QA system, specifically in the automotive domain, although it was limited by the fact that some of the data were transactional and required regular updating [61].

Recently, LOD has become a large KB, and interest in querying these data using QA techniques has attracted many researchers. Linked data possesses two known challenges: the lexical and semantic gaps [62]. The lexical gap is primarily the difference between the vocabularies used in an input question and those used in the KB, while the semantic gap is the difference between the KB representation and the information that needs expressing. The study conducted by [62] provided a novel method for querying these data, using ontology lexicon and dependency parse trees to overcome the lexical and semantic gaps. The system architecture component was constituted of six components. The syntactical analysis employed the NL as input and generated part of speech (POS) tags and dependency parse trees for the input. The second component was the named entity (NE), whose recognition aided the reference to NEs in the question and must be recognized and linked to KB resources. The template generation of the system architecture consisted of two parts: question classification and pseudo-SPARQL generation. The question classification classified questions into four categories: imperatives, such as list queries; predictive, such as 'wh-' questions; counting, such as 'how many-'; and affirmation, namely 'yes/no'. The classification was achieved using a certain classification rule. The pseudo-SPARQL generation queries were generated using a template generation algorithm, which helped to check the POS tags of the root word in the dependency parse tree generated from the

syntactical analysis phase. At the point of candidate selection, pseudo-SPARQL queries were converted into executable SPARQL queries, and all the entities mentioned in the pseudo-SPARQL query were required to replace the corresponding resources from the KB. In terms of the ranking component of the architecture proposed, the authors computed the score of the SPARQL query as the mean score for all the triples in the query. The triple score was computed as the product of the scores of its subjects, properties, and objects. Finally, the ranked SPARQL queries over the SPARQL endpoint of the KB were executed to identify the query with the highest score, and a non-empty result. The evaluation method used in this study was a recall, precision, and f-measure. The study was limited by the fact that the approach proposed could not answer questions for which there was no correct template, and also by the fact that the approach depended solely on an ontology lexicon.

One of the limitations of NLP for clinical applications is that concepts can be referenced in multiple forms across different texts. To surmount this challenge, a multi-ontology refined embedding (MORE) was proposed by [8]. This facilitated more accurate clustering of concepts across a wide range of applications, such as analysing patient health records to identify subjects with similar pathologies, to improve interoperability between hospitals. This was a hybrid approach framework for incorporating domain knowledge from many ontologies into a distributional semantic model that learned from a corpus of clinical text. Several methods were used in the course of the research. First, RadCore and MIMIC-III corpora were used to train the corpus-based component of the MORE, while the ontology aspect of the study employed the medical subject headings (MeSH) ontology and three state-of-the-art ontology-based similarity measures. The study also proposed a new learning objective, modified from the sigmoid cross-entropy objective function. The results were evaluated according to the quality of the word embedding generated, using two established datasets of semantic similarities among biomedical concepts pairs. The first dataset consisted of 29 concept pairs, with the physician establishing the similarity score and medical coders. The MORE's similarity score produced the highest combined correlation (0.633), which was 5.0% higher than that of the baseline model, and 12.4% higher than that of the best ontology-based similarity measure. The second dataset had 449 concept pairs, and the MORE's similarity score had a correlation score of 0.481, with a similarity rating of four medical residents, outperforming the skip-gram model by 8.1%, and the best ontology measure by 6.9%.

The availability of the semantic web has made structured data more accessible in the form of a KB. The goal of the QA system is to provide end-users with access to a large amount of data. As [63] explained, most current QA systems conduct KB queries using a single language (English), and current QA approaches are not designed to adapt easily to new KBs and languages. They therefore proposed a new approach that helped to translate NLQs to SPARQL queries, using a multilingual KB-agnostic approach for QA that was based on the fact that questions can be understood from the semantic context of the words in the question, and that the syntax components of the question are less important. In the study, the KB stored the semantic of the words in the question and provided the most likely interpretation of the question. The multilingual KB-agnostic approach was divided into four steps, the first of which was the question expansion step, involving the identification of all the name entities, properties, and classes to which the question referred. The second step was query construction, which involved the construction of a set of queries that represented a possible interpretation of a specific question with a given KB. At the end of the query construction step, all possible SPARQL queries were computed. The third step was query ranking, in which the query candidates computed were ordered according to their probability of answering the question correctly. Finally, at the end of the third step, a list of ranked SPARQL query candidates was computed. The candidate answer was computed if the SPARQL query was above a certain threshold. One of the limitations of the approach was that the identification of the relationships in the question relied heavily on the dictionary [63].

3.5. Current Mobile QA Systems for Pilgrimage Domain

Countless internet queries are made to obtain accurate information about the pilgrimage. However, the results of these searches produce millions of web addresses, and at times convoluted search results. According to [64], the emergence of web platforms with a vast pool of information web pages that is growing daily has transformed the ways in which users interact with the internet. Pilgrims are particularly reliant on news that is accessible on the web and have to use the internet to seek answers to related queries. Current internet queries require precise keywords to be identical or identically arranged [65,66], and this often affects the outcomes of individuals' searches for critical information. Moreover, in certain circumstances it can produce irrelevant search outcomes [67,68], causing misperceptions and confusion in the QA for related terms. There is currently a need to develop an approach that will semantically improve the quality of queries and ensure the provision of precise answers to queries by pilgrims in a mobile QA system.

Growth in the use of mobile phones while travelling has enabled the development of widespread mobile applications in all areas of life, including pilgrimage. Although the information that can be transmitted to and browsed for on such devices is limited [69], there are constraints on the kinds of questions asked by users of mobile devices regarding their pilgrimage. This is because, both appropriate websites and the types of questions that can be asked in order to obtain concise and precise answers are limited. Some efforts have been made to support mobile technologies and applications for pilgrims, such as in the Islamic domain, which is one of the main themes of ontology development. The semantic Quran [70] designed a multilingual RDF framework for the Quran structure with the associated Quran ontology, using aspects obtained from the Quran, such as places, living establishments, and events. Meanwhile, other studies addressed the interpretation of DBpedia to Arabic [71], and an ontology study explored the Mooney GeoQuery dataset, which contains documents about geographical locations in the United States (US) [72].

In their study, reference [73] proposed a Knowledge-Based Expert Systems (KBES) that can be used by pilgrims at different stages both before, during and after their pilgrimage. The main goal of their system was to support pilgrims in their learning and decision-making processes. Meanwhile, reference [74] presented a semantic system based on a QA system that enabled pilgrims to compose questions about the Umrah in an NL format. The system used ontology to represent the knowledge base of the entire ritual acts of the Umrah only. The various complexities of NL were observable in the system because the questions had to be mapped with the contents in the ontology. Among the more notable systems currently available are the Umrah and Educational modules. The Umrah module covers all the rules and facts in the ontology, and the Educational module is responsible for question and answer interactions.

Similarly, reference [75] proposed a dynamic knowledge-based approach that facilitated a decision support system used by pilgrims to query information from dynamic knowledge-based records. The system also enabled pilgrims to contact experts, in order to obtain suitable possible solutions to their queries. Meanwhile, the application known as M-Hajj DSS provided simple and advanced questions and answers using artificial intelligence (AI), based on decision tree and case-based reasoning (CBR). However, the system was only developed for Android smartphones, employed only a single language, and also failed to consider the universal user interface and mobile information quality. The issue of language was addressed by the authors of [76], who presented a mobile translation application to enable pilgrims to accomplish rituals that included different languages, and sought to assist pilgrims in connecting easily with other pilgrims who speak different languages.

Finally, the analytical study conducted by [77] reported that the majority of current pilgrimage applications' developers focused on basic usage, and their applications lacked interactive features, with users directed to visual applications, more than 87% of which support only one language. The study revealed certain weaknesses in the presentation of interactive features that support virtual communication between users, such as fatwa chatting, and found that existing pilgrimage mobile applications lack real-time map updates

to assist pilgrims, i.e., location awareness. Moreover, the study reported that no extant research has considered the use of ontology and knowledge-based approaches to resolve lexical–semantic ambiguity and enhance the accuracy of QA results in the domain.

4. The Challenges of the QA System

According to [8,78], extracting accurate information in response to a natural language question (NLQ) remains a challenge for QA systems. This is primarily because the existing techniques used by the system cannot address the effectiveness and efficiency of semantic issues in their NLQ analysis. Some scholars [11,79,80] reported positive outcomes to lexical solutions on the topic of word sense disambiguation (WSD), and several studies [42,62] appraised this issue in the context of QA, but only a few examined the role of WSD when returning potential answers [12,81].

A systematic review of NLP and text mining, conducted by [82], found that computational challenges, and synthesizing the literature concerning the use of NLP and text mining, present major challenges for QA. Meanwhile, a study conducted by [83] suggested that when real-world data are constructed and inserted into KGs, it provides an effective and convenient querying technique for end-users, a matter that requires addressing urgently. However, a greater challenge is the ability to understand the question clearly, in order to translate the unstructured question into a structured query in semantic representation form. This is due to the ambiguity associated with defining the precise word sense, which results in mismatches when mapping the expressions extracted from NLQ to the ontology elements available for the data.

Alternative systems discussed take into account QA as a semantic parsing issue and aim to train a semantic parser to plan input questions in suitable logical forms to be performed over a base of knowledge. These systems require identification of important data when training, which is time-consuming and labour intensive [62]. The QA systems based on IR engines require that text wording comprising an answer should be identical to the wording employed in the question [78].

Previous studies have revealed that existing QA systems concentrate on providing answering to factoid questions. As various sources of knowledge have advantages and disadvantages, a sensible direction to take would be one that exploits several kinds of underlying knowledge, so as to minimize the restrictions placed on the individual. It is indisputable that the principal quantity of factual information available on the web currently is enclosed by plain text (e.g., blogs, sites, documents, etc.). On the basis of the above, a direction that is worthy of further investigation would involve manipulating peripheral plain-text knowledge to enrich the content as well as the vocabulary of the KB (to develop tasks such as disambiguation, recalling the challenge of ambiguity). Moreover, pursuing an alternative direction is exciting, involving a search for more effective and flexible approaches for interlinking and structuring plain text information with organized representations of knowledge such as RDF graphs, which attempt to resolve the difficulties associated with distributed knowledge. As well as these major challenges, the processing of complex questions and multilingualism are exceedingly challenging issues, due to the multitude of languages, each of which possesses different structures and idioms [84]. In conclusion, assessment datasets and benchmarks are fundamental for attaining similar results reliably, which is essential when making comparisons of various QA systems' performance and the effectiveness of approaches. Thus, the assessment process represents another interesting research direction.

5. QA System Evaluation

Evaluation of the QA system method is a crucial component of QA systems, and as QA techniques are rapidly designed, a trustworthy evaluation measurement to review these applications is required. In the study by [14], the assessment measurement applied to QA systems was accuracy, and the harmonic mean function of precision and recall (F1 score). As shown in Formulas (1)–(3), a 2×2 possibility table enabled appreciation of these

metrics, which classified evaluation under two aspects: a) true positive (TP), denoting a section that was correctly selected, or false negative (FN), denoting a section that was correctly not chosen; and b) true negative (TN), denoting a section that was incorrectly not chosen, or false positive (FP), denoting a section that was not correctly selected. Precision was used to describe the measurement of the chosen review papers that were correct, while recall referred to the opposite gauge, namely the measurement of the correct papers chosen. Using precision and recall, the fact that there was a high rate of true negative was no longer important [14].

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

The evaluation of precision and recall involves understanding that they are contrary, and indeed there is a trade-off involved in all such measures, which means researchers must seek the best measure to calculate their particular system. Several QA systems have used recall as a gauge metric, since it does not require the evaluation of how false the positive rates are developed; if there is a high true positive rate, then the result will be improved. However, it can be argued that precision would be a more effective measure. To balance the trade-off, [14,43] introduced the 'f' calculation (Formula (4)).

$$F1 = \frac{2(Precision \cdot Recall)}{Precision + Recall} \quad (4)$$

This metric applied a weighted score, gauging precision and recall trade-off. According to [14,43,85], the following metrics (Formulas (5) and (6)) can be used to calculate QA systems, and include calculations such as the mean average precision (MAP), which is a universal assessment for IR, describing the (AveP), average precision, given by Formula (5). Meanwhile, the mean reciprocal rank (MRR) (Formula (7)) can be used to calculate relevance.

$$MAP = \sum_{Q=1}^{P=1} \frac{AveP(p)}{Q} \quad (5)$$

where AveP (Average precision) is given by the Equation (5). Mean Reciprocal Rank (MRR) is shown below, and it is used to calculate relevance.

$$AveP = \sum_{N=1}^{k=1} \frac{P(k) X_{rel}(q)}{relevant} \quad (6)$$

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

As shown in Table 2, the evaluation conducted for the present study produced a total of 83 review papers, of which 37 were relevant, ontology-based studies; thus, an average recall of 44.7% was achieved, and an average precision of 53.5%. It was observed that ontology-based methods/semantic web/link data yielded more precision than the QA, NLP, and other categories. This could be explained by the fact that ontology-based methods/semantic web/link data define several possible means of identifying entities, which helps when classifying the keywords for ontology entities. Surprisingly, the QA NLP returned a higher precision rate of 93.7%, while recall was 32.9%. This concurred with the claims made by [57], although, since their data were not publicly available it was not possible to compare the performance of the two systems. The results of the present study

employed the same method as [57], which is an approach used for knowledge identified in the ontology to manage the user’s question, and neither approach made rigorous use of complex linguistic and semantics methods. However, some elements were unique to the present study, which for example used estimate pairing and a user interface for entity mapping. This system required the participation of the user to refine their query and the use of a set of rules to certify and improve the SPARQL query.

Table 2. Evaluation results.

Category	Ontology-Based/Semantic Web/Link Data	QA NLP/ Others	Total Relevant
Total Number of Review Papers	37	46	83
Review papers Ontology QA	52	18	70
Review papers with Correct Ontology QA	23	17	40
Precision	41.2 %	93.6%	53.5%
Recall	59.1 %	32.9%	44.7%

6. Discussion

Reference [84] conducted a survey of the stateless QA system, with an emphasis on the methods applied in RDF and linked data, documents, and a combination of these. The author reviewed 21 recent systems and the 23 evaluation and training datasets most commonly used in the extant literature, according to domain-type. The study identified underlying knowledge sources, the tasks provided, and the associated evaluation metrics, identifying various contexts in terms of their QA system applications. Overall, QA systems were found to be widely used in different fields of human engagement, and their application spanned the fields of medicine, culture, web communities, and tourism, among others. The study also contextualized the current most efficient QA system, IBM Watson, which is a QA system that is not restricted to a specific domain, and which seeks to capture a wide range of questions. The IBM Watson competed against the former winner in the television game Jeopardy and won. However, despite the number of successes associated with the system, it possesses certain shortcomings and ambiguities.

There has recently been an increase in the adoption of QA-based personal assistants, including Amazon’s Alexa, Microsoft’s Cortona, Apple’s Siri, Samsung’s Bixby, and Google Assistant, which are capable of answering a wide range of questions [84]. Furthermore, there has been an increase in the application of QA systems at call centers and utilized in Internet Of Things (IoT) applications.

Here we discuss the classification of ontology-based approaches to QA, and present the categories created to classify the 83 papers reviewed by this study, which are associated with ontology methods that semantically enhance the QA. In order to extract and classify the data, the study extracted data at the full paper reading stage, creating five categories to classify the reviewed papers (Table 1). In total, of the 83 papers reviewed, only 42 mentioned ontology methods in QA directly, while 41 discussed QA NLP among other methods (Table 3); of the 83 papers reviewed, 42 (50%) employed an ontology-based paradigm, 32 (38%) used a QA NLP paradigm, and 9 (11%) employed a hybrid approach. There was a higher use of web-based and NLP approaches than hybrid paradigms (Figure 2), which concurred with the findings of [13].

Rich structured knowledge could be helpful for AI applications. Nonetheless, the way of integrating symbolic knowledge into a computational framework of practical applications remains a challenging task. Current advances in knowledge-graph-based studies concentrate on knowledge graph embedding or knowledge representation learning, which involves mapping relations as well as elements into low-dimensional vectors, while identifying their senses [86]. Knowledge-aware models obtain advantages from incorporating semantics, ontologies for the representation of knowledge, diverse data sources and multilingual knowledge. Therefore, some certifiable applications, like QA system have prospered, benefiting from their capacity for rational reasoning as well as perception. Some

genuine products, for instance, Google’s Knowledge Graph and Microsoft’s Satori have demonstrated a solid ability to offer more productive types of assistance.

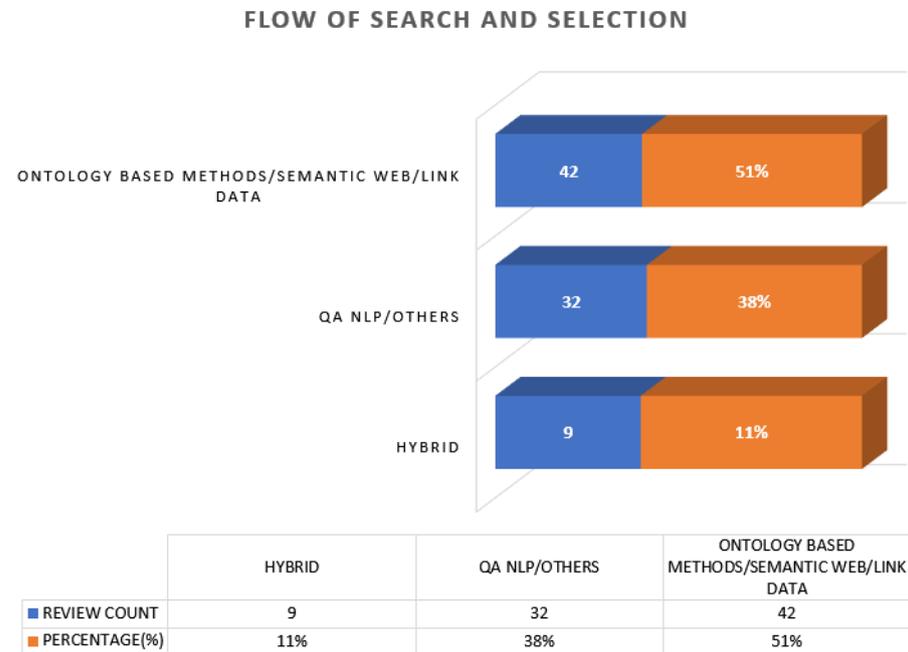


Figure 2. Flow of Search and Selection.

Table 3. Methods and techniques identified for the related literature.

Methods	Techniques	Technique for Evaluation	Ref.
Querying Agent	RDF, OWL, Spatiotemporal Information representations	Two evaluation stages: a. Air Quality Index calculation. b. Manual verification by health experts	[27,47]
Single, Multiple, Hybrid Ontology	Semantic information description. Design, Development, Integration, Validation, Ontology Iteration	No evaluation available	[55]
Semantic Search System	KB preparation, Query processing, Entity mapping, Formal query and Answer generation	Disambiguation evaluation: System achieved 103 correct answers out of 133 questions. 64% recall, 76% precision	[58]
Ontology. Decision Support system	Knowledge querying, Reasoning based on ontology model	Five sets of rollover stability data under different conditions used for evaluation: accuracy, effectiveness	[52]
RDF, SPARQL	NLP. Ontology entity mapping, Retrieve and Manipulate data in RDF	Quranic ontology, Arabic Mooney Geography dataset for evaluation. System achieved 64% recall, 76% precision	[52,58]
Semantic QA	Ontology reasoning, custom rules, Semantic QA system	Custom rules and query set for system evaluation. Result: Backward chaining ontology outperforms the in-memory reasoning system	[16]
IR	Quran’s concept hierarchy, vocabulary search system, Quranic WordNet, Knowledge repository, IR tools	Performance metrics evaluation: precision, recall, F-measure. Comparisons with similar frameworks	[56,59]
FrameSTEP	Raw trajectories, Annotation, Semantic graphs, Ontology	Segmentation Granularity Extent Context Type and availability	[27,47]
Ontology, Semantic Knowledge, Word2vec	OWL, Word embedding, fuzzy ontology	Bi-LSTM improved features extraction and text classification. Evaluation based on machine learning: SVM, CNN, RNN, LSTM. Metrics: Precision, Recall, Accuracy	[48]
Intelligent Mobile Agent	Ontology, DBPedia, WordNet.	IMAT was validated by Mobile Client Application implementation it helps testing of important IMAT features	[52,53,58,59]

Table 3. Cont.

Methods	Techniques	Technique for Evaluation	Ref.
IR	Math-aware QA system, Ask Platypus. A single mathematical formula is returned to NLQ in English or Hindi	Metrics: Precision, Recall, Accuracy, Function Measures	[56,59]
Linked Open data Framework	Concepts and Relation extraction, NLP, KB	Evaluation shows improved result in most tasks of ontology generation compared to those obtained in existing frameworks	[46]
IR	RDF, OWL, SPARQL	No evaluation available	[49,56,59]
Classification and Mining	NLP, Linguistic features,	Neural Networks capable of handling complexity, classified Hadith 94% accuracy. Mining method: Vector space model, Cosine similarity, Enriched queries—Obtained 88.9% accuracy	[64]
Semantic Search	IR, Quran-based QA system, Neural Networks classification	Evaluation of classification shows approximately 90% accuracy	[87]

With the ongoing progress in deep learning, and the inescapable utilization of distributional semantics to create word embedding for representation of words through vectors in neural networks, such corpus-based models have acquired tremendous prominence. One of the most widely recognized distributional semantics model is word2vec [8] for creating word embedding. The word2vec model is a neural network that is capable to enhance NLP with semantic and syntactic relationships between word vectors. While corpus-based measures have shown to be more adaptable and extensible than ontology-based measures, human curation can develop their precision further in domain applications.

Previous endeavors sought to integrate corpus-based and ontology-based similarity to well identify semantic similarities; in any case, no framework is presently available in the domain of pilgrimage delivering ontological information using a method creating word embedding for semantic similarity. Although the use of ontology for knowledge representation has captured the attention of researchers and developers as a means of overcoming the limitations of the keyword-based search, there has been little focus on the Islamic domain of knowledge, and specifically on QA for pilgrimage rituals. Those studies that do exist in this field have unique limitations, including the fact that existing religious content websites work on either language-dependent searches or keyword-based searches; the presence of only theoretical and conceptual notions of ontological retrieving agents [1]; and poor system usability, especially at the query level, that requires the user to manage complex language or interfaces to express their needs [88].

The present study addressed the research questions posed, and the ontology-based approach to QA systems identified is illustrated in Table 3, which shows that the majority of the extant studies in the field targeted a specific domain. The features unique to the majority of the works were their initial understanding of the domain (domain knowledge); the component that facilitated ontological reasoning (query processing); the mapping of the objects of interest in the domain (entities); the generation of appropriate responses, based on the query issued; and, finally, the acknowledgement that Hajj/Umrah activity is an all-year-round activity. In terms of the tasks to be performed, some are particular only to the Hajj, and in terms of sites to be visited, some are relevant to the Hajj, and others to both the Hajj and the Umrah. Several tasks occur across other domains, such as tourism, where activities and sites to be visited are essential factors. Therefore, the defined methodology with minimal modification of fuzzy rules can be implemented in such a domain.

7. Conclusions and Future Works

This study conducted a review of ontology-based approaches to semantically enhance QA for a closed domain. It found that 83 out of the 124 papers reviewed described QA approaches that had been developed and evaluated using different methods. The ontological approach to semantically enhancing QA was not found to be widely embraced,

as many studies featured NLP and IR processing. Most of the studies reviewed focused on open domain usage, but this study concerns a closed domain. The precision and recall measures were primarily employed to evaluate the method used by the studies.

Due to the unique attributes of the pilgrimage exercise, the methodology best suited to the development of a pilgrimage ontology is hybrid, the first component being the creation of fuzzy rules to analyze NLQs in order to define the exercise a pilgrim is performing at a specific location and time of year. The rules should be applied to the categories of words in the question, which will ultimately assist in generating some semantic structure for the text contained in the question. The second component is the semantic analysis, a component of the hybrid methods that will produce an understanding of the interpretation of words from the question. The generation of an ontology comprising these components, populated with dynamic facts using a KG, will help to answer questions specifically related to the pilgrimage domain.

We plan to evaluate the advancement of these segments within an extensible framework, by integrating a knowledge-based and distributional semantic model for learning word embedding from context and domain information, by using measures for semantic similarities on top of the Word2Vec algorithm. This is due to the algorithm's capability to encode high order relationships and cover a wider range of connections between data points. It is a significant factor in enhancing several tasks, such as semantic parsing, disambiguation of an entity that results in accuracy improvement in QA systems. This requires the development of new methods that suit a collaboration between these components to enhance data processing in QA systems. Finally, we aim to develop a mobile application prototype to evaluate effectiveness and address users' perceptions of the mobile QA application.

Author Contributions: Conceptualization, A.A. and A.S.; methodology, A.A.; software, A.A.; validation, A.A. and A.S.; formal analysis, A.A.; investigation, A.A.; resources, A.A. and A.S.; data curation, A.A.; writing—original draft preparation, A.A.; writing—review and editing, A.A. and A.S.; visualization, A.A.; supervision, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Authors can confirm that all relevant data are included in the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Khan, J.R.; Siddiqui, F.A.; Siddiqui, A.A.; Saeed, M.; Touheed, N. Enhanced ontological model for the Islamic Jurisprudence system. In Proceedings of the 2017 International Conference on Information and Communication Technologies (ICICT), Karachi, Pakistan, 30–31 December 2017; pp. 180–184.
2. Siciliani, L. Question Answering over Knowledge Bases. In *European Semantic Web Conference*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 283–293.
3. Hu, H. A Study on Question Answering System Using Integrated Retrieval Method. Ph.D. Thesis, The University of Tokushima, Tokushima, Japan, February 2006. Available online: <http://citeseerx.ist.psu.edu/viewdoc/download> (accessed on 20 April 2021).
4. Abdi, A.; Idris, N.; Ahmad, Z. QAPD: An ontology-based question answering system in the physics domain. *Soft Comput.* **2018**, *22*, 213–230. [[CrossRef](#)]
5. Buranarach, M.; Supnithi, T.; Thein, Y.M.; Ruangrajitpakorn, T.; Rattanasawad, T.; Wongpatikaseree, K.; Lim, A.O.; Tan, Y.; Assawamakin, A. OAM: An ontology application management framework for simplifying ontology-based semantic web application development. *Int. J. Softw. Eng. Knowl. Eng.* **2016**, *26*, 115–145. [[CrossRef](#)]
6. Vandenbussche, P.Y.; Atemez, G.A.; Poveda-Villalón, M.; Vatant, B. Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web. *Semant. Web* **2017**, *8*, 437–452. [[CrossRef](#)]
7. Hastings, J.; Chepelev, L.; Willighagen, E.; Adams, N.; Steinbeck, C.; Dumontier, M. The chemical information ontology: provenance and disambiguation for chemical data on the biological semantic web. *PLoS ONE* **2011**, *6*, e25513. [[CrossRef](#)]
8. Jiang, S.; Wu, W.; Tomita, N.; Gao, C.; Hassanpour, S. Multi-Ontology Refined Embeddings (MORE): A Hybrid Multi-Ontology and Corpus-based Semantic Representation for Biomedical Concepts. *arXiv* **2020**, arXiv:2004.06555.
9. Raganato, A.; Camacho-Collados, J.; Navigli, R. Word sense disambiguation: A unified evaluation framework and empirical comparison. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, Valencia, Spain, 3–7 April 2017; Volume 1, pp. 99–110.

10. Chaplot, D.S.; Salakhutdinov, R. Knowledge-based word sense disambiguation using topic models. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
11. Raganato, A.; Bovi, C.D.; Navigli, R. Neural sequence learning models for word sense disambiguation. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 7–11 September 2017; pp. 1156–1167.
12. Wang, Y.; Wang, M.; Fujita, H. Word sense disambiguation: A comprehensive knowledge exploitation framework. *Knowl. Based Syst.* **2020**, *190*, 105030. [[CrossRef](#)]
13. Soares, M.A.C.; Parreiras, F.S. A literature review on question answering techniques, paradigms and systems. *J. King Saud Univ. Comput. Inf. Sci.* **2020**, *32*, 635–646.
14. Yao, X. Feature-Driven Question Answering with Natural Language Alignment. Ph.D. Thesis, Johns Hopkins University, Baltimore, MD, USA, July 2014.
15. Mohasseb, A.; Bader-El-Den, M.; Cocea, M. Question categorization and classification using grammar based approach. *Inf. Process. Manag.* **2018**, *54*, 1228–1243. [[CrossRef](#)]
16. Shi, H.; Chong, D.; Yan, G. Evaluating an optimized backward chaining ontology reasoning system with innovative custom rules. *Inf. Discov. Deliv.* **2018**, *46*, 45–56. [[CrossRef](#)]
17. Green, B.F., Jr.; Wolf, A.K.; Chomsky, C.; Laughery, K. Baseball: An automatic question-answerer. In *Western Joint IRE-AIEE-ACM Computer Conference*; Association for Computing Machinery: New York, NY, USA, 1961; pp. 219–224.
18. Woods, W.A. Progress in natural language understanding: An application to lunar geology. In *Proceedings of the June 4–8, 1973, National Computer Conference and Exposition*; Association for Computing Machinery: New York, NY, USA, 1973; pp. 441–450.
19. Katz, B. From sentence processing to information access on the world wide web. In *AAAI Spring Symposium on Natural Language Processing for the World Wide Web*; Stanford University: Stanford, CA, USA, 1997; Volume 1, p. 997.
20. Kwok, C.C.; Etzioni, O.; Weld, D.S. Scaling question answering to the web. In Proceedings of the 10th International Conference on World Wide Web, Hong Kong, China, 1–5 May 2001; pp. 150–161.
21. Zheng, Z. AnswerBus question answering system. In Proceedings of the Human Language Technology Conference (HLT 2002), San Diego, CA, USA, 7 February 2002; Volume 27.
22. Parikh, J.; Murty, M.N. Adapting question answering techniques to the web. In Proceedings of the Language Engineering Conference, Hyderabad, India, 13–15 December 2002; pp. 163–171.
23. Bhatia, P.; Madaan, R.; Sharma, A.; Dixit, A. A comparison study of question answering systems. *J. Netw. Commun. Emerg. Technol.* **2015**, *5*, 192–198.
24. Chung, H.; Song, Y.I.; Han, K.S.; Yoon, D.S.; Lee, J.Y.; Rim, H.C.; Kim, S.H. A practical QA system in restricted domains. In Proceedings of the Conference on Question Answering in Restricted Domains, Barcelona, Spain, 25 July 2004; pp. 39–45.
25. Mishra, A.; Mishra, N.; Agrawal, A. Context-aware restricted geographical domain question answering system. In Proceedings of the 2010 International Conference on Computational Intelligence and Communication Networks, Bhopal, India, 26–28 November 2010; pp. 548–553.
26. Clark, P.; Thompson, J.; Porter, B. A knowledge-based approach to question-answering. *Proc. AAAI* **1999**, *99*, 43–51.
27. Meng, X.; Wang, F.; Xie, Y.; Song, G.; Ma, S.; Hu, S.; Bai, J.; Yang, Y. An Ontology-Driven Approach for Integrating Intelligence to Manage Human and Ecological Health Risks in the Geospatial Sensor Web. *Sensors* **2018**, *18*, 3619. [[CrossRef](#)]
28. Yang, M.C.; Lee, D.G.; Park, S.Y.; Rim, H.C. Knowledge-based question answering using the semantic embedding space. *Expert Syst. Appl.* **2015**, *42*, 9086–9104. [[CrossRef](#)]
29. Firebaugh, M.W. *Artificial Intelligence: A Knowledge Based Approach*; PWS-Kent Publishing Co: Boston, MA, USA, 1989.
30. Ziaee, A.A. A philosophical Approach to Artificial Intelligence and Islamic Values. *IJUM Eng. J.* **2011**, *12*, 73–78. [[CrossRef](#)]
31. Allen, J. *Natural Language Understanding*; Benjamin/Cummings Publishing Company: San Francisco, CA, USA, 1995.
32. Dubien, S. Question Answering Using Document Tagging and Question Classification. Ph.D. Thesis, University of Lethbridge, Lethbridge, Alberta, 2005.
33. Guo, Q.; Zhang, M. Question answering based on pervasive agent ontology and Semantic Web. *Knowl. Based Syst.* **2009**, *22*, 443–448. [[CrossRef](#)]
34. Cui, W.; Xiao, Y.; Wang, H.; Song, Y.; Hwang, S.W.; Wang, W. KBQA: Learning question answering over QA corpora and knowledge bases. *arXiv* **2019**, arXiv:1903.02419.
35. Dong, J.; Wang, Y.; Yu, R. Application of the Semantic Network Method to Sightline Compensation Analysis of the Humble Administrator’s Garden. *Nexus Netw. J.* **2021**, *23*, 187–203. [[CrossRef](#)]
36. Shortliffe, E. *Computer-Based Medical Consultations: MYCIN*; Elsevier: Amsterdam, The Netherlands, 2012; Volume 2.
37. Stokman, F.N.; de Vries, P.H. Structuring knowledge in a graph. In *Human-Computer Interaction*; Springer: Berlin/Heidelberg, Germany, 1988; pp. 186–206.
38. Dong, X.; Gabrilovich, E.; Heitz, G.; Horn, W.; Lao, N.; Murphy, K.; Strohmman, T.; Sun, S.; Zhang, W. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014; pp. 601–610.
39. Lin, Y.; Han, X.; Xie, R.; Liu, Z.; Sun, M. Knowledge representation learning: A quantitative review. *arXiv* **2018**, arXiv:1812.10901.
40. Dai, Z.; Li, L.; Xu, W. Cfo: Conditional focused neural question answering with large-scale knowledge bases. *arXiv* **2016**, arXiv:1606.01994.

41. Chen, Y.; Wu, L.; Zaki, M.J. Bidirectional attentive memory networks for question answering over knowledge bases. *arXiv* **2019**, arXiv:1903.02188.
42. Mohammed, S.; Shi, P.; Lin, J. Strong baselines for simple question answering over knowledge graphs with and without neural networks. *arXiv* **2017**, arXiv:1712.01969.
43. Zhang, Z.; Liu, G. Study of ontology-based intelligent question answering model for online learning. In Proceedings of the 2009 First International Conference on Information Science and Engineering, Nanjing, China, 26–28 December 2009; pp. 3443–3446.
44. Shadbolt, N.; Berners-Lee, T.; Hall, W. The semantic web revisited. *IEEE Intell. Syst.* **2006**, *21*, 96–101. [[CrossRef](#)]
45. Besbes, G.; Baazaoui-Zghal, H.; Moreno, A. Ontology-based question analysis method. In *International Conference on Flexible Query Answering Systems*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 100–111.
46. Alobaidi, M.; Malik, K.M.; Sabra, S. Linked open data-based framework for automatic biomedical ontology generation. *BMC Bioinform.* **2018**, *19*, 319. [[CrossRef](#)] [[PubMed](#)]
47. Nogueira, T.P.; Braga, R.B.; de Oliveira, C.T.; Martin, H. FrameSTEP: A framework for annotating semantic trajectories based on episodes. *Expert Syst. Appl.* **2018**, *92*, 533–545. [[CrossRef](#)]
48. Ali, F.; El-Sappagh, S.; Kwak, D. Fuzzy Ontology and LSTM-Based Text Mining: A Transportation Network Monitoring System for Assisting Travel. *Sensors* **2019**, *19*, 234. [[CrossRef](#)]
49. Wimmer, H.; Chen, L.; Narock, T. Ontologies and the Semantic Web for Digital Investigation Tool Selection. *J. Digit. For. Secur. Law* **2018**, *13*, 6. [[CrossRef](#)]
50. Singh, J.; Sharan, D.A. A comparative study between keyword and semantic based search engines. In Proceedings of the International Conference on Cloud, Big Data and Trust, Madhya Pradesh, India, 13–15 November 2013; pp. 13–15.
51. Alobaidi, M.; Malik, K.M.; Hussain, M. Automated ontology generation framework powered by linked biomedical ontologies for disease-drug domain. *Comput. Methods Programs Biomed.* **2018**, *165*, 117–128. [[CrossRef](#)]
52. Xu, F.; Liu, X.; Zhou, C. Developing an ontology-based rollover monitoring and decision support system for engineering vehicles. *Information* **2018**, *9*, 112. [[CrossRef](#)]
53. Dhayne, H.; Chamoun, R.K.; Sabha, R.A. IMAT: Intelligent Mobile Agent. In Proceedings of the 2018 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET), Beirut, Lebanon, 14–16 November 2018; pp. 1–8.
54. Nanda, J.; Simpson, T.W.; Kumara, S.R.; Shooter, S.B. A Methodology for Product Family Ontology Development Using Formal Concept Analysis and Web Ontology Language. *J. Comput. Inf. Sci. Eng.* **2006**, *6*, 103–113. [[CrossRef](#)]
55. Subhashini, R.; Akilandeswari, J. A survey on ontology construction methodologies. *Int. J. Enterp. Comput. Bus. Syst.* **2011**, *1*, 60–72.
56. Schubotz, M.; Scharpf, P.; Dudhat, K.; Nagar, Y.; Hamborg, F.; Gipp, B. Introducing MathQA: A Math-Aware question answering system. *Inf. Discov. Deliv.* **2018**, *46*, 214–224. [[CrossRef](#)]
57. AlAgha, I.M.; Abu-Taha, A. AR2SPARQL: An arabic natural language interface for the semantic web. *Int. J. Comput. Appl.* **2015**, *125*, 19–27. [[CrossRef](#)]
58. Hakkoum, A.; Kharrazi, H.; Raghay, S. A Portable Natural Language Interface to Arabic Ontologies. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*, 69–76. [[CrossRef](#)]
59. Afzal, H.; Mukhtar, T. Semantically enhanced concept search of the Holy Quran: Qur’anic English WordNet. *Arab. J. Sci. Eng.* **2019**, *44*, 3953–3966. [[CrossRef](#)]
60. Ullah, M.A.; Hossain, S.A. Ontology-Based Information Retrieval System for University: Methods and Reasoning. In *Emerging Technologies in Data Mining and Information Security*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 119–128.
61. Syarief, M.; Agustiono, W.; Muntasa, A.; Yusuf, M. A Conceptual Model of Indonesian Question Answering System based on Semantic Web. *J. Phys. Conf. Ser.* **2020**, *1569*, 022089.
62. Jabalameli, M.; Nematbakhsh, M.; Zaeri, A. Ontology-lexicon-based question answering over linked data. *ETRI J.* **2020**, *42*, 239–246. [[CrossRef](#)]
63. Diefenbach, D.; Both, A.; Singh, K.; Maret, P. Towards a question answering system over the semantic web. *Semant. Web* **2020**, *11*, 421–439. [[CrossRef](#)]
64. Saloot, M.A.; Idris, N.; Mahmud, R.; Ja’afar, S.; Thorleuchter, D.; Gani, A. Hadith data mining and classification: A comparative analysis. *Artif. Intell. Rev.* **2016**, *46*, 113–128. [[CrossRef](#)]
65. White, R.W.; Richardson, M.; Yih, W. Questions vs. queries in informational search tasks. In Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 18–22 May 2015; pp. 135–136.
66. del Carmen Rodriguez-Hernández, M.; Ilarri, S.; Trillo-Lado, R.; Guerra, F. Towards keyword-based pull recommendation systems. *ICEIS 2016* **2016**, *1*, 207–214.
67. Rahman, M.M.; Yeasmin, S.; Roy, C.K. Towards a context-aware IDE-based meta search engine for recommendation about programming errors and exceptions. In Proceedings of the 2014 Software Evolution Week-IEEE Conference on Software Maintenance, Reengineering, and Reverse Engineering (CSMR-WCRE), Antwerp, Belgium, 3–6 February 2014; pp. 194–203.
68. Giri, K. Role of ontology in semantic web. *Desidoc J. Libr. Inf. Technol.* **2011**, *31*, 116–120. [[CrossRef](#)]
69. Maglaveras, N.; Koutkias, V.; Chouvarda, I.; Goulis, D.; Avramides, A.; Adamidis, D.; Louridas, G.; Balas, E. Home care delivery through the mobile telecommunications platform: The Citizen Health System (CHS) perspective. *Int. J. Med. Inform.* **2002**, *68*, 99–111. [[CrossRef](#)]

70. Sherif, M.; Ngonga Ngomo, A.C. Semantic Quran: A multilingual resource for natural-language processing. *Semant. Web* **2015**, *6*, 339–345. [[CrossRef](#)]
71. Al-Feel, H. The roadmap for the Arabic chapter of DBpedia. In Proceedings of the 14th International Conference on Telecom. and Informatics (TELE-INFO'15), Sliema, Malta, 17–19 August 2015; pp. 115–125.
72. Hakkoum, A.; Raghay, S. Semantic Q&A System on the Qur'an. *Arab. J. Sci. Eng.* **2016**, *41*, 5205–5214.
73. Sulaiman, S.; Mohamed, H.; Arshad, M.R.M.; Yusof, U.K. Hajj-QAES: A knowledge-based expert system to support hajj pilgrims in decision making. In Proceedings of the 2009 International Conference on Computer Technology and Development, Kota Kinabalu, Malaysia, 13–15 November 2009; Volume 1, pp. 442–446.
74. Sharef, N.M.; Murad, M.A.; Mustapha, A.; Shishechi, S. Semantic question answering of umrah pilgrims to enable self-guided education. In Proceedings of the 2013 13th International Conference on Intelligent Systems Design and Applications, Salangor, Malaysia, 8–10 December 2013; pp. 141–146.
75. Mohamed, H.H.; Arshad, M.R.H.M.; Azmi, M.D. M-HAJJ DSS: A mobile decision support system for Hajj pilgrims. In Proceedings of the 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 15–17 August 2016; pp. 132–136.
76. Abdelazeez, M.A.; Shaout, A. Pilgrim Communication Using Mobile Phones. *J. Image Graph.* **2016**, *4*, 59–62. [[CrossRef](#)]
77. Khan, E.A.; Shambour, M.K.Y. An analytical study of mobile applications for Hajj and Umrah services. *Appl. Comput. Inform.* **2018**, *14*, 37–47. [[CrossRef](#)]
78. Rodrigo, A.; Penas, A. A study about the future evaluation of Question-Answering systems. *Knowl. Based Syst.* **2017**, *137*, 83–93. [[CrossRef](#)]
79. Al-Harbi, O.; Jusoh, S.; Norwawi, N.M. Lexical disambiguation in natural language questions (NLQs). *arXiv* **2017**, arXiv:1709.09250.
80. Navigli, R. Natural Language Understanding: Instructions for (Present and Future) Use. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18), Stockholm, Sweden, 13–19 July 2018; Volume 18, pp. 5697–5702.
81. Pillai, L.R.; Veena, G.; Gupta, D. A combined approach using semantic role labelling and word sense disambiguation for question generation and answer extraction. In Proceedings of the 2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAIECC), Bangalore, India, 9–10 February 2018; pp. 1–6.
82. Dreisbach, C.; Koleck, T.A.; Bourne, P.E.; Bakken, S. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *Int. J. Med Inform.* **2019**, *125*, 37–46. [[CrossRef](#)]
83. Zheng, W.; Cheng, H.; Yu, J.X.; Zou, L.; Zhao, K. Interactive natural language question answering over knowledge graphs. *Inf. Sci.* **2019**, *481*, 141–159. [[CrossRef](#)]
84. Dimitrakis, E.; Sgontzos, K.; Tzitzikas, Y. A survey on question answering systems over linked data and documents. *J. Intell. Inf. Syst.* **2020**, *55*, 233–259. [[CrossRef](#)]
85. Li, X.; Hu, D.; Li, H.; Hao, T.; Chen, E.; Liu, W. Automatic question answering from Web documents. *Wuhan Univ. J. Nat. Sci.* **2007**, *12*, 875–880. [[CrossRef](#)]
86. Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Yu, P.S. A survey on knowledge graphs: Representation, acquisition and applications. *arXiv* **2020**, arXiv:2002.00388.
87. Hamed, S.K.; Ab Aziz, M.J. A Question Answering System on Holy Quran Translation Based on Question Expansion Technique and Neural Network Classification. *J. Comput. Sci.* **2016**, *12*, 169–177. [[CrossRef](#)]
88. Fernández, M.; Cantador, I.; López, V.; Vallet, D.; Castells, P.; Motta, E. Semantically enhanced information retrieval: An ontology-based approach. *J. Web Semant.* **2011**, *9*, 434–452. [[CrossRef](#)]

Short Biography of Authors



Ammar F. Arbaeen received the B.Sc. (Hons) degree in Mobile and Wireless Computing from The University of Westminster, UK, in 2011 and the M.Sc. degree in Advanced Computer Science from The University of Manchester, Manchester, UK, in 2012. He is currently pursuing the Ph.D. degree in Computer Science at IIUM University, Malaysia. From 2013 to 2014, he was a Research Assistant with the Transportation and Crowd Management Centre of Research Excellence. In 2014, he joined Umm Al-Qura University as a lecturer and member of the Information and Scientific Services Department with The Institute of Hajj and Umrah Research, KSA. His research interests are to investigate how to semantically enhance question answering technology in order to understand and support human precisely includes NLP, semantics, and knowledge graph technology.



Asadullah Shah started his career as a Computer Technology lecturer in 1986. Before joining IIUM in January 2011, he worked in Pakistan as a full professor (2001-2010) at Isra, IOBM and SIBA. He earned his Ph.D. in Multimedia Communication from the University of Surrey UK in 1998. He has published 180 articles in ISI and Scopus indexed publications. In addition to that, he has published 12 books. Since 2012, he has been working as a resource person and delivering workshops on proposal writing, research methodologies and literature review at Student Learning Unit (SLeU), KICT, International Islamic University Malaysia and consultancies to other organizations. Professor Shah is a winner of many gold, silver and bronze medals in his career and reviewer for many ISI, and Scopus indexed journals, and other journals of high repute. Currently, he is supervising many Ph.D. projects in KICT in the field of IT and CS.