OntoDin: An Islamic Ontology of Quran and Hadith

Fairouz Bendjamaa
Department of Computer Science, Badji Mokhtar University
Algeria
bendjamaafairouz@gmail.com

Nora Taleb
Department of Computer Science, Badji Mokhtar University
Algeria
talebnr@hotmail.fr

Abstract: *In the religion of Islam, the Quran and Hadith are considered the principal sources of legislation, and they comprise a large amount of knowledge in an unstructured textual form. Islamic texts are organized in such a way that automatic exploitation is both challenging and difficult to achieve. Lately, researchers have been interested in developing a formal description of Qur'an and Hadith semantics based on ontologies to allow computer systems to leverage their knowledge. The main problem in this field is the lack of a complete ontology that allows for a thorough understanding of Islamic texts. In this article, we attempted to solve this problem by developing an ontology covering the maximum amount of data possible. For the Hadith ontology, we opted for a method using only protégé plugins, where we based our work on Comma-Separated Values (CSV) files containing Hadith texts in Arabic and English with all the related data. This method allowed us to ensure the correctness of the results. The final result of this research is a complete Hadith ontology containing a total of 6 classes and 58 487 individuals, of which 34 373 are Hadiths in Arabic and English, with all the related data such as the list of narrators, the book, the chapter mentioned, and the topic it discusses. To further extend the ontology, we merged the Hadith ontology with full pre-existing Qur'an ontologies to build the OntoDin ontology, representing the Islamic texts with 51 classes and 168 122 individuals in Arabic and English*.

Keywords: *Ontology engineering, semantic representation, protégé, religious texts*.

Received December 11, 2023; accepted September 3, 2024
https://doi.org/10.34028/iajit/21/5/1

1. Introduction

Nowadays, the number of religions spread around the world has increased to thousands. Some are well-known and spread, like Islam, Christianity, and Buddhism, and others are not.

Islam is considered one of the most followed religions, born in the early 7th century in Mecca. Muslims believe in the origin of most faiths and scriptures, and Islam is Allah's last word to humanity, transmitted by the last prophet Mohammed (peace be upon him). Muslims have extensive books and theological sources that help them learn and practice their faith. The first source of Islamic knowledge is the Qur'an, considered the holy book of Muslims, as Allah's words revealed to his prophet Mohammed (peace be upon him). The second source is Hadith, which is the prophet Mohamed's actions and sayings (peace be upon him). We can also find sources formed through means other than revelation, such as the companion's consensus (Ijama) and the analogical deduction (Quias).

Millions of Muslims worldwide applaud initiatives enhancing people's ability to learn, recite, memorize, comprehend, and grasp the beautiful teachings of Islamic legislative sources.

Aldhlan *et al*. [2] state that the Qur'an can be considered a gold standard for being the source of a considerable collection of analysis and text experimentation. Hence, several works were built on the Qur'an and Hadith texts.

In the literature, the methodologies employed to define and manipulate Islamic knowledge may be broadly divided into two main categories: semantic and keyword-based.

Semantically-based works are constructed as ontologies that formalize concepts, terminology, or relationships within a specific field. They provide us with information that is both human-readable and computer-readable.

The first definition of ontology was "an explicit specification of a conceptualization" [12]. According to Kazi1 *et al*. [17], the main purpose of an ontology is to capture and enable sharing knowledge about a domain of interest.

Dealing with the Arabic language automatically is considered difficult because of its morphological and structural characteristics, such as the multiple meanings and irregular forms of certain words.

These problems and limitations are more prominent when we talk about Islamic sources, especially the holy book, which is considered the most accurate text that presents the superiority and perfection of the Arabic language [5].

The fundamental issue with previous works is that they are limited in application because most of them

focus on a single domain and only provide a sample ontology. On the other hand, some books use a translated version of the Hadith rather than the original Arabic text. The main issue was the difficulty in handling the classical Arabic used in religious texts.

These limitations may have a direct effect on the outcomes and influence the meaning of the Islamic texts. The Qur'an and Hadith have a higher correlation because Hadiths complete and expand the knowledge encompassed within the Qur'an, which helps the reader fully understand it and avoid any misunderstanding, confusion, or misquotation. Moreover, sample domain ontologies provide difficulties in acquiring and representing knowledge, leading to erroneous and unintelligible search results and resources [20].

Additionally, in certain circumstances, employing a translated version might lose weight or change the meaning that the original text was attempting to transmit. In this study, we investigated whether a comprehensive Islamic ontology could yield better results to satisfy these constraints and whether doing so could enhance understanding of Islamic texts while maintaining their precise meaning.

To start, we provide a new ontology that encompasses the vast majority of Hadith knowledge from the most trustworthy Sunna books, including the entire Qur'an.

The Hadiths are included in the final ontology in their original Arabic form, as well as an English translation. We then merged it with a Qur'an ontology to have a complete coverage of Islamic knowledge spread over Hadith, Qur'an, and Tafsir, as this is one of our primary objectives for this project.

The final ontology is a global representation, gathering the main Islamic sources to provide the final user with a full grasp of any concept. This is primordial for understanding the Islamic texts, as a given concept can be discussed in different sources. For this reason, the outcome of this research can be used not only by simple users to answer their questions but also by domain experts to help them retrieve Islamic judgments to understand complex situations.

The rest of this article is organized as follows: The second section compiles a list of similar works that have been applied to Qur'anic and Hadith texts. The methodology of our ontology is described in the third section. We describe our ontology's methodology. The tests and results of our final application are given in the fourth section, and the study is wrapped up with a summary of the findings and suggestions for further research.

2. Islamic Ontologies

The following are ontological models based on Hadiths and Qur'an texts:

Al-Masri and AlAgha [3] present a method to support semantic search with complex queries by proposing a new ontology to model concepts from the Al-Shamela Digital Library (ADL). They build a Hadith ontology modeling the different entities in the Prophetic Medicine domain. For the evaluation process, they compared the results obtained from their system to the ADL results.

Basharat *et al*. [8] proposed the first linked data model to link Hadith with the Qur'an and other Islamic sources. They used their own Hadith ontology and based their work on 25,934 Hadith written in Arabic and 18,040 Hadith in English.

Azmi and Bin Badia [6] proposed an e-narrator system that parses the Hadith text and generates narrators' chains. The authors selected 90 different Hadiths from Sahih El-Bukhari and Sahih Muslim for the testing process. This system was able to achieve a success rate of 86.7%.

Baraka and Delloul [7] proposed a decision support system to judge Hadith Isnad using ontologies. They re-used the HadithRDF ontology presented in Rusli *et al*. [20] and expanded it by defining new concepts, properties, and instances to achieve their research goal based on the methodology used by Hadith scholars. They evaluated their system using El-Albani scholars and Hadith experts. The first approach achieved 75% accuracy, while the second achieved 81%.

Harrag [14] proposed a Hadith ontology based on Sahih Al-Boukhari. He used association rules to extract the relations between the concepts in Sahih Al-Boukhari.

Al-Rumkhania *et al*. [4] proposed a Hadith ontology for prophetic medicine 'الطب النبوي'. They used authentic Hadiths as a corpus and proposed future work to extend the ontology further to generate treatments for some diseases based on Prophetic medicine automatically.

Saad *et al*. [21] proposed a new methodology based on a combination of Natural Language Generation (NLP) techniques, information extraction, and text-mining technologies to generate ontologies representing Islamic knowledge. This work was applied to sources related to the obligatory prayers in Qur'an and Hadith.

Khan *et al*. [18] developed an ontology sample to represent knowledge related to creatures like animals and birds mentioned in the Qur'an. They used protégé and SPARQL Protocol and RDF query language to interrogate the ontology. They also made several suggestions for broadening this work to include all Qur'anic texts, such as dynamic contextual ontology and Islamic and Qur'anic WordNet development. They claimed that their study could be expanded to include other Islamic sources such as Hadith and Fiqh in the future

Hakkoum and Raghay [13] proposed an ontology covering the following subjects: Qur'anic chapters and verses, each word of the Qur'an and its root, and lemma to facilitate the keyword search. It is also updated with the related of Qur'an VOCabulary (QVOC) which

consists of a multilingual RDF representation of ontology links available and other knowledge related to the Qur'an, like people, events, and places.

Ta'a *et al*. [24] proposed a knowledge retrieval system based on a Qur'anic ontology developed in Java. They built a Qur'anic ontology and proposed a system to facilitate knowledge retrieval with accuracy. Their ontology used a translated version of the Qur'an implemented on Protégé-OWL and validated by Islamic experts.

Bendjamaa *et al*. [10] proposed a sample Hadith ontology representing volume 1 of "Sahih El-Boukhari". They used the Methontology methodology and a manual process on Protégé where they defined the list of concepts using the list of the most relevant words extracted on RapidMiner. The final ontology was evaluated using SPARQL queries and the authors state that the ontology can be combined with other ontologies in the future for a full Islamic ontology.

Junaidi *et al*. [16] proposed an ontology to expand the Hadith understanding. Differently than other works applied to Hadith texts, concentrated more on the comprehension of Hadiths and directed their efforts on Hadiths that used the word Sami'a, which refers to hearing in the literature. The work contains the concepts: of Allah, belief, and Shariah.

Sultana *et al*. [23] proposed a semantic search approach to retrieve food-related concepts from an English version of the Qur'an. They used 105 experimental queries for the evaluation, and the results showed 97.02% precision and 93.33% recall. The authors aim to expand their ontology to represent Food-related Hadiths too.

Abdelkader *et al*. [1] introduced and built an ontology for the main concepts and knowledge of Hadith Science in Arabic. They used Protégé, Java, and Semantic Query-Enhanced Web Rule Language of an Application Programming Interface (SqwrlAPI) for ontology implementation and querying, respectively.

Beirade *et al*. [9] proposed an ontology based on the semantic representation of Qur'anic terms classified into levels. The proposed ontology models 28 terms from the time names present in the Qur'an. They used Java to build a search engine to query the ontology. They used relevant Wikipedia pages to enhance their results. To obtain precise responses, they employed two distinct query expansion techniques.

Sadi *et al*. [22] proposed a new ontology covering all the Qur'anic verses related to the Nature domain. They used SPARQL Protocol and RDF query language Protocol and RDF query language to retrieve the concepts, the full text of the verse, and the verse number related to the user's query. This work uses an English translation of the holy Qur'an proposed by Sahih International.

Ghanem *et al*. [11] proposed a methodology for Islamic law concept extraction from Hadith text. They used the extraction pattern from bootstrapping, where they manually annotated the Hadiths from Sahih Muslim and Sahih Al-Bukhari.

In this little study, we noticed a lack of work focused on Arabic in general, and this deficiency becomes apparent when dealing with religious texts, as the need to computationally represent this information remains critical for all Muslims. This limitation is due to the richness of the Arabic language and its complex syntax, as automatically processing Arabic presents several obstacles and conflicts.

Recently, many scholars have been engaged in using Islamic legislative sources as a corpus with various techniques such as natural language processing, data mining, and machine learning.

The techniques used to represent and treat Islamic knowledge are divided into two categories: semantic-based and keyword-based. The first category searches the query words' contextual meaning, whereas the second retrieves the information by keyword matching or morphology search.

Some of these works, such as the works of Abdelkader *et al*. [1] and Basharat *et al*. [8], represent a large amount of knowledge, while others, such as those presented in Al-Masri and AlAgha [3] and Khan *et al*. [18], appear to focus on one particular domain and provide only a sample ontology.

We can also categorize these works based on the language of the Islamic scriptures, with some works using Islamic sources in their original standard Arabic, such as Harrag [14], which uses Hadith in Arabic, and Bendjamaa *et al*. [10], which uses the Qur'an in Arabic. Others, like Ta'a *et al*. [24]**,** employed a Malay translation of Hadith.

Table 1 below summarizes the similar works from the corpus used, the coverage of the final ontology, and outlines the main advantages and disadvantages of the ontology. After analyzing it, we can conclude that the most frequent weakness faced is the limitation in the coverage of Islamic knowledge and the lack of evaluation of the obtained ontologies. Starting from this point, we knew that this domain needed a full Islamic ontology rather than focusing our efforts on sample ontologies that could not give a complete and precise answer. Thus, we decided in this paper to present a full Islamic ontology covering the main two sources of Islamic Jurisprudence: the Qur'an and Hadith. It is critical to consider reusing existing ontologies to take advantage of them and give them more importance.

Table 1. Summary of similar works.

Paper	Corpus	Language of the corpus	Coverage	Strength	Weakness
[3]	Hadith-El Boukhari	Arabic	Partial: the prophetic Medicine domain	Uses the original Arabic text-enhanced the ADL search	A sample ontology-uses only keyword searching
[8]	Hadith	Arabic and English	25,934 Hadith	Covers an important number of Hadiths-uses the original Arabic corpus	No validation
[6]	Hadith-Sahih El Boukhari and Sahih Muslim	Arabic	Partial: 90 Hadiths	Study the narrator chain-uses the original Arabic corpus	A sample ontology
[7]	Hadith	Arabic	Partial	Reuse of the ontology in [20]: concept extension	A sample ontology-no evaluation
[14]	Hadith: Sahih El Boukhari	Arabic	Partial	Automatic concepts and relations extraction	A sample ontology-no evaluation
[4]	Hadith: Sahih El Boukhari	Arabic	Partial: the prophetic Medicine domain	Reuse of the ontology in [20, 3], concept extension	A sample ontology-no evaluation
[21]	Qur'an and Hadith	Malay	the obligatory prayers	Uses the Qur'an Indexes as the upper layer of the ontology	A sample ontology
[18]	Qur'an	Malay	Partial	Expert validation	A sample ontology
[13]	Qur'an	Arabic	Qur'an	A full Qur'an ontology-uses the original corpus	No evaluation
[10]	Hadith: Sahih El-Boukhari	Arabic	Partial: volume 1	Semi-manual building: high Precision	A sample ontology
[16]	Hadith	Arabic	Hadiths containing the word Sami'a	Novel work to enhance the understanding of Hadith using Isnad and based on the word Sami'a uses	A sample ontology-no evaluation
[23]	Qur'an	English	Partial: food-related verses	Good evaluation results, aiming to expand using related Hadiths	A sample ontology
[1]	Hadith	Arabic	Major Hadith books	A full Hadith ontology	No evaluative or validation
[9]	Qur'an		28 time-names	Best evaluation results compared to similar works	Using non-reliable sources such as Wikipedia pages
[22]	Qur'an	English	Partial: nature domain		A sample ontology-no evaluation
[11]	Hadith: Sahih Muslim and Sahih Al-Bukhari	Arabic	/	Extracting concepts using pattern bootstrapping	No evaluation or details on the developed ontology

3. The OntoDin Ontology

For our OntoDin ontology implementation, we chose to work on Islamic texts and Hadith specifically as we saw a huge lack in the ontology building side. After we studied similar works, we found an available full Qur'an ontology that can be reused, this led us to the idea of using it in our Islamic ontology and hence focusing our efforts on developing a full Hadith ontology and finally merging it into Qur'anic ontology for a complete representation of Islamic knowledge where Qur'an and Hadith are considered as the main sources of Islam legislation.

We named the obtained ontology "OntoDin;" an acronym for Ontology and the Arabic word of religion "Din".

3.1. Hadithonto

Numerous approaches to ontological engineering outline the steps involved in building an ontology.

For our case, we opted for a semi-automatic process using Protégé plugins mainly. This will help us be more productive by reusing pre-existing data and also testing the accuracy of Protégé in handling Arabic texts.

According to Hakkoum and Raghay [13], there is no one ideal approach to building an ontology; rather, the optimal approach is determined by the intended use of it and your future perspectives.

Since this ontology will eventually be combined with the Qur'an ontology in Hakkoum and Raghay [13] we decided to be inspired by the same process to make the final merging of both ontologies straightforward and prevent any heterogeneous or compatibility issues.

In light of this, we opted for the ontology development 101 guidance proposed by Standford University presented in Noy and McGuinness [19].

In our work, we chose to adhere to the phases in Figure 1 that were most suitable for our needs, taking into consideration the sources we are starting with and the results we are aiming to get at the end of this process:

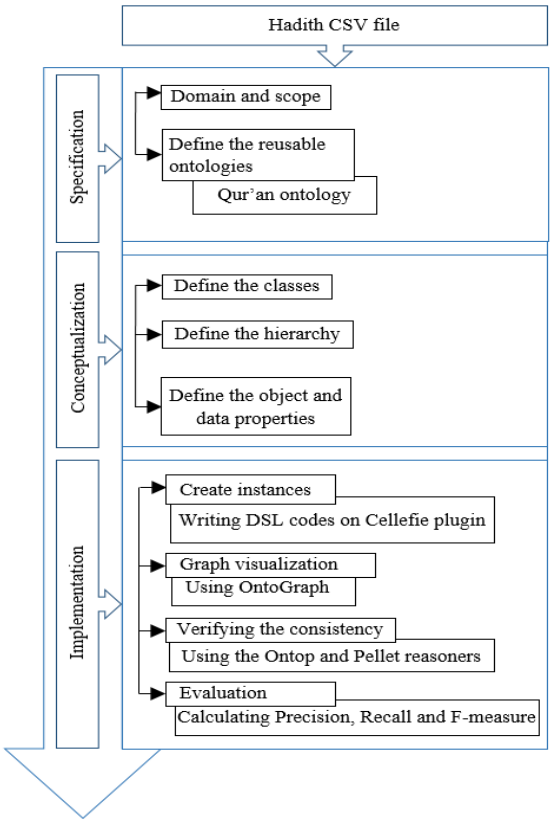

Figure 1. The ontology-building process for the HadithOnto.

3.1.1. Domain and Scope

- **Domain**: hadith knowledge which is the set of words and acts derived from the prophet of Islam Mohamed.
- **Objective**: the semantic representation of the Hadith to be used for understanding Islam and the automatic extraction of Islamic laws.
- **End-users**: anyone who plans to build a system based on the automatic extraction of Hadith text.
- **Source of Knowledge (Corpus)**: 6 major collections of Sunna: Sahih Bukhari, Sahih Muslim, Sunan Nasa'I, Sunan Abi Da'ud, Sunan Ibn Majah, Jami' al-Tirmidhi.

3.1.2. Define the Reusable Ontologies

The purpose of this search is to build an ontology by combining various ontologies that cover the same or related topics into a single, comprehensive ontology. We aim to use the HadithOnto to build a full Islamic Ontology where we chose to reuse the Qur'an ontology presented in Hakkoum and Raghay [13].

We chose this particular ontology for several reasons; we mention:

- The ontology models all the Qur'an in Arabic and English with related Tafsir which is beneficial for our objective to build a large Islamic ontology with all the main knowledge of the domain.
- The ontology is simple and easy to work with.
- The structure of the ontology is similar to our Comma-Separated Values (CSV) files and will facilitate the merging, considering the structure of the Hadith books divided into Chapters with themes.
- The availability of the ontology in its complete and correct owl file.

3.1.3. Define the Classes and the Hierarchy

According to Al-Yahya *et al*. [4], we can extract the important terms in two ways, either by going through the Hadiths one by one or based on the topics the Hadiths treat. Similarly, we chose to follow the second method as the Hadith books are already divided into chapters and each chapter treats a given topic. For an all-coherent ontology, we will build it using the following classes: hadith, collections, topic, historic person, and grade.

For the implementation, we chose Protégé 5.5.0 for the following reasons:

- A free and open-source platform.
- Among the most used tools in the literature.
- Regularly updated with new versions.
- A large user community with responsive assistance.
- Featuring many useful plug-ins, we aim to employ them in the next steps to facilitate and fasten our ontology-building process.

In Figure 2 we can find the class hierarchy of our ontology on Protégé.

Figure 2. The class hierarchy of HadithOnto.

3.1.4. Define the Object Properties and the Data Properties

As already elaborated in the previous step. We kept the list of the classes simple and clear where the relations are easy to draw. In Figure 3 we have a list of all the object properties and in Figure 4 the list of the data properties.

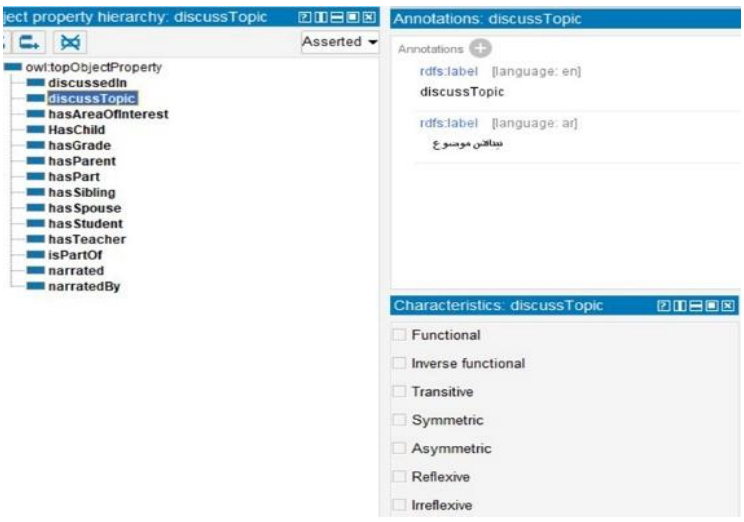

Figure 3. We defined the object properties of OntoHadith.

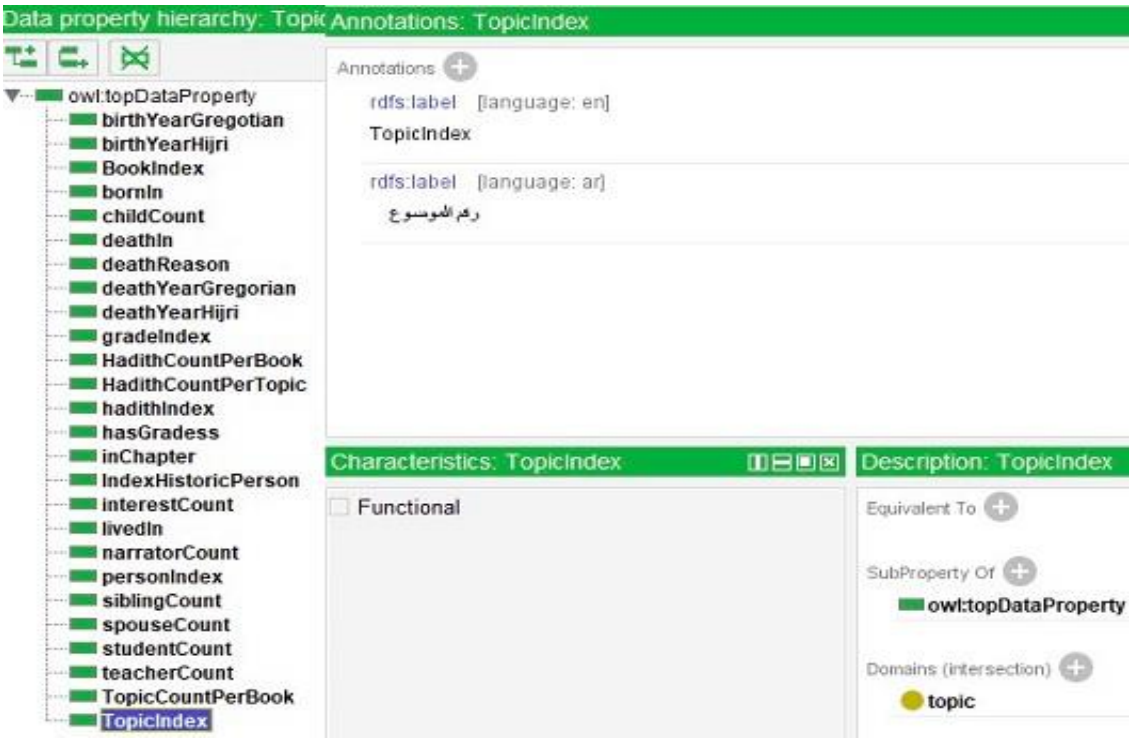

Figure 4. The data properties of HadithOnto.

3.1.5. Create Instances

To create the instances of each class, we used a CSV file containing the hadiths in lines where the data is separated by commas. To do so, we used the Cellefie plugin rules where we wrote our import script using Domain Specific Language (DSL) code to define the links between the spreadsheet content and the Web Ontology Language (OWL) ontology. This plugin helped us define all the mapping and import the data from the CSV file to our ontology and create the axioms. In Figure 5 we can see an example of DSL code used to import data to the Historic person class.

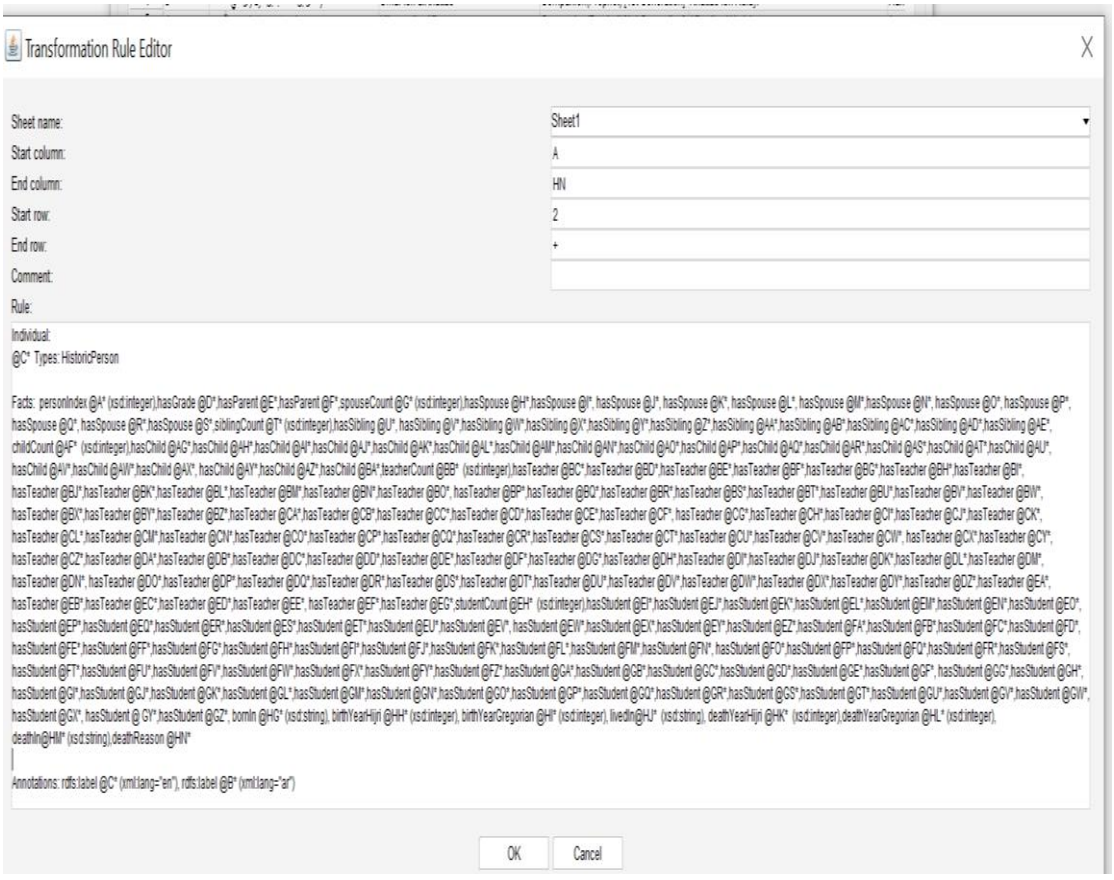

Figure 5. Importing the historic person CSV files to protégé using the Cellephie plugin.

Table 2. Ontology metrics.

Category	Count
Classes	6
Object property	15
Data property	27
Individual count	58 487

In Table 2, the final ontology build using protégé 5.5.0 was successfully completed with a total of 6 classes and 58487 individuals.

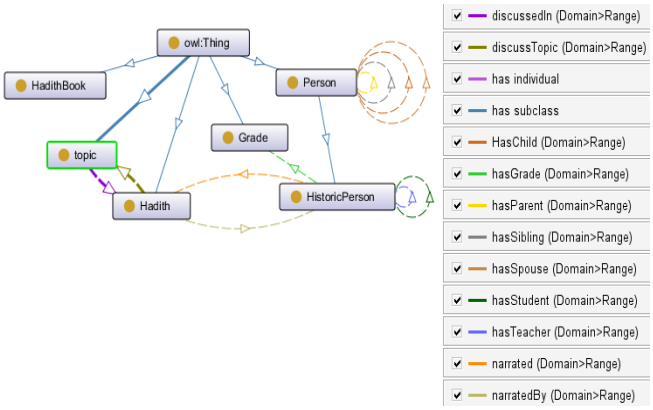

Figure 6. The HadithOnto graph using OntoGraph.

To visualize the ontology, we created the ontology graph in Figure 6 using the OntoGraph plugin.

3.2. The OntoDin

Ontology Merging is building an ontology by combining knowledge from two ontologies into an existing or new ontology. In this step, we opted to merge our Hadith ontology with a pre-existing Qur'an ontology into a full Islamic ontology using the "merge ontologies" plugin of protégé.

From the literature study, we chose the ontology presented by Hakkoum and Raghay [13] as it was a full Qur'an ontology up to date and always available on the web. The chosen ontology perfectly responded to our needs and was our inspiration in building our Hadith ontology to facilitate the merging step.

As known in the literature, the merging process is hard and both effort and time-consuming so from the beginning, we studied the Qur'an ontology and we tried to find similar knowledge between the ontology and the CSV files we started with to determine common concepts and draw our ontology building process based on it. This method helped us to easily merge the two ontologies and avoid any conflicts.

To accomplish this step, we chose another Protégé plugin for merging two ontologies using the following steps:

- *Step* 1. On protégé, we opened the Hadith ontology and added the Qur'an ontology in the direct imports in the same window, then from the "refactor" menu, choose the "merge ontologies" feature.
- *Step* 2. Select the first choice to import an ontology contained in a local file from the import ontology wizard and press continue. At the end of this step, we got the classes of both ontologies merged into one ontology. However, we noticed that the import wizard did not merge similar classes like "topic" from Qur'an ontology and "topic" from the Hadith ontology.
- *Step* 3. To eliminate the different duplicates presented in the ontology, we had to rename one of the classes to match the second one. For example, we changed the name of the class HistoricPerson to Historic Person by changing its Uniform Resource Identifiers (URI) class by selecting "rename entity" in the Refactor menu.

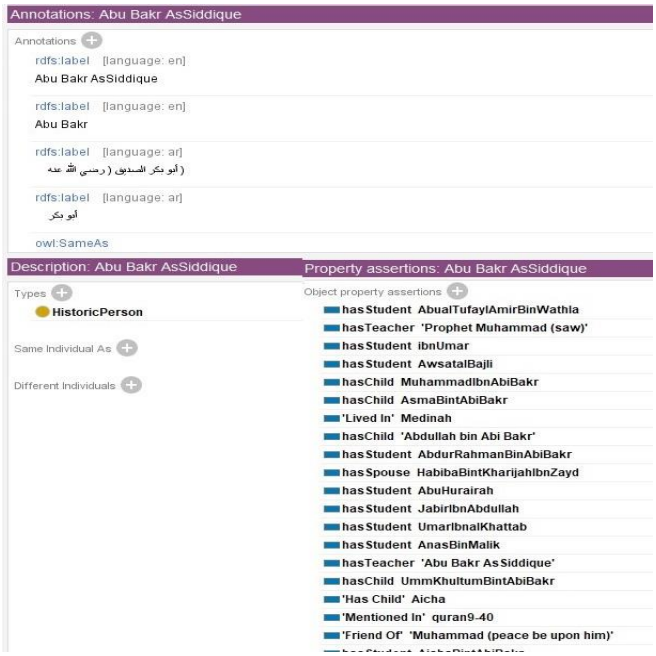

Figure 7. A merged individual.

- *Step* 4. We repeat the third step with the duplicate classes and individuals. As an example, in Figure 7, we got a complete individual, "Abu Bakr as Seddique," with all the data and object properties of both Hadith and Qur'an ontologies.

Table 3. Ontology metrics of OntoDint.

Category	Count
Classes	51
Object property	54
Data property	50
Individual count	168 122
Triples	822 526

In Table 3, the final Islamic ontology built using protégé 5.5.0 was successfully completed with a total of 6 classes and 58487 individuals.

To visualize the ontology, we created the ontology graph in Figure 8 using the OntoGraph plugin.

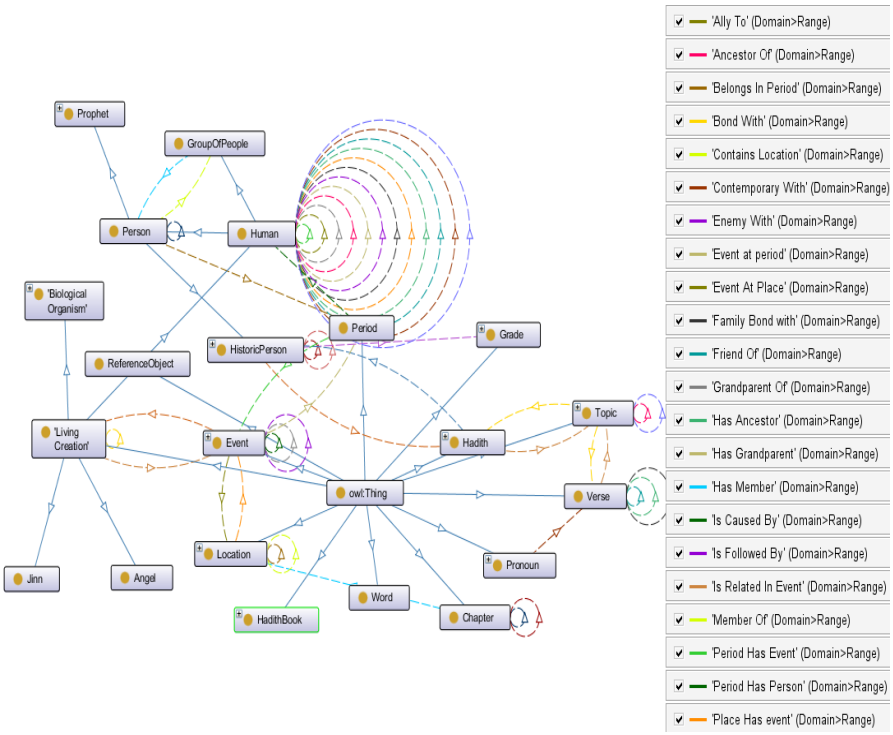

Figure 8. The OntoDin graph using OntoGraph.

3.3. Results and Evaluation

First, we started by validating our ontology by checking its logical consistency using Protégé reasoners like Ontop and Pellet, and the result was that the Ontology was consistent.

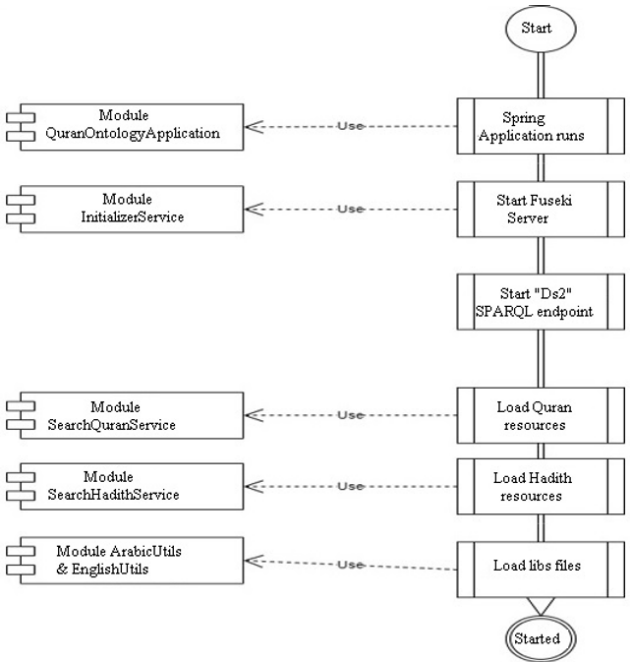

Figure 9. Life cycle of the AskIslam.

Then, for the evaluation part, we had to build our search engine that would facilitate questioning the ontology. The search engine named "AskIslam" was built using JAVA and the Jena Fuseki server. Figure 9 presents the life cycle of AskIslam.

The search engine of our web application is divided into two subcategories:

- Search by: topic, keyword, or Tafsir.
- Search in Qur'an, Hadith, or both.

Figure 10 presents a detailed diagram of the search engine.

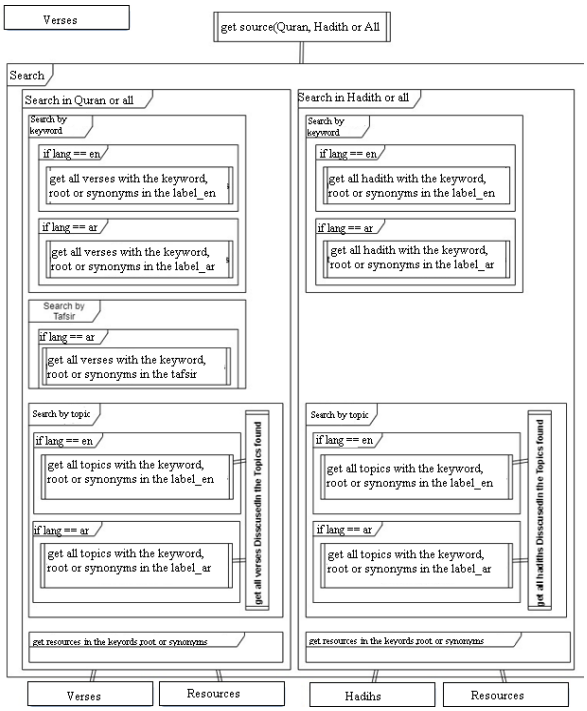

Figure 10. A diagram of the search model.

Figure 11 presents the full architecture of the search engine.

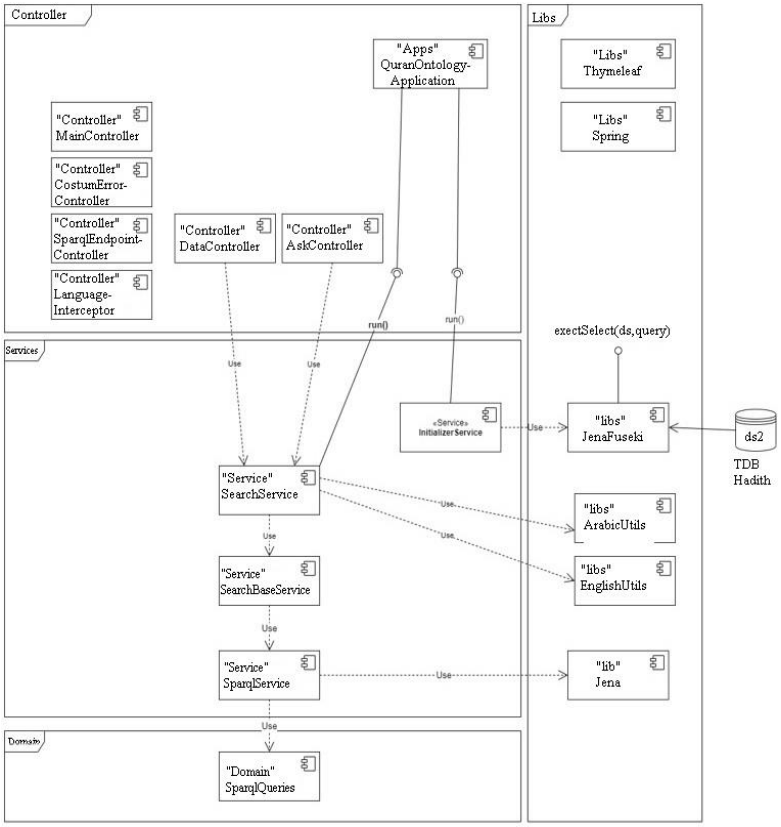

Figure 11. Logical architecture of AskIslam.

Figure 12 presents the search engine page of our AskIslam web application. The engine accepts direct questions and answers by displaying the sources where a term appears or where it was discussed as in Figure 13.

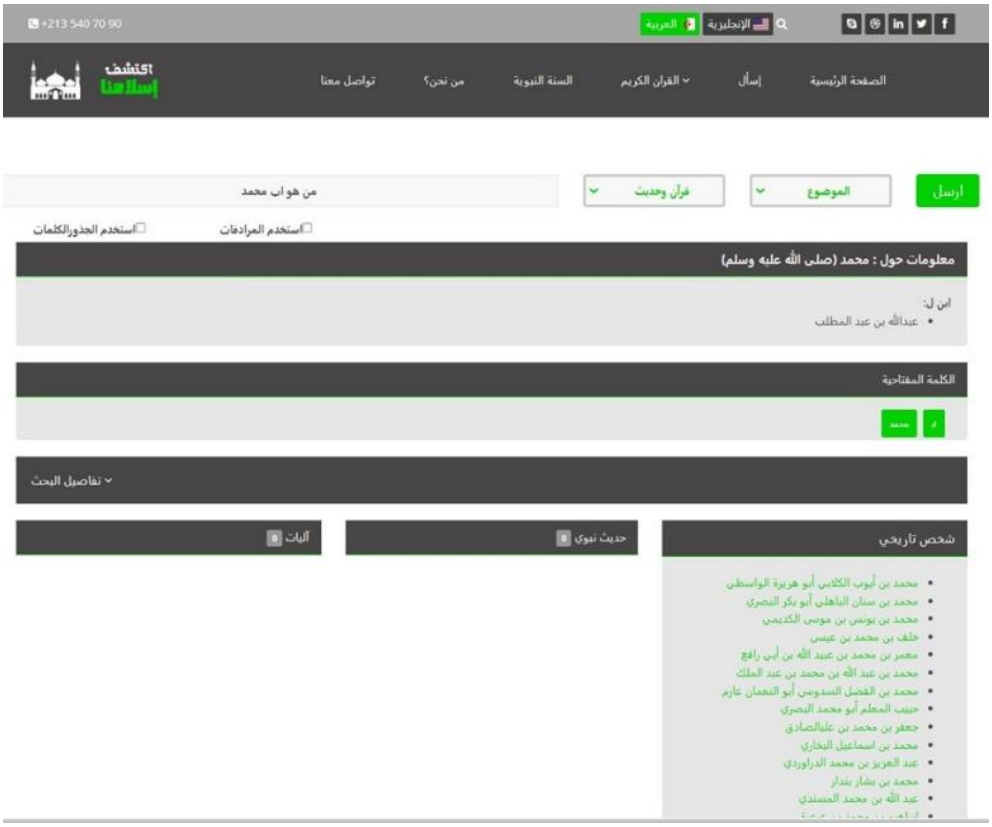

Figure 12. Answering a question directly from the ontology.

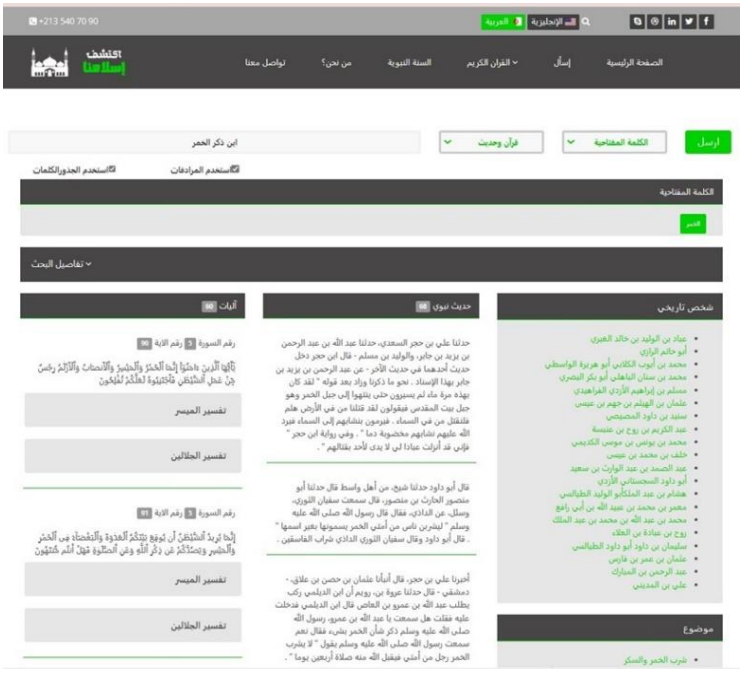

Figure 13. Citing all the sources discussing the topic searched.

For the evaluation process, we chose to calculate the recall, precision, and F1-measure metrics according to Harrag *et al*. [15], we calculated the results based on a total of 80 queries.

The questions used to evaluate our system were chosen based on the topics the most searched in Islam like Murder, Testament, marriage, and Ramadan.

Equation (1) calculates the precision metric is the total of the relevant answers divided by the total of all the retrieved answers,

$$Precision = \frac{Relevent \cap Retrieved}{Retrieved} \quad (1)$$

Equation (2) calculates the recall metric is the total of the relevant answers divided by the total of the relevant answers,

$$Recall = \frac{Relevent \cap Retrieved}{Relevant} \quad (2)$$

Equation (3) calculates the F-measure metric is calculated based on the recall and precision,

$$F1 - Measure = \frac{2\ (Precision\ X\ Recall)}{Precision + Recall} \quad (3)$$

We manually generated the metrics for each question by querying our ontology, going through all the retrieved responses (retrieved), and comparing them to the correct list of answers from Hadith and Qur'anic (relevant). From there, we were able to get the metric value for each question. Table 4 presents a sample of the queries and answers retrieved from our search engine AskIslam.

Table 4. Sample queries to evaluate the AskIslam.

Query	Answer	Precision	Recall	F-measure
طلاق	24 verses and 7 hadiths	0,88	1	0,93
ابليس	9 verses	1	0,81	0,9
سجد	20 verses and 19 hadiths	1	0,58	0,74

Table 5 presents the results of the metrics used to evaluate the effectiveness of our search engine AskIslam using 80 different queries. As shown, our system achieved 89% precision, 82% recall, and 82% F-measure.

Table 5. The metric results of evaluating the AskIslam.

Number of queries	Average precision	Average recall	F1-measure
80	89%	82%	82%

4. Discussion and Future Works

Our first Hadith ontology OntoHadith includes 34373 hadiths in Arabic and English, as well as all related data such as the list of narrators, the book and chapter mentioned in, and the topics it discusses, as shown in Figure 5 above.

Compared to the existing works, we attempted to eliminate the limitation of using a single topic of Hadiths by developing a complete ontology with over 34500 Hadiths in Arabic and English for a complete grasp and understanding of Hadith knowledge.

To further extend the ontology, we added the majority of Hadith's narrators as a separate concept, which will help in searching for information about the narrators but, more importantly, in studying the narration chain (Isnad) of a given Hadith.

Then, the second ontology, OntoDin, which we obtained by merging the OntoHadith and a Qur'an ontology, gave us a full Islamic ontology with the majority of data needed to fully grasp the meaning of each topic and allow for a high level of knowledge extraction when used in domains like Fatwa or Islamic research. Reusing pre-existing ontologies saved us a lot of time and effort, as well as allowing the Qur'an ontology to improve its capabilities when merged with Hadith.

As seen in the previous section, we did the majority of data collection and verification work manually,

because we believe that such critical texts must preserve their correctness to 100% and not even 99%. For this reason, we opted to create the OntoDin ontology. As a result, we chose to build an ontology using protégé plugins based on our collected and verified data.

It is noteworthy that applications based on Islamic texts have numerous flaws and challenges, with the main problem being that they use limited Islamic resources, resulting in incomplete or incorrect information retrieval.

Our mission was to build a domain ontology with the maximum of Islamic knowledge possible, and obviously, the only way to do it was to build an ontology covering the main sources of Islam the holy book with Tafsir and the major Hadith books. To facilitate this process and give the project the best conditions to obtain good results we decided to go with known and tested techniques and tools. As the final ontology was a result of merging our HadithOnto with a pre-existing Qur'an ontology, we found no better solution than using a close methodology to avoid any inconsistency in the final OntoDin ontology.

To evaluate our final ontology, we calculated precision, recall, and F-measure metrics on 80 queries, and the results were 89%, 82%, and 82%, respectively. Even though this is a good result for a start, it is not enough when considering the critical texts, we are dealing with.

Because of its simple structure, the OntoDin ontology is very comprehensive and easy to manipulate, with verses and hadith texts as the main concepts. This advantage can be seen as a disadvantage if you want specific concepts to be explicitly clear in the ontology.

As previously stated, we chose this path because of its feasibility, considering the size of the data we were dealing with and how time- and effort-consuming the second method may be. If necessary, this expansion can be added later for greater flexibility, which will improve accuracy and efficiency.

This drawback is evident when searching for one particular concept or term from the Qur'an or Hadith, as the texts of the sources are put together as instances of the verse or hadith concepts. However, introducing the class "topic" with the relationship "discussed in" allowed us to enrich the ontology and were capable of defining the main concepts discussed in the Qur'an and Hadith like "prayers," "medicine," "financial transactions" and more.

Ultimately, the methodology used was easy to work with, considering we had to perform the preprocessing work manually. This method allowed us to achieve a high precision percentage by checking and verifying the files and the important data in use.

The conceptualization was simple, given the inspiration we had from the chosen Qur'an ontology to be merged with.

Our expectations about time gain were not met by the decision to use exclusively Protégé plugins.

The creation of the instances using the Cellefie plugin was a straightforward and uncomplicated process; however, writing the DSL code for each concept, taking into consideration the significant number of columns we had, was a little bit time-consuming. However, the drawback of this method is the time we had to allow this step to fully upload all the instances. This step took more than three months to complete the importation of the 34373 hadiths and all their related concepts.

One of the major problems we faced and we were not expecting, was the ambiguity of the results as the final ontology was obtained through merging two ontologies, we found a lot of duplicates in concepts and instances. We eliminated all the duplicated topics and instances by changing their Internationalized Resource Identifiers (IRIs) to match and be merged as one concept or instance.

To further enhance the results obtained by our search engine, we aim to develop a more effective search tool and write more elaborate queries using the SPARQL Protocol and RDF query language; to fully utilize the ontology. The accuracy and completeness of the results will aid in understanding Islamic knowledge.

In the future, we aim to use Artificial Intelligent (AI) and deep learning techniques to help decision-makers in the field of Islamic studies and assist them in decision-making. This step will assist the Muftis in reading, understanding, and interpreting a text, as well as deciding whether an act is Halal or Haram.

Also, the OntoDin ontology can be used in a variety of other applications, such as chatbots to answer Islamic questions in Arabic and English, manual or automatic Hadith validation by studying its Isnad based on the HistoricPerson concept and the full relationships between them, or even to be used in the domain of comparative religions.

5. Conclusions

In this paper, we presented the entire process of building the OntoDin ontology. The final ontology models the knowledge presented in the Qur'an and Hadith in both Arabic and English. We built the Hadith ontology from scratch, and then we merged it with a pre-existed Qur'anic ontology.

For the Hadith ontology creation, we were inspired by the structure of the Qur'an ontology and we aimed to achieve a similar result for the Hadith ontology hence eliminating one of the merging difficulties when using ontologies of different structures.

Our Hadith Ontology was based primarily on the knowledge presented in a CSV file combining data from six major books of Sunnah. The whole process was semi-automatic, where we did the significant data preparation and verification manually, created the ontology manually, and then used the Cellefie plugin integrated into protégé to map the data to the concepts.

Finally, the merging process was also done semi-automatically, where at first, we used the merge ontologies feature from the refactoring tools in protégé. Then, we did the reasoning manually to eliminate any inconsistencies and redundant classes and individuals.

The OntoDin ontology covers all Qur'an and Hadith texts with related data like the prominent scholars and historic people in Islam and detailed biographies of each one.

In this study, we found a literature rich with methodologies, languages, and tools dedicated to ontology building which are developed, used, validated, and considered as a reference. Instead of taking challenges and presenting new techniques or approaches, we chose the safe path to put all our focus on the sources we are dealing with. For our case, the best process was to use a tested methodology and tool in the related works.

To summarize, there is a considerable need to create domain ontologies for Islamic legislative sources rather than focusing our efforts and time on sample ontologies.

For this reason, we presented in this paper an ontology covering all six major Hadith books with over 34 thousand Hadiths in Arabic and English, 24 thousand narrators and scholars, and all Qur'an texts with their 114 chapters, 6236 verses, and the related Tafsir in Arabic and English.

The next step in our research project would be the implementation of a decision-support system based on the created ontology using deep learning techniques to help the deciders of the Islamic law to ask complex questions and facilitate the process of knowing if an act is Halal or Haram by retrieving all the related resources with a predicted classification and precision percentage. The main purpose of building such a large ontology is to give a full representation of all Islamic texts related to a given concept.

6. Acknowledgement

The authors would like to thank the DGRSDT (General Directorate of Scientific Research and Technological Development)-MESRS (Ministry of Higher Education and Scientific Research), Algeria, for the financial support of LISCO Laboratory.

References

[1] Abdelkader A., Najeeb M., Alnamari M., and Malik H., "Creation of Arabic Ontology for Hadith Science," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 6, pp. 3269-3276 2019. DOI:10.30534/ijatcse/2019/96862019

[2] Aldhlan K., Zeki A., and Zeki A., "Data Mining and Islamic Knowledge Extraction: Hadith as A Knowledge Resource," *in Proceedings of the 3rd International Conference on Information and Communication Technology for the Moslem World*, Jakarta, pp. 21-25, 2010. DOI:10.1109/ICT4M.2010.5971934

[3] Al-Masri M. and AlAgha I., An Ontology-based Approach to Enhance Information Retrieval from Al-Shamelah Digital Library, Master Thesis, Islamic University-Gaza, 2015. https://search.emarefa.net/en/detail/BIM-610572-an-ontology-based-approach-to-enhance-information-retrieval-

[4] Al-Rumkhania A., Al-Razganb M., and Al-Faris A., "TibbOnto: Knowledge Representation of Prophet Medicine (Tibb Al-Nabawi)," *Procedia Computer Science*, vol. 82, pp. 138-142, 2016. DOI: 10.1016/j.procs.2016.04.021

[5] Al-Yahya M., Al-Khalifa H., Bahanshal A., Al-Odah I., and Al-Helwa N., "An Ontological Model for Representing Semantic Lexicons: An Application on Time Nouns in the Holy Qur'an," *Arabian Journal for Science and Engineering*, vol. 35, no. 2, pp. 21-35, 2010. file:///C:/Users/user/Downloads/An_Ontological_Model_for_Representing_Semantic_Lex.pdf

[6] Azmi A. and Bin Badia N., "e-Narrator-An Application for Creating an Ontology of Hadiths Narration Tree Semantically and Graphically," *Arabian Journal for Science and Engineering*, vol. 35, no. 2, pp. 51-68, 2010. https://www.researchgate.net/publication/262687526

[7] Baraka R. and Delloul Y., "Building Hadith Ontology to Support the Authenticity of Isnad," *International Journal on Islamic Applications in Computer Science and Technology*, vol. 2, no. 1, pp. 25-39, 2014. file:///C:/Users/user/Downloads/674-754-1-PB.pdf

[8] Basharat A., Abro B., Arpinar I., and Rasheed K., "Semantic Hadith: Leveraging Linked Data Opportunities for Islamic Knowledge," *in Proceedings of the Workshop on Linked Data on the Web Co-Located with 25th International World Wide*, Montreal, pp 1-10, 2016. https://ceur-ws.org/Vol-1593/

[9] Beirade F., Azzoune H., and Zegour D., "Semantic Query for Quranic Ontology," *Journal of King Saud University-Computer and Information Sciences*, vol. 33, no. 6, pp. 753-760, 2021. https://doi.org/10.1016/j.jksuci.2019.04.005

[10] Bendjamaa F., Taleb N., and Arari N., "An Ontological Model of Hadith Texts," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 4, pp. 367-371, 2020. https://thesai.org/Publications/ViewPaper?Volume=11&Issue=4&Code=IJACSA&SerialNo=51

[11] Ghanem M., Mouloudi A., and Mourchid M., "Creation and Populating of an Islamic Knowledge Ontology Using Extraction Pattern Bootstrapping," *in Proceedings of the 3rd National*

Day on Engineering, Networks, and Telecommunications, Kenitra, pp. 36-39, 2015. https://ouci.dntb.gov.ua/en/works/4v3ndpel/

[12] Gruber T., "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquisition*, vol. 5, no. 2, pp. 199-220, 1993. https://doi.org/10.1006/knac.1993.1008

[13] Hakkoum A. and Raghay S., "An Ontological Approach for Semantic Modeling and Querying the Qur'an," *International Journal on Islamic Applications in Computer Science and Technology*, vol. 4, no. 1, pp. 37-45, 2016. file:///C:/Users/user/Downloads/IMAN.pdf

[14] Harrag F., "Text Mining Approach for Knowledge Extraction in Sahih Al Bukhari," *Computers in Human Behavior*, vol. 30, pp. 558-566, 2014. https://doi.org/10.1016/j.chb.2013.06.035

[15] Harrag F., Alothaim A., Abanmy A., Alomaigan F., and Alsalehi S., "Ontology Extraction Approach for Prophetic Narration (Hadith) Using Association Rules," *International Journal on Islamic Applications in Computer Science and Technology*, vol. 1, no. 2, pp. 17-26, 2013. https://www.sign-ific-ance.co.uk/index.php/IJASAT/article/view/483

[16] Junaidi J., Jamal I., Ghazali N., Ahmad H., and Salam R., "Expanding Hadith Understanding Using Ontology," *Journal of Computational and Theoretical Nanoscience, Advanced Science Letters*, vol. 23, no. 5, pp. 4611-4614, 2017. DOI:10.1166/asl.2017.8939

[17] Kazi1 Z., Kazi L., Radulovic B., and Bhatt M., "Ontology-based System for Conceptual Data Model," *The International Arab Journal of Information Technology*, vol. 13, no. 5, pp. 542-551, 2016. https://ccis2k.org/iajit/PDF/vol.13,%20no.5/8333.pdf

[18] Khan H., Saqlain S., Shoaib M., and Sher M., "Ontology-based Semantic Search in Holy Quran," *International Journal of Future Computer and Communication*, vol. 2, no. 6, pp. 570-575, 2013. DOI:10.7763/IJFCC.2013.V2.229

[19] Noy N. and McGuinness D., Ontology Development 101: A Guide to Creating Your First Ontology, Technical Report, 2001. https://www.bibsonomy.org/bibtex/12a496bb63d728d96264a50c0a4960e9a

[20] Rusli A., Ridzuan F., Zaki Z., Sayuti M., and Salam R., "A Systematic Review on Semantic-based Ontology for Quranic Knowledge," *International Journal of Engineering and Technology*, vol. 7, no. 15, pp. 80-86, 2018. DOI:10.14419/ijet.v7i4.15.21376

[21] Saad S., Salim N., Zainal H., and Muda Z., "A Process for Building Domain Ontology: An Experience in Developing Solat Ontology," *in Proceedings of the International Conference on Electrical Engineering and Informatics*, Bandung, pp. 1-5, 2011. DOI:10.1109/ICEEI.2011.6021572

[22] Sadi A., Anam T., Abdirazak M., Adnan A., Khan S., Rahman M., and Samara G., "Applying Ontological Modeling on Quranic "Nature" Domain," *in Proceedings of the 7th International Conference on Information and Communication Systems*, Irbid, pp. 151-155, 2016. DOI:10.1109/IACS.2016.7476102

[23] Sultana Z., Rahman M., Uddin M., and Arfat M., "Developing a Semantic Search Method for Retrieving Food-Related Verses and Concepts from the Holy Quran Using Ontology," *in Proceedings of the 5th International Conference on Electrical Engineering and Information and Communication Technology Military Institute of Science and Technology*, Dhaka, pp. 1-5, 2021. DOI:10.1109/ICEEICT53905.2021.9667817

[24] Ta'a A., Abidin S., Abdullah M., Ali A., and Ahmad M., "Al-Quran Themes Classification Using Ontology," *in Proceedings of the 4th International Conference on Computing and Informatics*, Sarawak, pp. 28-30, 2013. DOI:10.1109/ICTC.2014.6983090

Fairouz Bendjamaa Ph.D. student, working on ontologies and decision-making systems and their application to sacred texts. Graduated from Badji Mokhtar University in 2013 with a Master's Degree in Complex System Engineering.

Nora Taleb professor at the Computer Science Department, University of Badji Mokhtar University. Specialised in: Ontologies, Acquisition of Knowledge, Knowledge Modeling.