Evaluation of Search Engines in Retrieving Relevant Document on Prophet Muhammad

O. J Hassan[1], Zainab Binti Abu Bakar [2]

[1] *Faculty of Computer and Information Technology, Al-Madinah International University, Malaysia, o.jhassan@gmail.com*
[2] *Faculty of Computer and Information Technology, Al-Madinah International University, Malaysia, zainab.abubakar@mediu.edu.my*

Received 09 May 2017; accepted 01 06 2017

Abstract

The main purpose of this experiment is to explore the concept information retrieval on search engines. Identifying the search engine that returns the most relevant and accurate results is sought which best satisfies a number of queries regarding the prophet Muhammad (SAW). The concept of IR tries to achieve is to obtain relevant required information resources from a collection of information resources and at the same time, returning results with minimal irrelevant documents. An example of an IR system which is prevalently used today are search engines. IR is the technology which is been used to power and make search engines as efficient as they are today. Over the decades, different companies have sprouted into existence developing their own search engine. With this, users have at their disposal different types of search engine, it is just a matter of preference that determines which search engine the users then intend to make use of. The concept of using search engines in retrieving documents has over the years gained a reputation by many and a reason for such rapid growth would be attributed to efforts and researches which have been made in order to make them very efficient. Amongst many of the things searched over the internet via the use of search engine are matters and questions which has to do with religion. Narrowing it down to the context of this dissertation, the person of Prophet Muhammad is the scope of this research. Many a times, individuals intending to have certain question answered about him resort to the search engine online. But a question which arises is that how do users then get to know what search engine of repute would best give in return relevant documents from site links. Absence of such knowledge leaves the users confronted with the myriad of websites which the search engines gives in return and as a result, leaving the users with so many relevant and irrelevant documents. And this translates to an uncontrolled dissemination of wrong information's about the prophet. During the course of the research, test collections about the prophet were explored. This experiment entailed trying to gather both relevant judgement documents as well as result documents gotten from some selected search engines. Then with these documents, some human and computational judgements were made to see examine the closeness of the search engine results in accurately answering the queries on the subject matter earlier stated. Although in writing this paper. The experiment is yet to be completed, so the initial three chapters would be discussed and then a summary explanation of the earlier stages of the experiment would be discussed and likewise a brief discussion of the expected results would be briefly discussed too.

*Key*words: information retrieval, search engines, relevant judgement

1. Introduction

Information retrieval (IR) identifies documents, from a larger collection, which are expected to be relevant or correlate with respect to some query. IR identifies documents which matches a query presented to the system and in return, the result may or may not contain the desired information. With the concept of information retrieval, it allows documents, web pages, online catalogs, records which are either structured or semi-structured to all be represented, stored, organized and easily accessible. By doing this, it paves way for information of interest for individuals to be made available with easy accessibility. IR's purpose is to obtain relevant required information resources from a collection of information resources and at the same time, returning results with minimal irrelevant documents. (Yates and Neto, 2011].

On a day-to-day basis, plethora of documents are made available over the Internet. Some of these available documents/information needs to be verified and vetted for its correctness due to reasons that they emanated from a source believed to be duly correct beyond doubts and as such, there is no room for these available documents to contradict their main sources of information from which it emanated from. An example of these document are the Quran and the Hadith that came from the Islamic world and is read by all Muslims and sometimes by people of other belief. (Shatnawi et al., 2010). The Quran is the believed to be the speech of God while the Hadith contains the sayings and approved actions of the prophet Muhammad (PBUH) which further explains the content of the Quran. The duos are important reference points for the Muslims in completely understanding the religion of Islam. Going beyond the realm of just making these resources printed on books, efforts have been made in making them available on the Internet via search engines and even on desktop/mobile applications with embedded search engines. Efforts have been made to ensure that the documents available over the Internet as well as in-offline dedicated search engines are authentic, relevant and efficient. Ongoing researches are still been made to improve the reliability and retrieval of such document.

Knowing about the prophet is also believed to be a topic much revered to the Muslims, which could be for various reasons. Some been to better understanding the religion, taking him as a role model, adopting a part of his life style or intimately knowing his natural behavior not necessarily stipulated by the religion. As such, it makes it appropriate and needful for making information about the prophet (PBUH) available on the Internet, which would answer possible queries of people via search engines. But much more necessary would be ensuring that these documents relating to the prophet (PBUH) are from sources considered genuine. For quite a number of the documents would be relevant to a query and vice-versa and there are tendencies that there would be documents that aren't of reliable content. As such, verifying content and sources of these type of documents is a necessary endeavor.

There are some available information retrieval systems or search engines for hadith and Quran collection, which are either online or application. But same does not apply for the biography/history of the prophet which implies also there are no repository sources which contains collated documents serving as answers for possible queries about the prophet (PBUH). Although some reputable Islamic websites provide search engines which allows users to search but these search engines aren't dedicated to only about the prophet (PBUH) implying more irrelevant feedbacks and excluding one from chances of exploring other sites which have the answers. So narrowing down reliable search engines is imminent. As of the point of this research, there is no available recommendation for the best search engine to use for the above stated reason or a referral to a sort of offline application dedicated for the same purpose.

Studies have shown that the usage of search engines have become the sole means for a number of people to fulfill their information needs and the use of these available search engines has become a part of peoples' daily routines (Kelly and Ruthven, 2010).

A number of researches have been made to ensure that the collection of documents related to the Quran as well as the hadith is readily available on the Internet coupled with efforts to ensure that reliable and relevant documents are presented after queries have been processed via search engines. Search engines like Google and Bing are popular sources from which relevant documents are gotten. Also, there are some websites that have within them search engines that makes querying the Quran and hadith possible.

The same is not the case for documents about the prophet (PBUH). There is no available means of clearly stating that results gotten from website links via search engines contain information which are deemed to be authentic or not. This shows the inability of users to know precisely which website to resort to for the purpose of retrieving relevant document about the prophet(PBUH). This in most cases leaves the user with a myriad of document to sift through without even knowing how relevant and accurate a document about the prophet is. Available in the internet space are a plethora of documents about the prophet(S.A.W)

Knowing the best search engine to use to query about the prophet(PBUH) also is a going to be an issue for the average user of the internet having in mind that search engines have become a primary means through which a myriad of people use the available search engines to get answers for their information need (Gan et al., 2008),it is essential to know which search engine would best provide these sought relevant documents.

Also is the issue that there are no known test collections relating to the said subject, which is about the prophet (PBUH). The absence of this collection of documents could be a hampering reason for dedicated search engines to not have been developed yet.

With inability to find available test collection documents and likewise the inability to firmly vouch the relevance of documents from a search engine's result, this research intends to find methods and means in evaluating search engines which best satisfies queries on the basis of providing relevant documents.

1.2. RESEACH QUESTIONS

1. What are the relevant documents on the prophet(PBUH).

2. What are the available search engines that best gives relevant results to a query about the prophet (PBUH)?

1.3. RESEARCH OBJECTIVES

The research will be conducted to find and address the following

1. To find out the level applied by the academic library to address IT strategic planning process in automation of information and communication technology at federal college of education, technical Gusau.

2. To find out the current challenges which are connected with IT strategic planning process in the library?

1.4. RESEARCH AIM

The research intends to know the best search engine that provides relevant results to possible queries about the prophet (PBUH).

1.5. SIGNIFICANCE OF THE RESEARCH

The intended significance of this research would be to suggest search engines that best give relevant documents based on evaluation made. This would make answers sought about the prophet (PBUH) to be easily retrieved with reduced chances of reasons to get worried about content authenticity. This would culminate in promoting people to adopt the way and lifestyle of the prophet

considering they can easily submit queries about his life with feedback of relevant documents.

2. Related Studies

2.1. INTRODUCTION

Obtaining useful and relevant document using search engines is one of the major tasks that search engines does. This is in line with the goal that information retrieval systems try to achieve. This chapter aims in expounding on the concept of information retrieval likewise the subject of search engines. In furtherance, some related work with regards to the concept information retrieval was discussed.

As far as knowledge management is concerned, information retrieval is a key technology that guarantees access to large repository of unstructured text. It is the basic technology behind web search engines and as such, it is a technology that is of great use to many web users. Information retrieval deals with the storage and representation of knowledge and the retrieval of information relevant for a special user problem. Information retrieval systems need to react to queries that are typically a few words of a natural language. The query is compared to document representations that were extracted during the indexing phase. The most similar documents are presented to the users who can evaluate the relevance with respect to their problem. (Tawileh, Mandl & Griesbaum, 2009).

IR addresses the retrieval of documents from an organized as well as relatively static repository, which is referred to as a "collection". These collections could come in different forms. They could be a stream of messages, e.g., E-mail messages, faxes, news dispatches, flowing over the Internet or some private network. (Greengrass, 2000). The term "information retrieval" was first coined in the year1950 but research in this area has been actively pursued for at least the last 100 years. Developed at the end of the 19th and the beginning of the 20th centuries, the first automatic retrieval systems used mechanical solutions to speed up lookup in library catalogues (Sanderson & Croft, 2012).

Information Retrieval (IR) is concerned with the representation, storage, access and retrieval of information. In the past, the approach of information retrieval was traditionally such that it was a manual process, mostly happening in the form of book lists in libraries, and in the books themselves, as tables of contents, other indices etc. These lists/tables usually contained a small number of index terms (e.g. title, author and perhaps a few subject headings) due to the tedious work of manually building and maintaining these indices. (Vester & martiny, 2005).

But with the advent of digital computer in the 20th century to this present day, this approach began to evolve and this was due to the fact that man was able to store, search and retrieve textual information. As a result, information retrieval has grown well beyond its previous limited form, mostly concerned with indexing and searching books and other kinds of textual information. The amount of information available in electronic form on the Web has grown exponentially. Almost any kind of desired information is present and made available over the internet. In furtherance, the amount of documents managed in organizational intranets that represent the accumulated knowledge of the organizations is also quickly growing, and efficient access to these documents is indeed of paramount need which plays a role in the success of modern organizations. (Vester, K. L., Martiny. M. C., 2005).

2.2. BASIC PROCESS OF INFORMATION RETRIEVAL

There are three processes an information retrieval system has to support: the representation of the document's content, the representation of the user's information need, and the comparison of the two representations. Below is a diagram depicting the process of information retrieval.

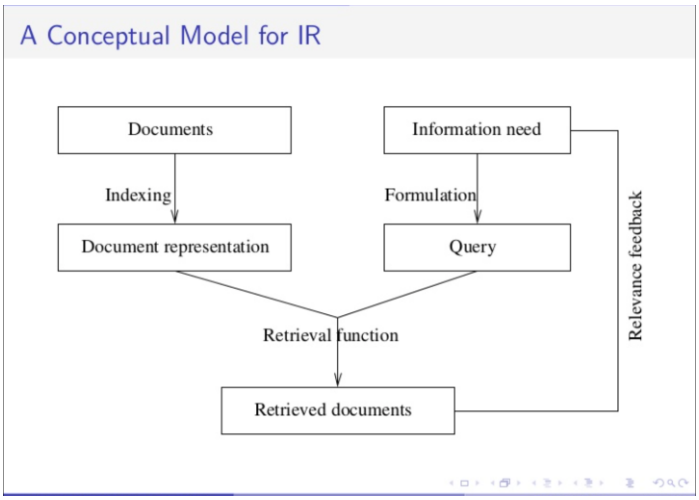

Figure 2.1 Information retrieval process (Lamas,2011)

Representing the documents is usually called the indexing process. The indexing process results in a formal representation of the document. The indexing process may include the actual storage of the document in the system, but often documents are only stored partly, for instance only title and abstract, plus information about the actual location of the document. The representation of the information problem or need is often referred to as the query formulation process. The resulting formal representation is the query. (Hiemstra, 2000).

2.3. SEARCH ENGINES

In it's initial stages, the term search engine was originally used to refer to specialized hardware for the purpose of text search. However, during the mid-1980s and onward, it gradually came to be used in preference to "information retrieval system" as the name for the software system that compares queries to documents and produces

ranked result lists of documents. (Croft, W. B., Metzler, D., & Strohman, T., 2015).

The origins of information retrieval can be traced way back during World War II as massive quantities of documentation and reports about weaponry were produced (Chu, 2003). Representation of these large collection of documents was indeed a very tedious task. The sheer size of the task has been described in Vannevar Bush's famous publication about the memex (Bush, 1945). With the growth of the Internet and the WWW which has led to the provision of huge volumes of available information, tools were devised to aid in the representation of these documents on the internet, but even at that, representation of these documents was a challenging task . (Kowalski, 1997) . Systems that provide these tools are usually known as search engines. A search engine is indeed a practical application of information retrieval techniques today.

Search engines can be used with small collections, such as a few hundred emails and documents on a desktop, or extremely large collections, such as the entire

Web. Search engines could come in different forms. Some of the ways in which they come for usage are as web search engines, Enterprise search engines, Desktop search engine, open source search engine and so on. (Croft, W. B., Metzler, D., & Strohman, T., 2015).

They all have some functionality in common but at the same time have their individual core functionality. There are a number of critical features of search engines that result from their deployment in large-scale, operational environments. Some among these features is the performance of the search engine in terms of measures such as response time, query throughput, and indexing speed. By response time, this means the delay encountered between submitting a query and receiving the result list, throughput measures the number of queries that can be processed in a given time, and indexing speed is the rate at which text documents can be transformed into indexes for searching. An index is a data structure that improves the speed of search. There may be only a few users of a given application, or many thousands. Scalability is clearly an important issue for search engine design. Designs that work for a given application should continue to work as the amount of data and the number of users grow.

Search engines have become a medium through which a plethora of people use in order to get relevant results for their information needs (Gan et al., 2008). And this trend is seen to be resorted to on a daily basis thus making this approach peoples' daily routines (Kelly & Ruthven, 2010). A trend amongst members of religious groups have been observed to frequently make use of the Internet to undertake activities, such as seeking advice and information, sharing experiences and opinions, listening to sermons, social networking and even shopping for religious merchandise (McKenna and West, 2007; Cheong et al., 2009; Campbell, 2005; Hoover et al., 2004; Helland, 2002; Foltz & Foltz, 2003; Casey, 2001; Larsen & Rainie, 2001).

Højsgaard and Warburg (2005) reported that by the year 2004 there were approximately 51 million religious websites on the Internet, disseminating information and communicating with followers. With all these fact and within the context of this research, ensuring that information sought out for from commercial search engines about the prophet should give and provide results which are undoubtedly accurate and relevant. For this is how the religion has been known to be preserved. That is by ensuring authenticity of content as well as narrator of content.

2.3.1 HOW WEB SEARCH ENGINES WORK

Undoubtedly, web search engines have evolved such that they have now become part and parcel of the Internet. They provide an easy and effective way of locating information on the World Wide Web (WWW), provided that the information is available to the search engines. A search engine operates in the order as stated below:

- Web crawling
- Indexing
- Search and retrieval

There is an automated computer program that browses/crawls the WWW in a systematic order and then retrieves web pages. Information about these web pages are stored by web search engines.

The main function of this robot is to store the information about the web pages it has visited for later processing. It uses two ways of navigating through a website which is either by following links on each visited web page or by making use of the Sitemap of the website if this is provided. Despite this method, it should be noted that this way of navigating the web limits how much of the web a robot can discover, in that if a webpage is not linked to from pages on the "known" web, the robot cannot discover it on its own. However, most search engines provides a way around this problem by offering a Sitemap service, where a site owner can submit a Sitemap of their website directly to the search engine.

There are search engines that usually allows the owner of a site to submit a link to his or her website with the intent requesting and allowing the robot to visit and index one or more pages on their site. Websites that have not been visited by the robot yet can use this method for possible reason like it been very new or because the robot missed it for some reason, or for pages that are already indexed by the search engine, but has received an important update which hasn't yet been reflected in the search engine results pages. As such, this could also help reduce the time it takes for a robot to discover a webpage and index it. (Aleksander. J., & Valla.O, May 2013).

Robots(also known as web crawler) use the User-agent field of a HTTP request to make themselves known to a web server. The moment a search query is submitted to the search engine, a comparison is made between the search

query to the index and retrieves the pages considered most relevant. A query processor is responsible for making this comparison. Search engines do make use of an algorithm that it uses for assigning importance to indexed items. One of the well-known algorithm is PageRank which is implemented by Google. Many search engines also offer users the ability to cut off or narrow the search. The resulting set, which generally contains links to all the resources that met the search specifications, is then presented in order of presumed relevance to the user. This set of results is often referred to as the search engine results pages (SERP).

Some copy of search pages also are stored by search engines. This is called caching. By caching, this allows the users to be able to either visit the original page or view the cached copy. An advantage the cache copy has is that in situations where the original page is temporarily down, the users can view the cached copy, and the download time is fast. But worthy of noting is that a down side to this is that the content of the cached copy might be obsolete.

2.3.2 EVALUATION OF SEARCH ENGINES

The most common methodology of information retrieval evaluation is based on the "traditional" evaluation approach of the Cranfield University first used in the 1960s(Weimar,2014,). Retrieving accurate and relevant information for users over the internet has become vital considering the large amount of diverse content existing on the internet. Evaluation of search engine performance has been under the spot light now for a while. (Leroy et al., 2007). In order for the deployment of search engines to thrive, employing some means of evaluating it is key. For with evaluation, possible issues as well as further improvements are realized and promptly addressed. The focal concerns raised during evaluation are the search engine's effectiveness as well as its efficiency. The first deals with how the search engine finds the right information, in other words the sequence of steps taken to achieve it and while the efficiency deals with how prompt the action of getting information is.

The evaluation of search engines paves way for the acquisition of new knowledge about existing search engines. With this new knowledge, it undoubtedly enhances the understanding of search engines by verifying assumptions that have been made and also discovering new aspects that have not been considered or thought through deeply. For a given query and also for a specific definition of relevance, effectiveness can be described as how well the ranking produced by the search engine corresponds to a ranking based on user relevance judgments. Efficiency is defined in terms of the time taken as well as the space requirements for the algorithm that produces the ranking. Viewed more generally, however, search is an interactive process involving different types of users with different information problems.

There are a couple of influencing factors/aspects, which would have to be considered in order to have a justifiable and accurate result. For these factors do influence the effectiveness of an IR system. Just to mention a few include:

- number of queries that would judge the retrieved results,
- selecting meaningful queries that serves as a form of user's information need,
- number of relevant documents the IR system can process and the list goes on. (Kirchhoff. L., 2010).

Over the years, after series of deliberations and evolution of the realm of search engines, proposed possible measures for the success of an IR system came into existence. Different experts had their own list of criteria's, but in general, evaluation categorization evolve around these namely –

- ✓ Relevance,
- ✓ User interface
- ✓ And system performance evolution.(Lewandowski 2007).

In an environment where the process of searching is involved with different types of users all with their individual information need (query), there are quite a number of factors that would affect the effectiveness and efficiency of search engines which could vary from the type of interface that was/is used for displaying search results and query refinement techniques, such as query suggestion and relevance feedback. So as such, evaluation is more typically done in tightly defined experimental settings. (Croft, W. B., Metzler, D., & Strohman, T., 2015).

When the deployment of either effectiveness or efficiency affects either of the other, such deployment technique would not be introduced. This fact does make effectiveness and efficiency to be related in their types of techniques. Generally speaking, however, information retrieval research focuses on improving the effectiveness of search, and when a technique has been established as being potentially useful, the focus shifts to finding efficient implementations. This is not to say that research on system architecture and efficiency is not important.

The focus on effectiveness is based on the underlying goal of a search engine, which is to find the relevant information. A search engine that is extremely fast is of no use unless it produces good results.

Emphasis of this research would be on the relevance. That is the relevant results retrieved that satisfies chosen queries about which constitutes information's about the prophet (PBUH).

2.3.3 A Search Engine Evaluation Method

An experiment held decades back which is referred to as Cranfield experiement paved way for the establishment of principles used in evaluating an IR system. Prior to evaluation, a blueprint of three requirements were required

before the evaluation could be achieved (Cleverdon, 1970). They are:

- a document collection (index)
- a set of information needs (queries)
- a set of relevance judgments is necessary (reasoning algorithm)

These requirements defined in the Cranfield experiment have been refined since then in numerous experiments and research efforts. The system evaluation experiment is conducted by using a corpus that consists of a document collection as well as information needs (queries) that are run against a document collection in batch mode. Relevance judgments of the results are created by experts, which list the documents that should be returned for each query.

2.3.4 Precisions And Recall

Precision and recall are two prominent techniques used in the measurement of an IR system. From the documents retrieved by an IR system, precision measures how many of those retrieved documents are relevant to a query, in other words that satisfies/answers an information need whereas recall measures how many relevant items are delivered by the IR system (Kowalski, 1997). In other words, the precision of a retrieval system for a certain query is the proportion of results that are relevant while recall is the proportion of relevant results that have been retrieved. (Prof. Wolfgang G. Stock,2012). These evaluation measures were first proposed as part of the Cranfield Project (Cleverdon and Keen 1968).

$$\text{Precision} = \frac{\text{relevant retrieved documents}}{\text{Overall retrieved results}}$$

$$\text{Recall} = \frac{\text{relevant retrieved documents}}{\text{Relevant results in database}}$$

An advantage which both the precision and recall evaluation measurement technique have is their permissibility for allowing evaluation of IR systems under different conditions in the sense that one measurement might appear be more important than the other in different circumstances. Siting an example would be in a scenario where a user of a search engine would like to retrieve all relevant results on the first page (high precision), but does not want to look at every relevant document. In contrast to that example also is say when a professional searcher might be very concerned about getting all relevant documents (high recall) and will most probably tolerate low precision. This means that precise measurements can be given for various information needs. (Lars Kirchhoff , 2010).

2.4 RELATED WORK

This section of this research tries to find out some of the related research that have previously been taken place which in one way or the other, has some sort of relationship with this present research. With this, it helps narrow down what the researcher should be focusing on. Articles, conference research papers as well as some books were resorted to in order to accomplish this task.

Abdul Rauf Saeed and Saeed and Syed Waqar Jaffry took out a research which commenced by first carrying out a comprehensive review of existing applications that offered different types of Quran and hadith information retrieval processes. Afterwards, a framework that is based on text mining technique was proposed. The framework uses stemming and related algorithms to create inverted index as well as term frequency and inverted document frequency. Methodology that was adopted included data collection and preprocessing, searching and ranking of results. This was demonstrated by developing an application that entailed a Quran and Hadith information retrieval framework. Culminating this research, users were tasked with evaluating this application that was aided with the use of a questioner. The result of the research was on the basis of the application's efficiency as well as it's usability. The whole intent of the research was to improve the text mining of Quran and Hadith information retrieval process and users were used as a means of evaluating the application.

Roslina Othman and Fauziah Abdul Wahid in their research paper tried to evaluate issues in evaluating the retrieval performance of the mutli-script indexing of translated texts of the Quran. The objective of the paper was to examine retrieval performance of these texts and evaluate the retrieval performance of the translations with high level of semantic which were available in two different scripts which are the Malay Rumi and Malay Jawi which were both built upon Pimpinan ar-rahman(malay translated version of the Quran), Indri and Jawi software. A goal of this research work was to solve issues with regards to handling test collections, involving parallel corpus in the area of cross language IR in the Muslim world. Queries were gotten from questions asked on newspapers by readers in both Malay Rumi and Malay Jawi scripts while test relevance judgment was made based on answers to questions from a panel of experts. Adopted techniques that were used, as measures were recall, precision and overlap. They were computed in both scripts original selection. These three paved way for issues to be identified, issues like standardized text, translation and transliteration, construction of queries.

3. Methodology

3.1 INTRODUCTION

As indicated in the past chapter, it's aim was to try to provide a general overview on the concept information retrieval and other facets related to it which would be pertinent to this research. A similar approach is adopted in this chapter, which is to make known the intended research methods, which entails the test collection chosen and evaluation process.

3.2 TEST COLLECTION

In this research, a number of varying requisites before commencing this project had to be sorted out first. These includes gathering of queries which would be used on the search engines, collating relevant judgments to those queries and selecting a number of preferred search engine which would be used for evaluation purpose.

3.2.1 QUERY SELECTION

Queries are the information need of users that is usually typed into search engines with expectation the user gets in return a set of results which hopefully best satisfies the information need/query. The mode of collection of queries used in this research was by two approaches.

Requesting from random users information needs of their preference with regards to the prophet. This was done via audio messaging on whatapp platform. The responses were transcribed.

Verbally engaging and asking users, requesting from them possible information needs they are likely ask assuming they had at their disposal a repository containing documents only about the prophet. The response was written down and later typed.

A total of 102 questions/queries was gotten in return. Some were repeated so a bit of query cleaning had to take place. This was by avoiding duplicated questions. After removing duplicates, 93 questions were gotten. Below is a table containing some sample queries that would be used as part of the evaluation processes.

3.2.2 RELEVANT JUDGMENT

From the list of queries already collected, ten queries would be chosen for the purpose of this research. During the query selection process, it would be a combination of both oft-repeated queries as well as queries which excites the interest to the researcher. Finding relevant documents serving as answers to this selected query is a required step of action. Sources used in obtaining relevant and accurate document for these queries would be two websites as well as a book which is in the form of a pdf file . The book been used acclaimed for itself a grand prize from the Muslim World League in Saudi Arabia which is a Pan Islamic NGO thus making the book known in the minds of Muslim scholars as well as amongst the Muslim commoner.

In ensuring that the site used for getting relevant and genuine answers to this queries, a site belonging to a well renowned teacher and public speaker of Islamic studies alluded to the that fact that in seeking sound authentic knowledge from the internet, one should tread this option of using the internet with caution noting to the fact that there are a lot of quack sites claiming to be genuine sources for Islamic content and information.(Bilal Philips Website). On his site, he mentioned some websites, which he believes to be genuine with regards to the Islamic information provided for user consumption. From the list of recommended sites, the researcher did select one website which would be used for obtaining relevant results to these queries. The site chosen also is a site, which the researcher had previously known and used for years. Worthy of mentioning also is a fact that this site which the researcher would use for obtaining accurate answers for the queries has within itself it's own built-in search engine. This would make querying on the website even seemingly efficient for the researcher due to the fact that the website would give in return many answers to just a single query. So with a hindsight knowledge of the already known site to the researcher which was recommended on the website of the scholar, the research felt it apt to go ahead choosing the site to him.

3.3 METHODS

3.3.1 Query Cleaning

After devising means of collating the queries, they are saved into a folder. It would be almost inevitable for some of these queries to not be repetitive in nature considering the diverse background of each individual from whom the queries were obtained from. Scrutiny of these queries would be done with the intent of seeking out for repetitions. From queries that seem to be repeated, the best comprehensive one would be left while excluding the other repetition form the list of query collection. This eliminating approach would be applied to all queries that are noticed to have a repetitive pattern. A new folder would be created to save this new collection of already vetted queries.

3.3.2 Evaluation Phases

From the collection of queries cleaned as mentioned previously, a total ten queries would be selected for the purpose of this evaluation. Ensuing would be finding relevant documents which best satisfies this queries. Sources used for the purpose of obtaining this accurate answers has already been stated earlier in this chapter which would be a book containing the biography of the prophet as well as two Islamic websites whose authenticity for relevant documents is backed with scholarly support and approval.

All relevant answers to each query would be retrieved and saved into files in a folder. This folder would contain only relevant documents that have within them answers to the ten queries respectively. This entails the first step of the experiment.

After the above mentioned initial experiment phase, employing the search engine to retrieve results for each of the ten selected queries ensues. Then first two pages of results for each of the search engine would be retrieved. A table would be used to keep track of the results. The results of each search engine would be tested against the collection of relevant judgment collection. By doing this, information about both relevant and irrelevant documents would be obtained. A table representation would be used to have a detailed organization of this process. This table entails information about each query, each search engine, and results retrieved by each search engine, either it been

relevant or irrelevant documents. A sample of the table is shown below.

Query	Relevant Document	Irrelevant Document	Search Engine
1.			
2.			
3			

Figure 3.4:Sample table to record results from search engines

From the table above, the precession and recall of the search engines would be calculated.

4. Descriptive Statistical Analysis

4.1 INTRODUCTION

The chapter aims at providing a structured discourse with regards to the experiment that was carried out. Personal observation during the experiment would be expounded upon and also there would be a discussion of the results obtained.

4.2 CHARACTERISTICS OF TEST COLLECTIONS

Been that the result for documents obtained are from the Internet alludes to the fact that it's more of representation is likely to come in different format representation since these documents came from different sources, in other words the documents were made available from different people from around the world. In the quest of answering these queries, not everyone chose to adopt the conventional was of using text alone. The reason for making mention of this is that it was expected that the document results answering queries would be a single page text formats that would be copied and saved for text processing purposes, which was not the case. Other forms of document representations were used. Below is a summary of the characteristics of the document results gotten from the search engines.

a. Texts documents
b. Google books format
c. Videos
d. Forum pages (containing both relevant and non-relevant documents at the same time).

4.3 PROCESS OF DOCUMENT COLLECTION FROM INDIVIDUAL SEARCH ENGINES

In gathering retrieved documents from the search engines, a list of ten queries were used for querying purposes against the search engines. Although the total query gotten during the inception of this research amounted to a total of one hundred and two queries (102). It was noticed though that there appeared to be some duplicated amongst these queries, and this necessitated the queries to be cleaned. Post cleaning the initial queries, the number total number dropped to ninety-three (93) as a result of removing duplicate queries. The total query gotten would be included in the appendix. Below is a table that shows the collection of queries that was made use of for retrieving documents from the respective search engines.

Query No	Query
1	What is the name of the Bedouin woman that nursed prophet muhammad in the Arabian desert?
2.	Upon what religion was prophet Muhammad before the revelation first came to him?
3.	At what age was Aisha when the prophet married her and why?
4.	Apart from Ramadan, which month did prophet Muhammad fast the most?
5.	Which sahabah divorced his wife and the prophet Muhammad married her and what is her name?
6.	What are the names of prophet Muhammad's children?
7.	What jobs did prophet Muhammad engage in before his prophethood?
8.	Why did prophet Muhammad marry so many wives?
9.	How many wives did prophet Muhammad have?
10.	What are the names of Prophet Muhammad's wives?

Table 4.2 Table showing the list of queries that were used in retrieving documents from the search engines

Each of the above query was searched on each search engine. The first twenty links, in other words the first two pages for each of these queries was used for experimental purposes. For each query that was searched on each search engine, the links from the result was saved in a table. Below shows a sample table for one of the queries.

No	Document Number	URL
1	Q1-Doc1	https://global.britannica.com/biography/Halimah-bint-Abi-Dhuayb
	Q1-Doc2	http://www.oocities.org/mutmainaa/quiz/quiz_qa.html

	Q1-Doc3	https://blogofthebeardedone.wordpress.com/2015/04/20/barnaby-rogerson-on-the-wet-nurse-of-the-prophet-muhammad/
	Q1-Doc4	http://www.musalla.org/Articles/Seerah/seerah6.htm
	Q1-Doc5	https://en.wikipedia.org/wiki/Aminah
	Q1-Doc6	https://en.wikipedia.org/wiki/Halimah_bint_Abi_Dhuayb
	Q1-Doc7	https://en.wikipedia.org/wiki/Muhammad_in_Mecca
	Q1-Doc8	http://www.lastprophet.info/4-the-prophet-s-childhood-and-adolescence
	Q1-Doc9	http://www.ummi.co.uk/the-prophets-saw-breastfeeding-experience/
	Q1-Doc10	http://www.soundvision.com/book/muhammad-prophets-childhood-the-infant

Table 4.3 A table showing links of retrieved result form one of the search engine

From the above table, this is the retrieved result for just one query, same would be repeated for the nine (9) other queries for the same search engine which then amount to a total of two hundred earlier mentioned. This process of querying and tabulating information's regarding the query is also applied on the other search engines.

After having the website URL to the retrieved documents from each search engines. A collection of the content of each of these results was retrieved. This was done by copying the content for each retrieved webpage and saving them. Folder were created for collected content of each search engines.

Citing an example of how this was achieved was by having a folder named "Google", within this folder, ten folders were created each for an individual query. For each of these folders, twenty word documents were retrieved from the search engine, the content for each webpage was saved individually within each of the word documents respectively together with their URL links as well. Considering that a lot of documents were going to be retrieved entailed that properly saving arranging the collection was key to taking out the experiment. At this juncture though, not all content of the webpage could be retrieved due to reason with regards to the representation mode which was used. As earlier stated, some webpage represented their information via videos, Google books,, forums which enabled copying content from such page to be impossible. So as such, saving only the url link to these webpages was the best only option that remained. For each search engine, a total of 200 documents were retrieved. A final table was created which depicts the relevance of documents retrieved from the search engine against the relevant judgment. This table would be discussed at a later section of this chapter.

4.3.1 EVALUATING RETRIEVED DOCUMENTS AGAINST RELEVANT JUDGMENT

The document retrieved for each search engine was a combination of both relevant and irrelevant documents. As a means of sorting through the documents in determining if contents of retrieved documents were in correlating context with the appropriate relevant judgment documents, two approaches were adopted in achieving this which were by using:

I. Human Judgment approach.

II. Computational judgment approach.

I. HUMAN JUDGMENT APPROACH:

After all necessary documents required for the experiment are in place, the relevant judgment document as well as the search engine results whose content were retrieved and saved in Microsoft word documents were involved in this section of the experiment. Using the ten queries that was earlier discussed, each of them have their relevant documents stored as earlier discussed. So for each query result from the search engines (which already has it's content stored in a word document), the results were visually compared against the appropriate document from the relevant judgment collection. A search engine result document is said to be relevant if it's content is of the same context with the content of the relevant judgment document. In other words, say for example if a search result for query one returned from Google search engine has its contents to be of the same context with a relevant judgment document called RJ001, then that query result is considered relevant.

Been that a visual comparison between relevant judgment documents against search engine results were used, that made this experiment to be of a human judgment approach. Although considering the possibility of human error which could be involved, trying to minimize and best have an efficient result was a matter of concern. Due to the just stated concern, a second approach was adopted . The intent here is to compliment the human judgment evaluation.

II. COMPUTATIONAL JUDGMENT APPROACH:

In this experiment, a program had to be developed with the intent of achieving a more precise accurate evaluation result. Varying functions were developed and implemented. Some of the implementations involved removing stop words, punctuations, keeping track of word frequency and ultimately for comparing documents gotten from the search engine against relevant judgment documents. The list of stop words used were gotten from a collection of stop words compiled and made available online by the university of Glasgow. Although worthy of mentioning is that there were other sites that had provided a collection of stop words, choosing which stop word lists from online would probably be a choice of preference backed by reliable and genuine information providers which one

seems deem fit. The language used for processing these documents was python. Below is a graphical representation that depicts a summary of the processes/steps involved in processing the documents involved.

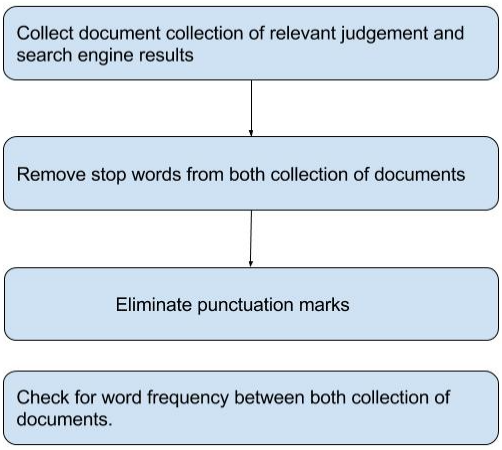

Figure 1 Graphical representation of steps taken in comparing documents between search engine and relevant judgment documents.

From the diagram above, it can be seen that three steps were involved in processing the documents. The initial two steps were implemented on both the relevant judgment documents and the documents from the search engine results. With the use of python, stop words were first removed from both collections of documents. This ensures that words that do not necessarily have a role to play in understanding the context of a explanation is removed leaving only words that helps in understanding concepts within any form of explanation. After removing stop words, punctuation marks were removed as well. The goal of the above two steps is to try to clean the two collection of documents to the minimum level. This aids in reducing anomalies and ensuring that the final results gotten are as accurate as possible. Having done the above two documents, the final steps is responsible for providing results which draws the experiment closer to finding results that provides information regarding the evaluation of the search engines stated earlier. After cleaning the text documents, the search engine documents would evaluated/compared against the relevant judgment documents. The frequency of word matching was used is comparing the documents. The more frequency of word appearance between the relevant judgment documents and the search engine documents, the more relevant these documents are considered to be.

5. PROPOSED EXPERIMENT AND RESULTS

As earlier stated, the experiment was yet to be completed during the writing of this experiment paper. The next sections would try to expound on phases that have begun and end this paper with a brief discussion on what the expected result would be like.

Prior to this stage, all retrieved documents from the search engines would be examined and then tabulated in order using the sample table below.

Query No	Relevant Docs	Irrelevant Docs	Recall	Precision

Table 5.1 A table showing summary results of retrieved documents

From the list of documents retrieved from the search engines, both relevant results as well as irrelevant results would be identified by comparing them against the collection of relevant judgment documents. The sample table above shows the structure of how the data would be inputted. The human judgment approach was employed in getting the information of this table. This table would be filled up in reference to all documents gotten from the search engines.

In reference to the formula for calculating recall, due to the inability of knowing the total amount of data available in the database of each search engine, having an accurate data for recall would not be possible using the formula stated earlier in the chapter.

Complimenting this seemingly flaw, further evaluation is carried out using computational approach. The python programming language is used to achieve this. With the use of python, stop words were first removed from both collections of documents. This ensures that words that do not necessarily have a role to play in understanding the context of a explanation is removed leaving only words that helps in understanding concepts within any form of explanation. After removing stop words, punctuation marks were removed as well. The goal of the above two steps is to try to clean the two collection of documents to the minimum level. This aids in reducing anomalies and ensuring that the final results gotten are as accurate as possible. Having done the above two documents, the final steps is responsible for providing results which draws the experiment closer to finding results that provides information regarding the evaluation of the search engines stated earlier. After cleaning the text documents, the search engine documents would evaluated/compared against the relevant judgment documents. The frequency of word matching was used is comparing the documents. The more frequency of word appearance between the relevant judgment documents and the search engine documents, the more relevant these documents are considered to be.

This computational approach was deployed on all documents returned from the search engines. The results gotten would then pave way for knowing the search engine that produces documents deemed relevant to the queries.

References

[1]. Log Rita Wan Chik, Mark Sanderson and Paul Clough , Investigating Religious Information Searching through the Analysis of a Search Engine log.

[2]. Dirk Lewandowski ,The Retrieval Effectiveness of Web Search Engines: Considering results Descriptions.

[3]. http://www.bilalphilips.com/islamic-websites/

[4]. Aleksander, J., & Olsen Valla, O., (2013,May)Privacy by Design: Adding Privacy toPublishing Platforms Joakim Aleksander Olsen Valla.

[5]. Ramaraj Palanisamy ,(2015) Evaluation of Search Engines: A Conceptual Model and Research Issues.

[6]. Zheng Zhu, (2011,May)Improving Search Engines via Classification.

[7]. Weimar,(2014,August). The Ideal User of a Search Engine Modeling and Evaluation Issues.

[8]. Ed Greengrass ,(2000,November)Information Retrieval: A Survey ,30 November.

[9]. Kenneth Lolk Vester, Moses Claus Martiny,(2005,August). Information retrieval in document spaces using clustering.

[10]. Thomas Mandl . Recent Developments in the Evaluation of Information Retrieval .Systems: Moving Toward Diversity and Practical Applications.

[11]. Kirchhoff.L,(2010) Applying Social Network Analysis to Information Retrieval on the .World Wide Web: A Case Study of Academic Publication Space.Croft,B.W., Metzler,D., & Trevor Strohman,T.(2015). Search Engines Information Retrieval in Practice, 2015.

[12]. Prof. Wolfgang G. S,(2012). On Search Engine Evaluation Metrics.

[13]. Othman,R., Wahid,F.A.,(2011) Issues in Evaluating the Retrieval Performance of .Multiscript Translation of Al-Quran.

[14]. Rauf S.A., Jaffry, S.W.,(2013) Information Mining from Muslim Scriptures

[15]. Lalmas,M.,(2011). Introduction to Information retrieval and Model, Retrieved from

[16]. Ricardo Baeza-Yates Berthier Ribeiro-Neto. (2011). Modern Information Retrieval, 2nd edition, 2011.

[17]. Stefan Bu¨ttcher, Charles L. A. Clarke, Gordon V. Cormack. (2010). Information Retrieval, Implementing and Evaluating Search Engines.

[18]. W. Bruce Croft, Donald Metzler and Trevor Strohman.(2015) Search Engines, Information Retrieval in Practice, 2015

[19]. Abdul Rauf Saeed and Syed Waqar Jaffry.(Information Mining from Muslim Scriptures, 2013).

[20]. Mohammed Q. Shatnawi ,Qusai Q. Abuein ,Omar Darwish, (2010). Verification Hadith Correctness in Islamic Web Pages Using Information Retrieval Techniques.

[21]. Rita Wan Chik, Mark Sanderson and Paul Clough (2008).Investigating Religious Information Searching through the Analysis of a Search Engine Log.