

Article

Using 5WH Competency Questions for Validating Information Retrieval from Al-Quran: A Methodological Approach

Nazia Nishat¹, Rosalina Abdul Salam¹, Yusuf Mahbubul Islam² and Zulkifly Mohd Zaki¹

¹Faculty of Science & Technology, Universiti Sains Islam Malaysia, 71800, Nilai, Negeri Sembilan, Malaysia.

²Department of Computer Science & Engineering, SouthEast University (SEU), Dhaka-1208, Bangladesh.

Correspondence should be addressed to:

Nazia Nishat; nishat.nazia12@gmail.com

Article Info

Article history:

Received: 18 November 2024

Accepted: 15 January 2025

Published: 23 February 2025

Academic Editor:

Shahrina Ismail

Malaysian Journal of Science, Health & Technology

MJoSHT2025, Volume 11, Issue No. 1

eISSN: 2601-0003

<https://doi.org/10.33102/mjosht.v11i1.450>

Copyright © 2025 Nazia Nishat et al.

This is an open access article distributed under the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

Abstract— Various methods are used to validate information retrieved (IR) from unstructured text. Measuring the effectiveness of IR systems is an important focus of IR research, based on sample retrieval for testing combined with evaluation measures. Evaluation measures include both quantitative and qualitative aspects. Quantitative measures assess precision, recall, and benchmark scores, while qualitative measures focus on relevance, user satisfaction, and usability. As the quality of information retrieved is of importance to humans, the use of competency questions (CQs) has been introduced, particularly, in the case of information retrieved from Al-Quran - to validate the output. As the design of CQs usually relies on expert support, which can limit usability for general users seeking guidance, this research proposes using basic fact-finding 5WH questions. These are fact finding why, what, how, where, when and who questions. It is proposed to use these as CQs to test the relevance of Verses/Ayats retrieved on the basis of common text-patterns found in Al-Quran. We propose this methodological approach using 5WH CQs to test the competency of retrieved information, focusing on popular concepts like *Salat*, *Zakat*, and the Sun. It is found that sample Ayats retrieved on the basis of text-patterns found answer fundamental questions such as WHY? For example, why pray *Salat*? or why was the Sun created? Results show that this method evaluates the relevance and usability of IR systems. Further work based on repeating text-patterns in the Al-Quran is recommended to test whether this evaluation approach can help ontology learning.

Keywords— ontology learning; competency questions; text-patterns

I. INTRODUCTION

Evaluation methods form an integral part of Information Retrieval (IR) systems. Such methods assess how well the retrieved results, based on a particular index, search engine, or organized database, satisfy a user's needs. Measuring the effectiveness of results retrieved based on user needs is therefore an important focus of IR research. Evaluation

methods are therefore fundamental to the development and success of IR systems in both online and offline platforms.

The relevance of results retrieved is an important determinant of a systems' effectiveness for users [1]. Evaluation measures include both quantitative and qualitative aspects. Quantitative measures assess precision, recall, and benchmark scores, while qualitative measures focus on relevance, user satisfaction, and usability. The evaluation

methodology for an information retrieval system should also include a validation of the measure itself. Essentially, an assessment of the fitness of purpose, in other words, how well the measuring system fits its intended use case [2]. As the quality of information retrieved is of importance to users, the use of competency questions (CQs) [4] has been introduced, particularly, in recent years, in the case of information retrieved from Al-Quran - to validate retrieved output. Generally, in the design of CQs domain experts are involved. The CQs so created, understandably, require expert insight. A user, particularly one seeking guidance from Al-Quran, on the other hand, is unlikely to have background knowledge of the subject. Such a user is likely to have more fundamental questions. In other fields, journalists [9] and the police [9] use fundamental 5WH questions for fact finding. Police use 5WH questions to find facts to solve a crime. Sample questions could be: What is the crime? When and where did it happen? How did it take place? Who committed the crime? Why was the crime committed? Covering all 5WH questions forces looking for facts that help solve a crime. Similarly, before reporting an event, journalists are taught to ask 5WH questions to help write a fact based report. For example, why was the event held, what was in the event, where and when was the event held, who held the event, and how was the event held?

In general, competency Questions (CQs) as well as descriptions of reasoning tasks are used for setting up the validation rules and cases, however, such approaches often lack focus on how to verify whether requirements are properly met [3][5]. CQs are natural language questions outlining and constraining the scope of knowledge represented by an ontology [10][13] and may consist of query questions [6] that are used both to lay down requirements as well as to test the ontology. However, development of appropriate and answerable CQs has been shown to be difficult and time-consuming, due to, among others, manual authoring, relevance, answerability, and re-usability [7]. To test domain ontologies, as an alternative, Nasution [8] proposes that the ontology must contain answers to fact seeking 5WH questions as the extract in Figure 1 shows. Likewise, to validate whether the information retrieved presents basic facts about a concept, this research proposes to use 5WH questions as competency questions (CQs). Such CQs would help validate whether the information retrieved brings out basic facts.

The study [8], proposes basic 5WH CQs for evaluating ontology information on domains shown in Figure 1. These are simply directed fact finding questions having what, who, where, when, why and how on a chosen topic. In other fields, the goal of the questioning technique is to gain factual answers to each 5WH question. It is expected that answers to all six questions should give clarity to what a questioner is trying to discover, for example, the solution to a problem, the answer to a mystery, or even the best way to build a product [9]. Overall, directed 5WH questions help bring out essential fundamentals of a topic. As shown in Figure 1, ontological descriptions should contain answers about what, who, where, when, why and how, that is, 5WH. Drawing from this, the present research proposes the use of 5WH as CQs to compare sample domain ontologies available for Al-Quran as a test of the methodology and then to validate information retrieved on the basis of concept-based common text-patterns found in Al-Quran.

As test cases, Table I shows designed basic fact-finding 5WH CQs for the Al-Quran domains of *Salat*, *Zakat*, and the Sun. Rather than using keywords of *Salat*, *Zakat* and the Sun, this research uses repeating text-patterns (multiple word sequences) found in Al-Quran, in the same domain areas, as the basis of information retrieval. The 5WH CQs would then need to be designed on the whole text-pattern concept itself. For example, 5WH CQs may be designed on the text pattern “the sun and the moon each running for a term”. So, a basic fact-finding 5WH question may be why do the sun and the moon run for a term? The sample test cases used are popular concepts for which guidance is sought from Al-Quran. The basis of selection of the retrieved *Ayats* would be common text-patterns that contain the concept under study.

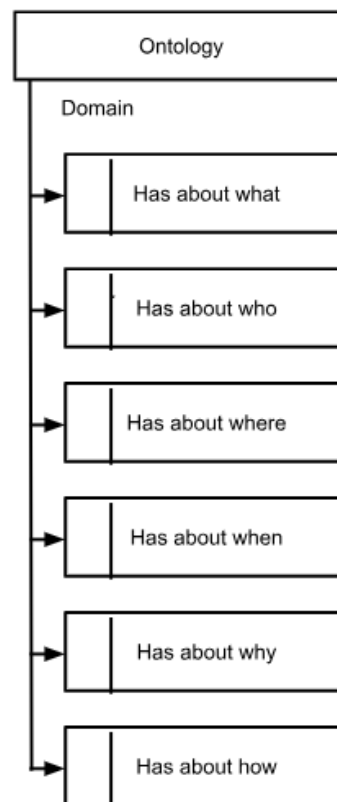


Figure 1. 5WH Propositions for Ontology Domain [8].

As a first step, this research proposes to test the proposed 5WH CQ methodology. It is important to understand that if a user searches directly from an index the possibility of error in retrieved results reduces. For example, searching from the back of the book index. As the index is designed from the text itself, both precision and recall are likely to be high. However, a retrieved list of sentences with an indexed concept as the keyword may contain the concept present either as the subject or the objective of the sentence. When a user is looking for the properties of the concept, sentences where the concept appears as the subject may be relevant while sentences where the concept appears as the object may be irrelevant. For example, if the subject “Sun” (شمس) is looked up for its properties in the Quranic ontology of Dukes [14] 32 *Ayats* are listed where 19 *Ayats* carry “Sun” as the subject and 13 *Ayats* where “Sun” is presented as an object. If these 13 *Ayats* are considered irrelevant, then the list contains $13/32 = 40\%$ irrelevant *Ayats*.

Precision and recall methods of validation for Information Retrieval are unable to directly solve the issue of relevance therefore, an alternative method involving fact-based 5WH CQs is proposed.

Current methods for validating information retrieval systems, particularly for religious texts like Al-Quran, often require domain expertise and lack a standardized, scalable framework for validation. This paper introduces a user-friendly approach based on 5WH questions, which not only reduces the reliance on domain experts but also improves the accessibility of IR systems for general users.

Previous works [3][5] have highlighted challenges in automating competency question generation and validation in ontology learning. These studies have mainly focused on expert-driven models, which limit scalability. By using the 5WH methodology, our approach simplifies the generation of competency questions, allowing for a more systematic validation process that is scalable.

II. VALIDATION USING BASIC 5WH ONTOLOGY COMPETENCY QUESTIONS (CQS)

Existing ontologies of three domains, that is, Islamic prayer or *Salat*, charity or *Zakat* and the Sun, are chosen for validating the proposed methodology of 5WH CQs as crafted in Table I. These are subsequently used as 5WH CQs for validating text-pattern based IR.

As recommended by the study [8], it is observed that these CQs are fundamental to the nature and purpose of each domain, for example, why should *Salat* be offered? Or why give *Zakat*? Two of Al-Quran domain ontologies examined in this paper make use of their own expert designed CQs to specify, guide, and test the output.

The expert designed CQs are shown in Table II in the central column together with proposed sample fact finding

CQs. The ontology domains include *Zakat* [10], *Salat* ([11], [13]), and the Sun [14]. Two deal with pillars of the religion of Islam and one with a Sign (miracle of creation) of Allah.

As the CQs of these ontologies ([10], [11], [13], [14]) have already been published, it is proposed to contrast these with the proposed 5WH CQs on these three domains. Table I shows sample proposed 5WH CQs for these three domain ontologies. To test the proposed *Zakat* ontology [10], the researchers designed CQs to validate the relevance of the various branches of the ontology they proposed. As Hadith was also considered along with Al-Quran for this work [10], the competency of Al-Quran Ayats retrieved was not tested separately. As shown in Table II, the CQs [10] for testing the *Zakat* ontology appear to be more in-depth, for example, such questions can be asked only by someone who already knows about *Zakat* and the history of Islam, e.g., was *Zakat* implemented before the migration of Prophet Muhammad (SAW)? as shown in the second column of Table II. The proposed 5WH CQs based on the T-box ontology model [8] are independent of prior knowledge and include the fundamentally important “why” question, for example: why is *Zakat* needed? Similarly, the competency questions [13] on *Salat* domain are from persons who already know about *Salat*, as opposed to basic questions like why is *Salat* needed or what is *Salat*? (Column 3, Table II). The processes that use CQs both as part of development and testing could be said to be analogous to “Teach to the Test” philosophy and thereby show a reasonable degree of accuracy. The research [14] on the other hand claims creating an ontology on Astronomical Bodies just by listing *Ayats* on the Sun. The methodology of 5WH CQs [8] does not require prior knowledge of the subject matter and can be easily adopted on any topic to test and validate whether an ontology provides fundamental information on a domain. The proposed fact-finding 5WH CQs were validated on 4 published domain ontologies. After checking, it is proposed that the 5WH CQs can also be used for validating text-pattern based Information Retrieval from Al-Quran.

TABLE I. 5WH CQS FOR TESTING OF THREE PUBLISHED ONTOLOGY DOMAINS

5WH	<i>Zakat</i> Domain	<i>Salat</i> Domain	Sun Domain
Why	Why give <i>Zakat</i> ? Why is <i>Zakat</i> needed?	Why should <i>Salat</i> be offered?	Why was the Sun created?
What	What is <i>Zakat</i> ?	What is <i>Salat</i> ?	What is the Sun?
How	How is <i>Zakat</i> given?	How is <i>Salat</i> offered?	How does the Sun work?
When	When should <i>Zakat</i> be given?	When should <i>Salat</i> be offered?	When does the Sun appear?
Where	Where should <i>Zakat</i> be given?	Where should <i>Salat</i> be offered?	Where is the Sun?
Who	Who should give <i>Zakat</i> ? Who should receive <i>Zakat</i> ?	Who should offer <i>Salat</i> ?	Who created the Sun? Who benefits?

TABLE II. COMPARISON OF CQS/VALIDATION METHODS FOR AL-QURAN ONTOLOGY-LEARNING

Al-Quran Domain Ontology	CQs/other Methods Used by the Researchers to Check the Validity of the Resulting Ontology	Proposed Sample 5WH CQs for Domain Ontology- Testing
Zakat Ontology [10]	Who deserves <i>Zakat</i> ? Was <i>Zakat</i> provision allowed in ancient civilization? Was <i>Zakat</i> implemented before the migration of Prophet Muhammad (SAW)? What are the conditions of <i>Zakat</i> ? What are the benefits of <i>Zakat</i> and warning for deniers? [10] <i>Salat</i>	Why is <i>Zakat</i> needed? What is <i>Zakat</i> ? How should <i>Zakat</i> be given? Where is <i>Zakat</i> applicable? When should <i>Zakat</i> be given? Who should give <i>Zakat</i> ? (CQs proposed by this research)
<i>Salat</i> [11]	Expert validated [11] six themes of <i>Salat</i> .	Six themes: History, Synonyms, PartOf, Actor, Methods and Consequences on <i>Salat</i> .
<i>Salat</i> [13]	What are the different invocational prayers in Islam? What are the daily prayers that Muslims perform? What are the optional prayers that Muslims perform daily? What prayers are offered in congregation? What are the obligatory prayers that Muslims perform daily? What are the components of funeral prayer? What are the impermissible time for <i>Salah</i> ? What are the different schools of thought? Which prayers have <i>Khutbah</i> ?	What is <i>Salat</i> ? Why is needed? How should <i>Salat</i> be offered? When should <i>Salat</i> be offered? Where should <i>Salat</i> be offered? Who should pray <i>Salat</i> ? (CQs proposed by this research)
Sun From Quranic Arabic Corpus [14]	The research [14] used the fundamental concepts from hadith and the tafsir (Quranic exegesis) of Ibn Kathir to build this concept ontology. A portion of Astronomical Bodies concept ontology is taken here for comparison. As this ontology uses expert knowledge from the beginning it is considered to be validated by experts. Selecting the Sun in the ontology produces an unclassified list of 33 Ayats that have both the Sun as the subject as well as the object of the Ayats.	Why was the Sun created? What is the Sun? What is the purpose of the Sun? How was the Sun created? Where is the Sun? When may the Sun burn out? Who created the Sun? (CQs proposed by this research)

III. TESTING THE 5WH CQ METHODOLOGY ON AYATS RETRIEVED ON THE BASIS OF CONCEPT-BASED REPEATING TEXT-PATTERNS FOUND IN AL-QURAN

Word sequences in Al-Quran are used precisely [15], for example, the context and use of Heavens and the Earth (السموات والأرض) is different from the Earth and the Heavens (الأرض والسموات). The study [16] therefore does a text-pattern study of up to 40 sequential words to create a dictionary of repeating patterns from the first 5 chapters of Al-Quran. Following up on text-patterns, proposed research shows that *Ayats* with larger text-patterns, for example, *Ayats* with the same 5 sequential words (or 5-grams) help to narrow down the precise context in which Al-Quran words are used. For example, 3 *Ayats* in the Arabic Al-Quran 13:2, 35:13 and 39:5 are found with the five-gram, “the Sun and the Moon each running for a term” (الشمس والقمر كل يجري لأجل) that contain the concept of the Sun. These 3 *Ayats* are shown in Table III along with their concordance on the two sides of each *Ayat*. The possible phrases in which the answers to the 5WH CQs may be found are underlined on the two sides in Table III.

A. Formulating 5WH CQs to Validate Retrieved *Ayats* based on Text-Patterns

Once the *Ayats* to which the specific text-patterns belong are extracted, the fact-finding 5WH questions are applied one by one systematically. For example, in the case of the “Sun” domain, the *Ayats* containing the text-pattern “the Sun and the Moon each running for a term” are retrieved and their relevance tested by applying 5WH CQs such as, “Why do the Sun and the Moon run for a term?”, “What does running for a term mean?” etc. The *Ayats* retrieved are then validated to ensure relevance to the text-pattern subject matter.

As shown in Table III, in answer to the CQ *why* do the Sun and the Moon run for a term, each of the three *Ayats* carry the word “appointed” (مسمى). The next word “term” implies a limited time. This chosen text-pattern on the Sun consists of six repeating words. Continuing with the question, “why”, the remaining phrase of *Ayat* 13:2 explains, “so that you believe with certainty”, on the “when” question “in the meeting with your Lord!” Similarly, the underlined phrases in the other two *Ayats*, shown in Table III, continue to add specific contexts on who question – “such, is Allah”, “your Lord”, on How question “is He not exalted in power?” (35:13, 39:5) Mapping all the phrases in the three *Ayats* against CQs, Figure 2 has been drawn up. Thus it is observed that extracting phrases on the basis of 5WH CQs may help support ontology learning.

TABLE III. CONCORDANCE OF AYATS WITH FIVE-GRAMS TEXT-PATTERN ON “SUN

Ayat no.	Part of Ayat after the Common five- gram	Common five- gram text-pattern and corresponding 5WH CQs	Part of Ayat leading to the Common five- gram
13:2	<p>مسمى يدبر الأمر . يفصل الآيات لعلمك بقاء ربكم تؤمنون</p> <p>appointed. He doth regulate all affairs, explaining the signs in detail, that ye may believe with certainty in the meeting with your Lord.</p>	<p>الشمس والقمر كل يجري لأجل</p> <p>the sun and the moon each running for a term</p> <p>5WH Qs on the 5-gram text-pattern can be:</p> <ul style="list-style-type: none"> • Why do the Sun and the Moon run for a term? • What does running for a term mean? • What would happen at the end of the term? • How do the sun and moon run? • When will the term end? • Where do the sun and moon run to? • Who makes the sun and the moon run? 	<p>الله الذي رفع السماوات بغير عمد ترونها ثم استوى على العرش وسخر</p> <p>Allah is He Who raised the heavens without any pillars that ye can see; is firmly established on the throne (of authority); He has subjected</p>
35:13	<p>مسمى ذلكم الله ربكم له الملك والذين تدعون من دونه ما يملكون من قسطير</p> <p>appointed. Such is Allah your Lord: to Him belongs all Dominion.</p>		<p>يولج الليل في النهار ويولج النهار في الليل وسخر</p> <p>He merges Night into Day, and he merges Day into Night, and he has subjected</p>
39:5	<p>مسمى ألا هو العزيز الغفار</p> <p>appointed. Is not He the Exalted in Power - He Who forgives again and again?</p> <p>(The Arabic Quran is collected from Tanzil.net [18])</p>		<p>خلق السماوات والأرض بالحق يكور الليل على النهار ويكور النهار على الليل وسخر</p> <p>He created the heavens and the earth in true (proportions): He makes the Night overlap the Day, and the Day overlap the Night: He has subjected</p> <p>(Translation used from Abdullah Yusuf Ali [17])</p>

By comparison, in the Quranic Arabic Corpus [14] website the option "Ontology of Concepts" contains ontology "Concept Maps". In "Concept Map" the root entity "Concept" is divided into 12 sub- categories, for example, "Artifact", "Astronomical Body", etc. Among them, "Astronomical Body", has the 6 sub concepts "Sun", "Moon", "Earth", etc., is given. When "Sun" (شمس) is selected, 33 Ayats are listed, as an undivided long list, containing the concept of "Sun" from the Quran corpus, where one of the Ayats is repeated twice, bringing a unique total of 32. Such a listing is essentially a keyword based listing, and does not support context-specific IR from Al-Quran. Compared to the concept map given in [14], the 5WH CQs provide better support for ontology learning. Figure 2 contains phrases that potentially answer some of the 5WH questions generated based on text patterns. 5WH are used to test the relevance of Verses/Ayats retrieved on the basis of common text-patterns found in Al-Quran. This method evaluates the relevance and usability of IR systems. Through this method Ayats are validated by checking whether one or more of the 5WH questions are answered.

B. Comparing Existing Salat Ontology CQs with 5WH CQs

A number of researchers have attempted to build ontologies based on the domain of Salat - it being one of the main pillars of the religion of Islam. For example, extracted terms, concepts and relations using rules [12] to identify patterns inside the text [11] for relation extraction from Al- Quran to derive the ontology of Salat. Authors in [13] additionally used test-driven ontology development methodology using Competency Questions (as shown in Table II) to derive the Salat ontology.

Their information sources in addition to Al-Quran included, for example, Hadith and Sunnah.

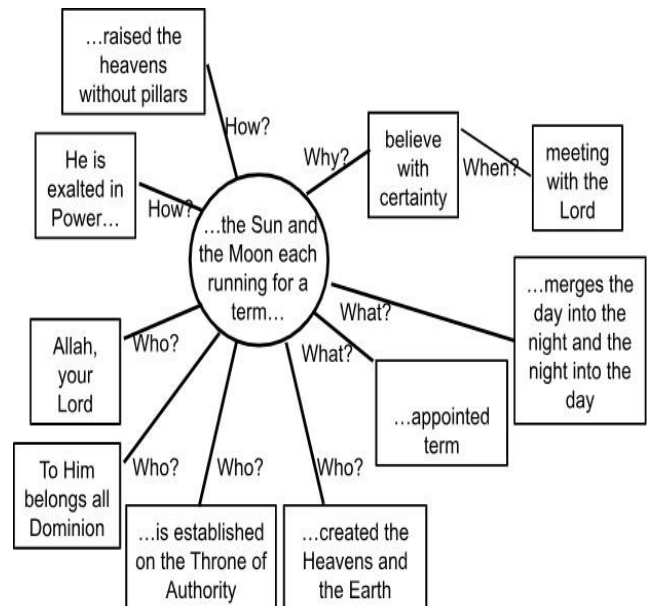


Figure 2. Sample learning phrases in which answers to basic 5WH questions can be found for ontology-learning related to the Sun.

Authors in [11] extracted Salat concepts using grammar patterns published earlier [12] to create the ontology shown in Figure 3. The Salat Ontology [11] is classified into six expert given themes: History, Synonyms, PartOf, Actor, Methods and Consequences. For the purposes of comparing with the

output [11] shown in Figure 3, *Iqamatas Salat* was chosen from the theme “Synonym” as a text-pattern. *Ayats* from Al-Quran were retrieved using this particular bi-gram text-pattern.

Three *Ayats* 20:14, 29:45 and 11:114 were found with this particular text-pattern as shown in Table IV. Possible 5WH CQs on the bi-gram *Iqamatas Salat* (and establish prayer) are given in the middle column of Table IV. Phrases, found on the two sides of the text-pattern (bi-gram) in the *Ayats*, that may answer the CQs, are underlined. Rather than using themes to classify the contexts found, which may require inputs from domain experts, this research takes a look at what type of IR support can be provided by using 5WH CQs on the concept. The concordance of the *Ayats* (Table IV) after the bi-gram mainly answers the question, *why establish prayer?* Three answers “for my remembrance”, “prayer prohibits immorality and wrongdoing” and “good removes evil” clearly come out as underlined. The leading parts of the *Ayats* shown on the right of the Table IV also answer the basic *why* question: “no deity except me” and “to recite the Book” revealed through Muhammad (pbuh).

So to answer a *why*-question on the bi-gram text-pattern *وأقم الصلاة* (And establish prayer) the phrases → “to remember Allah”, “to worship Allah alone”, “to help prevent immorality and wrongdoing” and “as a reminder” phrases from the *Ayats* (Table IV) may be used. The structured way in which *Ayats* retrieved based on a text-pattern brings out the answers to a basic 5WH question, in this case, a “*why*” question brings out yet another organizational beauty of Al-Quran, affirming its non-human origin. Text-patterns based on other forms of *Salat* found in Al-Quran (*والصلاة, بالصلاة, صلواتهم*) may help to answer other 5WH CQs.



Figure 3. Concept of relations in *salat/prayer* [11].

As IR based on text-patterns appears to help answer basic 5WH questions, the 5WH CQ methodology may lend support to the process of testing IR from Al-Quran. Authors in [13] on the other hand used CQs that require rudimentary knowledge of the subject as shown in Table II. Understandably, as the CQs were part of the initial design of what the ontology [13] should answer, the resulting ontology answered most of the CQs. However, the *Salat* ontologies ([11], [13]) do not answer the all important CQ based on the “*why*” question.

C. Comparing Zakat Ontology CQs with 5WH CQs

The study [10] also used CQs based on prior basic knowledge to develop their ontology on *Zakat*. For example: Was *Zakat* provision allowed in ancient civilization? Was *Zakat* implemented before the migration of Prophet Muhammad (SAW)? What are the conditions of *Zakat*? What are the benefits of *Zakat* and warning for deniers?

The *Zakat* ontology [10] was 100% accurate in 4 out of 5 competency questions as the competency questions were answered from Al-Quran, Hadith and other sources. These Competency Questions demand in-depth answers and are from the point-of-view of a person who already knows about *Zakat* rather than from a novice seeking guidance from Al-Quran point-of-view. This time, to compare with *Zakat* ontology [10], the stem (زكاة) was used to retrieve *Zakat* topics (الزكاة, الزكاة) based on text-patterns with “Charity/*Zakat*” from Al-Quran.

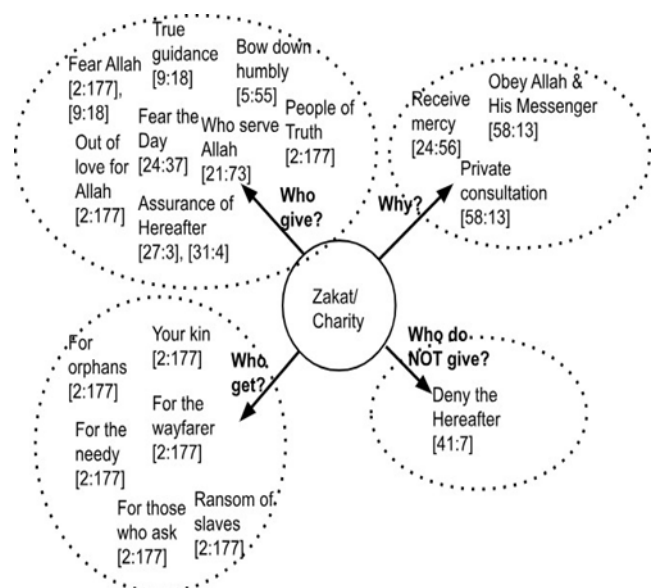


Figure 4. 5WH *Zakat* ontology learning from 4 selected text-patterns (Table V) and corresponding *Ayats* retrieved.

The total number of retrieved *Ayats* are 21 from two-word text-patterns related to the topics. For this purpose, *Ayats* were retrieved based on four different text-patterns shown in Table V to test against 5WH CQs. The *Ayats* are shown in Table V with a corresponding 5WH CQs concept map in Figure 4. It is observed that the 10 *Ayats* retrieved answer four questions, i.e., “*Why* give charity?”, “*Who* should give charity?”, “*Who* should get charity?” and “*Who* does not give charity?” Phrases from the 10 *Ayats* that potentially answer basic 5WH questions are underlined in Table V and reproduced in Figure 4.

It is also seen that each set of *Ayats* having the same text-pattern cluster consistently answer a specific type of 5WH question - another beauty of the structure of Al-Quran. By comparison, it is observed that the Competency Questions [10] do not contain the all important “*Why*” question as shown in column 2 in Table II. All told, it is observed that *Ayats* retrieved on the basis of text-patterns can potentially answer basic 5WH CQs to support IR from Al-Quran and may support ontology learning on basic facts.

TABLE IV. CONCORDANCE OF AYATS HAVING BI-GRAMS OF وَأَقِمِ الصَّلَاةَ (AND ESTABLISH PRAYER)

Ayat no.	Part of Ayat after the Common n-gram	Common bi-gram & CQs	Part of Ayat leading to the Common n- gram
20:14	لذكري for <u>My remembrance</u> (answer to why establish prayer?)		إني أنا الله لا إله إلا أنا فاعبدني Indeed, I am Allāh. There is <u>no deity except Me</u> , so worship Me (answer to why establish prayer?)
29:45	إن الصلاة تنهى عن الفحشاء والمنكر ولذكر الله أكبر والله يعلم ما تصنعون Indeed, prayer <u>prohibits immorality and wrongdoing</u> , and the remembrance of Allāh is greater. And Allāh knows that which you do. (answer to why establish prayer?)	(<i>Iqamatas Salat</i>) وأقم الصلاة and establish regular Prayers	اتل ما أوحى إليك من الكتاب <u>Recite</u> , [O Muḥammad], what has been revealed to you of <u>the Book</u> (answer to why establish prayer?)
11:114	طرفي النهار وزلفا من الليل إن الحسنات يذهبن السيئات ذلك ذكرى للذاكرين <u>at the two ends of the day and at the approaches of the night: For those things, that are good remove those that are evil:</u> Be that the <u>word of remembrance</u> to those who remember (their Lord): (answer to why establish prayer?) (also answers when to pray?)	CQs: Why should prayers be established? What are prayers? How to pray? When to pray? Where to pray? Who should pray?	Ayat starts with the bi-gram <i>Iqamatas Salat</i> , therefore there is no leading part

IV. RESULT

The comparison in Table II show that basic 5WH CQs outperform other competency questions in the ease of generation. To develop 5WH CQs there is no need for prior knowledge. 5WH CQs include a variety of WH questions with a single concept at a time for example, What is *Zakat*? Who should give *Zakat*? On the other hand, the research on *Zakat* ontology [10] developed questions with multiple concepts, for example, “Was *Zakat* provision allowed in ancient civilization?” This question has two concepts: one is permission of *Zakat*, and another is ancient times. To design this question the person would already have the the basic knowledge of *Zakat* and seeks the background of *Zakat*. 5WH CQs includes all sorts of WH questions, especially why questions applied to a single concept. For example, “Why should charity be given?” This question is included as one of the 5WH CQs in Table V, which is answered by Surah 24 *Ayat* 56 and Sura 58 *Ayat* 13. On the other hand, the “Why” question is not designed by the researchers mentioned in Table II on *Zakat* [10] and *Salat* [13] which shows fact-finding 5WH CQs better performance while testing *Ayats* retrieved on the basis of repeating text- patterns along with the ease of generation. Table VI shows the number of relevant *Ayats* retrieved using common text-patterns, i.e., text-patterns that repeat. The table shows the number of *Ayats* found that are relevant, i.e., where the text-patterns appears as a subject of the *Ayat*.

It is found that all such *Ayats* answer at least one of the 5WH questions, bring out the value of using text-patterns for information retrieved.

V. BENEFIT

In this research work, the proposed 5WH CQs methodology was tested against existing Al-Quran ontological outputs. It was also found that *Ayats* retrieved on the basis of text- patterns are in a position to answer one or more fact-finding 5WH questions. While the 5WH CQs were compared with the *Salat* domain ontology by Saad et al [12] as well as the retrieved *Ayats* based on a common text-pattern *Iqamatas Salat* differences were found in the “Why” *salat* question. Saad et al [12]’s domain ontology misses the fundamental “why” question although the text-pattern used for retrieval is the same. The important “Why” question is also answered by the retrieved *Ayats* based on 4 repeating text-patterns on the *Zakat* domain. However, Daud et al [10] *Zakat* ontology overlooks answering the fundamental “why” question. As a methodological approach it appears that fact-finding 5WH CQs can provide a useful measure to check IR output based on text-patterns in Al-Quran. Therefore, basic fact-finding 5WH questions may support information needs of general users to satisfy their understanding of basic facts on a topic. The method is scalable as 5WH CQs are easily generated. There is no need of expert help and no need of having prior knowledge of the subject. Anyone seeking basic knowledge about the topic can generate such fact-finding 5WH questions.

TABLE V. TWO-WORD TEXT-PATTERNS BASED ON ZAKAT (زكاة)

<i>Ayat Retrieved</i>	Keywords from <i>Ayat</i> context	Surah: <i>Ayat</i>	What types of 5WH questions are potentially answered ?
Based on text- pattern الزكاة وهم	Your (real) friends are (no less than) Allah, His Messenger, and the (fellowship of) believers,- those who establish regular prayers and regular charity , and they <u>bow down humbly</u> (in worship).	5:55	Who should give charity?
	Those who establish regular prayers and give in regular charity , and also <u>have (full) assurance of the Hereafter</u> .	27:3	Who should give charity?
	Those who establish regular Prayer, and give regular charity , and have (in their hearts) the <u>assurance of the Hereafter</u> .	31:4	Who should give charity?
	Those who practice not regular charity , and <u>who even deny the Hereafter</u> .	41:7	Who does not give charity?
Based on text- pattern وإيتاء الزكاة	And We made them leaders, guiding (men) by Our Command, and We sent them inspiration to do good deeds, to establish regular prayers, and to practice regular charity ; and they <u>constantly served Us</u> (and Us only).	21:73	Who should give charity?
	By men whom neither traffic nor merchandise can divert from the Remembrance of Allah, nor from regular Prayer, nor from the practice of regular charity : Their (only) <u>fear is for the Day</u> when hearts and eyes will be transformed (in a world wholly new),-	24:37	Who should give charity?
Based on text- pattern الزكاة وأطيعوا	So establish regular Prayer and give regular charity ; and obey the Messenger; <u>that ye may receive mercy</u> .	24:56	Why should charity be given?
	Is it that you are afraid of spending sums in charity before your <u>private consultation</u> (with him)? If, then, ye do not do so, and Allah forgives you, then (at least) establish regular prayer; practice regular charity ; and <u>obey Allah and His Messenger</u> . And Allah is well-acquainted with all that ye do.	58:13	Why should charity be given?
Based on text- pattern وأتى الزكاة	It is not righteousness that ye turn your face towards East or West; but it is righteousness - to believe in Allah and the Last Day, and the Angels, and the Book, and the Messengers; to spend of your substance, out of love for Him, for your kin, for orphans, for the needy, for the wayfarer, for those who ask, and for the ransom of slaves ; to be steadfast in prayer, and practice regular charity ; to fulfill the contracts which ye have made; and to be firm and patient, in pain (or suffering) and adversity, and throughout all periods of panic. Such are the <u>people of truth, the Allah- fearing</u> .	2:177	Who should get charity? Who should give charity?
	The masjids of Allah shall be visited and maintained by such as believe in Allah and the Last Day, establish regular prayers, and practice regular charity , and <u>fear none (at all) except Allah</u> . It is they who are expected to be on <u>true guidance</u> .	9 :18	Who should give charity?

TABLE VI. NUMBER OF RELEVANT AYATS FOUND USING COMMON TEXT-PATTERNS

Domain	Text-patterns	Number of Ayats found	Number of Ayats that answer 5WH CQs
Salat	واقم الصلاة and establish regular Prayers	3	3
	Text-patterns with Zakat الزكاة وهم وايتاء الزكاة الزكاة وأطيعوا وأتى الزكاة	10	10
Sun	الشمس والقمر كل يجري لأجل the sun and the moon each running for a term	3	3

VI. CONCLUSIONS

The proposed fact-finding 5WH CQs appear to work well with text-patterns based IR of the Al-Quran. The text-patterns of Al-Quran that repeat have therefore revealed a very interesting find. Ayats retrieved based on the same repeating text-pattern tend to answer individual basic/fundamental fact-finding 5WH Competency Questions. These CQs need not be created by a domain expert who has the benefit of prior knowledge of the domain. The fact that 5WH CQs are answered shows scalability of the method that may be used to test Information Retrieval from Al-Quran, where a repeating text-pattern is the subject. To further test the 5WH CQ methodology itself, the methodology was tested against existing Al-Quran ontologies. The test found that these ontologies mainly do not contain the answer to the basic “why” question. By structuring queries around user needs for fundamental information, a scalable method for improving ontology learning and enhancing the user experience in Quranic IR systems has been proposed. 5WH competency questions on the search concept helps to identify whether the information retrieved provides basic facts on the concept. Information retrieved on the basis of text-patterns has been shown to answer fact-finding 5WH Competency Questions. As such it is recommended that IR based on text-patterns be further studied. In conclusion, this research demonstrates the potential of using 5WH competency questions to validate the relevance of information retrieved from the Al-Quran.

VII. LIMITATIONS AND FUTURE WORK

The 5WH CQ method appears to work with text-pattern based information retrieval from Al-Quran. Finding the required text-patterns, however, would be difficult for a user. Therefore, there is a need to create an index of repeating text-patterns to solve the problem. The study by Oktaviani et al [16] only listed the longest repeating text-patterns from 5 Al-Quran chapters, confirming that a large number of repeating text-patterns exist in Al-Quran. A list of available text-patterns would have to be found in Al-Quran and be made available as an index. The user would then have to select from the list of repeating text-patterns available in Al-Quran.

The 5WH CQs validation method serves the need of general users rather than experts. Experts may also use this method of IR based on text-patterns to support ontology learning. While the 5WH method has proven effective for text-pattern based information retrieval, integrating this method with AI and natural language processing tools could enhance the scalability of Quranic information retrieval. Which of the 5WH questions is answered has to be decided by the user. Subsequently, Machine Learning may be applied to match user queries to relevant text-patterns.

NOMENCLATURE

CQ	Competency Questions
IR	Information Retrieval
5WH	5 WH questions (what, who, where, when, why and how)
n-gram	n sequential words
bi-gram	two sequential words as a text-pattern
Teaching to the test	A colloquial term for any method of education whose curriculum is heavily focused on preparing students for a standardized test.
Text-pattern	A fixed sequence of words is labelled as a text-pattern when the identical sequence appears in multiple places in Al-Quran.
T-box ontology model	Ontology testing model

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

ACKNOWLEDGEMENT

This research is part of continuing PhD work of Nazia Nishat bearing ID 4150171 at USIM. I would like to thank the USIM authorities for their support to do the PhD work.

REFERENCES

- [1] Carterette, Ben, and Ellen M. Voorhees. "Overview of information retrieval evaluation." In Current challenges in patent information retrieval, pp. 69-85. Springer, Berlin, Heidelberg, 2011.
- [2] Karlgren, Jussi. "Adopting systematic evaluation benchmarks in operational settings." Information Retrieval Evaluation in a Changing World: Lessons Learned from 20 Years of CLEF (2019): 583-590.
- [3] Blomqvist, Eva, Azam Seil Sepour, and Valentina Presutti. "Ontology testing-methodology and tool." In Knowledge Engineering and Knowledge Management: 18th International Conference, EKAW 2012, Galway City, Ireland, October 8-12, 2012. Proceedings 18, pp. 216-226. Springer Berlin Heidelberg, 2012.

- [4] Potoniec, Jędrzej, Dawid Wiśniewski, Agnieszka Ławrynowicz, and C. Maria Keet. "Dataset of ontology competency questions to SPARQL-OWL queries translations." *Data in brief* 29 (2020): 105098.
- [5] Bezerra, Camila, Fred Freitas, and Filipe Santana. "Evaluating ontologies with competency questions." In *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, vol. 3, pp. 284-285. IEEE, 2013.
- [6] Wiśniewski, Dawid, Jędrzej Potoniec, Agnieszka Ławrynowicz, and C. Maria Keet. "Analysis of ontology competency questions and their formalizations in SPARQL-OWL." *Journal of Web Semantics* 59 (2019): 100534.
- [7] Antia, Mary-Jane, and C. Maria Keet. "Automating the generation of competency questions for ontologies with AgOCQs." In *Iberoamerican Knowledge Graphs and Semantic Web Conference*, pp. 213-227. Cham: Springer Nature Switzerland, 2023.
- [8] Nasution, Mahyuddin KM. "Ontology." In *Journal of Physics: Conference Series*, vol. 1116, p. 022030. IOP Publishing, 2018.
- [9] "5 Ws and H," ProductPlan website, [Online]. Available: <https://www.productplan.com/glossary/5-ws-and-h/>.
- [10] Daud, Ali, Muhammad Hafeez Ullah, Ameen Reda Banjar, and Abdulrahman A. Alshdadi. "Ontological modeling and semantic search in quran." *IJCSNS* 22, no. 5 (2022): 771.
- [11] Saad, Saidah, Naomie Salim, Hakim Zainal, and S. Azman M. Noah. "A framework for Islamic knowledge via ontology representation." In *2010 International Conference on Information Retrieval & Knowledge Management (CAMP)*, pp. 310-314. IEEE, 2010.
- [12] Saad, Saidah, Naomie Salim, and Hakim Zainal. "Pattern extraction for Islamic concept." In *2009 International Conference on Electrical Engineering and Informatics*, vol. 2, pp. 333-337. IEEE, 2009.
- [13] Islam, Noman, Kashif Laeeq, Junaid Sheikh, Haseeb Ahmed, and Ghazala Shafi Sheikh. "Salaat Ontology: A domain ontology for modeling information related to prayers in Islam." *Indian Journal of Science and Technology* 12 (2019): 31.
- [14] K. Dukes, The Qur'anic Arabic Corpus. [Online]. Available: <https://corpus.quran.com/qurandictionary.jsp>.
- [15] Bentrucia, Rahima, Samir Zidat, and Farhi Marir. "Extracting semantic relations from the Quranic Arabic based on Arabic conjunctive patterns." *Journal of King Saud University-Computer and Information Sciences* 30, no. 3 (2018): 382-390.
- [16] Oktaviani, Devi, Moch Arif Bijaksana, and Ibnu Asror. "Building a Database of Recurring Text in the Quran and its Translation." *Procedia Computer Science* 157 (2019): 125-133.
- [17] Yusuf Ali, Abdullah. "The holy Qur'an: Text, translation and commentary." RSA, Islamic Propagation Centre International, Durban (1946).
- [18] "Tanzil Quran Search," Tanzil, [Online]. Available: <https://tanzil.net/#search/quran/%D8%A7%D8%A8%D9%86>.